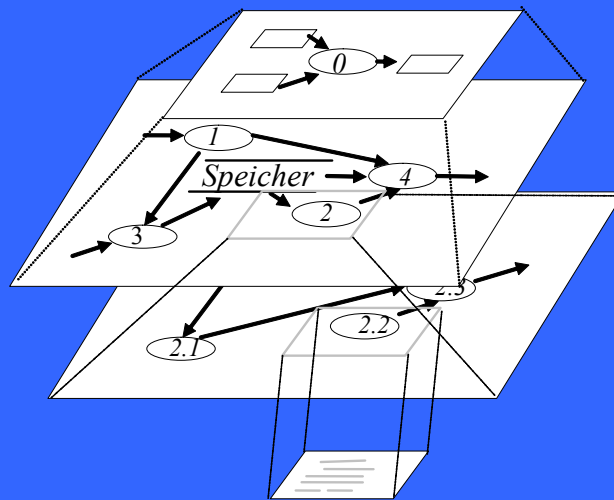


Adaptive Modellierung und Simulation

Rüdiger Brause



Rüdiger Brause

Adaptive Modellierung und Simulation

Grundlagen und Konzepte

Skript zur Vorlesung
SS 2015

© R.Brause, Goethe-Universität 2015

Prof. Dr. Rüdiger Brause
Adaptive Systemarchitektur
Institut für Informatik
J. W. Goethe-Universität Frankfurt am Main
60054 Frankfurt

Email: R_Brause@informatik.uni-frankfurt.de
Home: <http://www.informatik.uni-frankfurt.de/asa>
Folien: <http://www.informatik.uni-frankfurt.de/asa/ms/vorlesung>

(C) Johann Wolfgang Goethe-Universität,
Institut für Informatik

Dieses Skript ist nur für die interne Verwendung der Universität nach §52a Urheberrechtsgesetz bestimmt. Es enthält copyright-geschütztes Material und darf deshalb nicht ausserhalb der Hochschule benutzt werden. Eine Kopie ist untersagt.

Inhaltsverzeichnis

Inhaltsverzeichnis	VII
1 Einleitung	1-1
2 Black-Box-Modellierung	2-1
2.1 Input-Output-Tabellen	2-2
2.2 Lineare Modellierung	2-3
2.2.1 Parameterbestimmung: Verrauschte Messungen	2-3
2.2.2 Parameterbestimmung: Minimierung einer Zielfunktion	2-5
2.2.3 Parameterbestimmung: Multiple Regression	2-5
2.2.4 Parameterbestimmung: TLMSE	2-6
2.3 Die Grenzen des linearen Modells	2-10
2.3.1 Linearisierung	2-10
2.3.2 Parameterschätzung	2-11
2.3.3 Zur Rolle des Störterms	2-12
2.3.4 Multikollinearität	2-13
2.3.5 Testen von Einflüssen	2-14
2.4 Nichtlineare Modellierung mit Polynomen	2-17
2.5 Nichtlineare Modellierung mit Neuronalen Netzen	2-18
2.6 Backpropagation-Netze	2-20
2.6.1 Die Funktionsarchitektur	2-20
2.6.2 Lernen im Netzwerk	2-21
2.7 Training, Validierung, Testen	2-26
2.8 RBF-Netze	2-28
2.8.1 Approximation mit Glockenfunktionen	2-30
2.8.2 Normierung der Eingabevariablen	2-31
2.9 RBF-Lernverfahren	2-34
2.9.1 Anpassung der Parameter der ersten Schicht	2-34
2.9.2 Anpassen der Parameter der zweiten Schicht	2-38
2.10 Anwendungsbeispiel: Reaktormodell	2-39
2.11 Approximationsleistung neuronaler Netze	2-43

3	Wissensbasierte Modellierung	3-1
3.1	Einführung in dynamische Systeme	3-1
3.1.1	Modellierung mit Regeln	3-2
3.1.2	Modellierung des Textes	3-4
3.1.3	Stabilität	3-9
3.2	Dynamische Modellierung	3-15
3.2.1	Subsystem 1: Bevölkerungsentwicklung	3-15
3.2.2	Subsystem 2: Umweltbelastung	3-17
3.2.3	Subsystem 3: Konsumententwicklung	3-19
3.2.4	Das Gesamtsystem: Ein Mini-Weltmodell	3-20
3.2.5	Diskussion	3-24
3.3	Grundelemente dynamischer Modellierung	3-24
3.3.1	Zustandsnormierung	3-28
3.3.2	Systeme ohne Zustandsvariablen	3-29
3.3.3	Systeme mit einer Zustandsvariablen	3-30
3.3.4	Systeme mit zwei Zustandsvariablen	3-34
3.3.5	Systeme mit drei Zustandsvariablen	3-46
3.4	Gleichgewichtssysteme	3-47
3.5	Schätzung der Parameter	3-52
4	Hierarchische Systeme	4-1
4.1	Hierarchiebildung durch Aggregation	4-1
4.2	Hierarchiebildung durch lineare Multi-Level-Modellierung	4-6
4.3	Hierarchiebildung und Strukturanalyse	4-12
4.3.1	Graphenbasierte Analyse	4-13
4.3.2	Zeitverschobene Einflüsse	4-21
4.3.3	Reduktion der Komplexität von Modellen	4-22
5	Simulation	5-1
5.1	Deterministische Simulation	5-1
5.1.1	Analytische Modellierung vs. Simulation diskreter Systeme	5-1
5.1.2	Zeitkontinuierliche vs. zeitdiskrete Modellierung und Simulation	5-4
5.1.3	Ereignisorientierte Modellierung und Simulation	5-7
5.2	Stochastische Simulation	5-12
5.2.1	Eigenschaften von Zufallszahlen	5-12
5.2.2	Messen von Zufallszahlen	5-14
5.2.3	Erzeugung und Simulation von Zufallszahlen	5-16
5.2.4	Erzeugung einer Verteilung	5-19
5.2.5	Test der Zufallszahlen	5-29
5.3	Simulation mit Monte Carlo-Verfahren	5-32
5.3.1	<i>Hit-or-miss</i> Monte Carlo-Verfahren	5-33
5.3.2	<i>Crude</i> Monte Carlo-Verfahren	5-37

5.4	Der Simulationsrahmen	5-39
5.4.1	Problemidentifikation	5-39
5.4.2	Modellspezifikation und –detaillierung	5-40
5.4.3	Modellintegration realen Systemverhaltens	5-41
5.5	Auswertung der Simulationsergebnisse	5-46
5.5.1	Vergleich zweier Systeme	5-46
5.5.2	Antithetische Zufallsvariable	5-48
5.5.3	Synchronisierung der Zufallsvariablen	5-49
5.5.4	Auswertung mit Konfidenzintervallen	5-50
Literaturverzeichnis		1
Abbildungsverzeichnis		1
A Der Fixpunktalgorithmus für TMSE		A-1
Index		1

1 Einleitung

Unsere Welt wird immer komplexer und undurchschaubarer. Wie können wir damit umgehen? Eine plausible, oft genutzte Strategie besteht darin, das betrachtete System auf wenige, wichtige Wesenszüge zu reduzieren.: Wir machen ein Modell. Ähnlich einem Gebäudemodell in der Architektur, das nicht aus dem gleichen Material oder mit allen Details wie das geplante Gebäude bestehen muss, ist unser Modell nur eine Abstraktion des Systems, nicht das System selbst. Deshalb sind alle Verallgemeinerungen vom Modell zur Wirklichkeit mit Vorsicht zu betrachten. Trotzdem erhalten wir aber wichtige Vorteile durch die Modellbildung:

- Das Modell kann auch Bedingungen und Parameterwerten unterworfen werden, die in der Realität nur selten oder gar nicht vorkommen, etwa Erschütterungen (Erdbeben) bei Hochhausmodellen.
- Die Konsequenzen dieser beliebigen Parameter sind problemlos für das reale System. Damit sind Experimente am Modell möglich, die in der Realität zu riesigen materiellen Schäden führen würden, etwa den Einsturz des Hauses bei Erschütterungen.
- Der zeitliche Ablauf eines Modells kann wesentlich schneller oder langsamer als in der Realität vor sich gehen. Damit kann man das wesentliche Verhalten genauer untersuchen.
- Es ist wesentlich billiger, einfacher und schneller, Modelle und Simulationen einzurichten und zu ändern, als die Realität. Damit kann eine optimale Steuerung der Realität ausprobiert und ihre Konsequenzen erforscht werden, ohne dies in teuren und zeitaufwändigen Untersuchungen, Projekten und realen Aufbauten zu tun.

Wir alle führen ständig eine Modellierung der Realität durch, bereits wenn wir unseren Tag planen. Alle Aktionen, die wir tun (oder nicht tun) und ihre Konsequenzen versuchen wir, in eine sinnvolle Reihenfolge mit möglichst wenig Kosten zu bringen. Wir planen unsere Frühstückszeit, um den Bus zur Arbeit nicht zu verpassen, und planen unsere Fahrzeit, um nicht zu spät zur Arbeit zu kommen. Dabei modellieren wir Frühstück und Busfahrt jeweils nur bezüglich ihres Zeitrahmens; Details wie die Frage, ob wir vorn oder hinten sitzen im Bus, bleiben meist im Busmodell unberücksichtigt. Es stellt sich also nicht die Frage, ob man

modellieren soll oder nicht; wir alle tun dies täglich zu Hause und im Betrieb, sondern nur, wie.

Ein wichtiger Aspekt der Modellierung ist die Reduzierung der Information, die über das zu modellierende System existiert, auf wenige, wichtige Aspekte. Welche Aspekte sind wichtig? Diese Frage entscheidet sich bei der Fragestellung, die man mit dem Modell klären will: Die Anforderungen, der Sinn des Modells, bestimmen die Art des Modells; es gibt kein Modell ohne Nutzer.

Beispiel: Ozean

Wir können den Ozean verschieden modellieren, je nach Aufgabe. Wollen wir die beste Route für Segelschiffer finden, so modellieren wir ihn am besten mit Windmodellen und Strömungskarten. Ist er als Klimasystem interessant, so modellieren wir seine Wärmekapazität, seine Erwärmungszonen (wieviel Volumen wird in welcher Zeit auf welche Temperatur durch Sonneneinstrahlung erwärmt?) und seine Strömungen (Golfstrom!). Für die Ernährung ist dagegen sein Fischreichtum (Welche Arten, wo, wann), Algenbewuchs und Planktonkonzentration wichtig. Für Unterwasserfotografie sind nur Trübung, Farbänderung je nach Wassertiefe und Lichtbrechung im Medium „Wasser“ entscheidend.

Betrachten wir nun große Systeme wie unsere Volkswirtschaft, von der wir alle abhängen, so kann man ohne ein gutes Modell keine ernstzunehmenden Aussagen mehr treffen. Erst eine quantitative Modellierung ökonomischer Zusammenhänge gestattet es, die zum Teil gegenläufigen Effekte von wirtschaftspolitischen Maßnahmen auf ihre resultierenden Gesamtwirkung zu untersuchen. Alle Versuche, mit ideologischen Aussagen („Was gut ist für die Wirtschaft, ist auch gut für den Staat“) oder mit gewohnten Denkschablonen und gewohnten Rechenhilfen (Taschenrechner bei Tarifverhandlungen!?) die Entwicklung und Reaktion eines solchen komplexen Systems vorherzusagen, sind von Anfang an zum Scheitern verurteilt und führen teilweise zum Gegenteil von dem, was man „guten Willens“ beabsichtigt hatte.

Für eine erfolgreiche Modellierung eines komplexen Systems müssen wir berücksichtigen, wie viel wir von dem betrachteten System wissen. Dazu unterscheiden wir drei Fälle:

- Wir wissen nichts über die inneren Mechanismen des Systems. Dies bedeutet, dass wir es nur über das Eingabe-Ausgabeverhalten beschreiben können, siehe Abb. 1.1. Dies ist der Fall, in dem wir das System als „schwarzen Kasten“ (*black box*) mit unbekanntem Inneren betrachten. Wir können also nur eine Verhaltensmodellierung durchführen.
- Wenn wir im Gegensatz dazu alles wissen, woraus das System besteht wie bei der Mechanik einer Uhr, so können wir das Eingabe-Ausgabeverhalten sehr genau beschreiben. In diesem Fall sprechen wir von einer *white box*

oder *glass box*. Im Unterschied zur Verhaltensmodellierung benötigen wir wesentlich weniger Daten und können auch das dynamische Verhalten des Systems bei bisher unbekanntem Parameterwerten vorhersagen.

- Meist sind einige Zusammenhänge bekannt, aber nicht alle. In diesem Fall besteht das System aus bekannten und unbekanntem Komponenten. Für uns ist das System also nur halb durchsichtig (*opaque, grey box*). Wir können zwar von unserem Wissen profitieren, müssen aber unbekannte Subsysteme durch Verhaltensmodellierung approximieren. Dies ist der Regelfall.

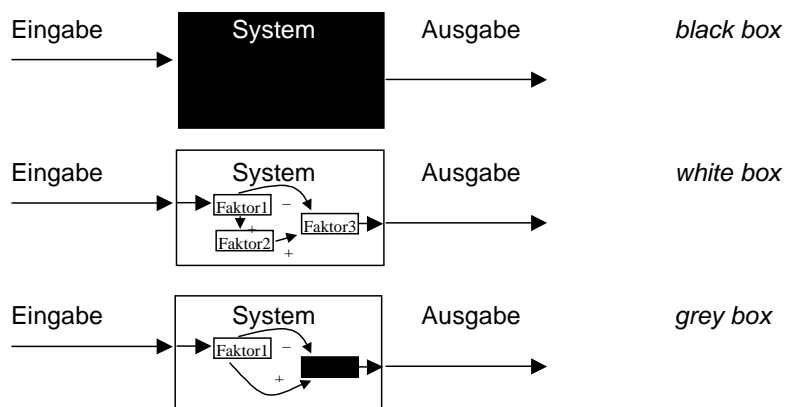


Abb. 1.1 *black box, white box und grey box* -Verhalten

In den folgenden Kapiteln wollen wir uns zunächst mit der Verhaltensmodellierung und dann mit der dynamischen Modellierung beschäftigen. Die Einbindung verschiedener Komponenten bei Mischmodellierung sowie die Techniken zur Simulation der Modelle schließen diesen Text ab.

Aus der großen Menge an möglichen Modellen möchten wir uns zunächst auf kontinuierliche Modelle mit beliebig vielen Zuständen beschränken. Zustandsorientierte Modelle wie endliche Automaten werden bereits in den Grundlagen der Informatik behandelt und sollen deshalb in diesem Text nicht weiter berücksichtigt werden.

2 Black-Box-Modellierung

Der einfachste Ansatz, die Welt von morgen zu beschreiben, besteht darin, einfach von der Welt von heute auszugehen und in Kenntnis aktueller Trends sie auf morgen zu fortzuschreiben. Dies ist die Aufgabe der Trendanalyse oder Vorhersagen (*forecasting*) wie sie gern von Börsenanalysten und Finanz“experten“ verwendet wird. Meist wird die Vorhersage mit Funktionen der Zeit (*Zeitreihen*) beschrieben, etwa bei Wechselkursen, Umweltverschmutzung oder Fieberkurven.

Es gibt sehr viele Arten der vorausschauenden Modellbildung, siehe Abb. 2.1.

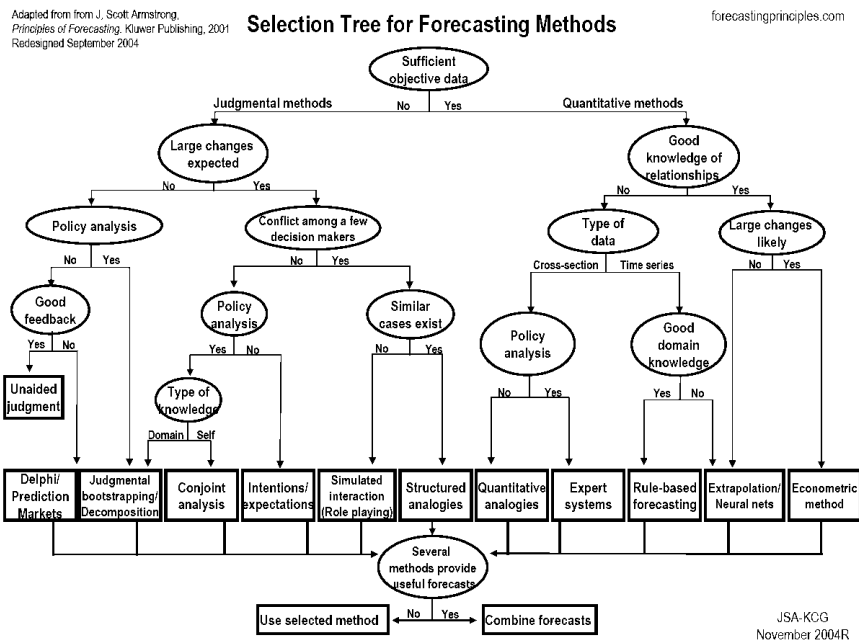


Abb. 2.1 Methoden vorausschauender Modellierung (nach [Arm01])

Haben wir zu wenige Messdaten für eine sichere Aussage, so sind alle Methoden von menschlichen Experten ebenfalls problematisch, da dann objektiv keine sichere Aussage möglich ist. Die linke Hälfte des Entscheidungsbaumes in Abb.

2.1 ist deshalb sehr kritisch zu beurteilen. Unterscheiden sich ihre Ergebnisse von reiner Kaffeesatzleserei, so sollte es auch möglich sein, die wenigen Daten mit mathematisch fundierten Methoden auszuwerten.

Wir möchten uns hier auf wenige Methoden für kontinuierliche Systeme beschränken, die sich auf eine Modellierung stützen, bei der die Parameter eines vorgegebenen Modells an beobachtete Messdaten angepasst werden. Die typische Modellbildung besteht darin, dass ein Modell gewählt wird, das möglichst einfach ist und doch noch die typischsten Verläufe der betrachteten Zeitreihen beschreiben kann.

2.1 Input-Output-Tabellen

Die einfachste Modellierung eines Systems besteht darin, alle Daten, die in das System hineingehen sowie alle Daten, die herauskommen, systematisch zu erfassen und in einer Tabelle zu auflisten. In Abb. 2.2 ist für die Darstellung des Zusammenhangs zwischen Konsumarten und benötigte Güterherstellung die Form einer Input-Output-Tabelle gezeigt.

		Konsum- verwendung			Saldo
		1	...	43	
Güter- gruppen	1				
	...				
	59				
Saldo					

Abb. 2.2 Makroökonom. Input-Output-Tabelle der Konsumverflechtung

Dabei ist die Tabelle für das Beispiel von 59 Arten von Produktionsgütern gezeigt, die den Input bilden für 43 Konsumkategorien. Eine solche Input-Output-Tabelle stellt immer nur einen Schnappschuss der Situation für einen Zeitpunkt dar, auch wenn sie über einen Zeitraum (etwa ein Jahr) gemittelt wurde. Zu einem späteren Zeitpunkt ist die Tabelle wieder anders; die Dynamik, die in einer echten Funktion enthalten ist, wird hier auf einen Zeitpunkt, eine Stützstelle der Input-Output-Funktionen beschränkt.

2.2 Lineare Modellierung

Ein wichtiger Schritt über die tabellarische Input-Output-Modellierung hinaus besteht in einer Generalisierung der Zusammenhänge durch eine Funktionsgleichung. Im einfachsten Fall nehmen wir einen linearen Zusammenhang zwischen der *exogenen*, ursächlichen Variable x und der *endogenen* Variable y an. Bei der Modellierung durch eine lineare Funktion

$$y = f(x) = y_0 + ax \quad (2.1)$$

besteht die Aufgabe darin, die Parameter a und y_0 bei gegebenen Daten zu finden. Ist $y_0 = 0$, so spricht man von einer *homogenen* Form, sonst von einer *inhomogenen* Form.

Beispiel Ökonomie

Nimmt man für den Zusammenhang zwischen Einkommen x und Konsum y ein Modell der Form (2.1) an, so liegt dem die Annahme zugrunde, dass ein Konsumsockel existiert, etwa durch Sozialfürsorge gesichert, und darüber hinaus zusätzliches Einkommen zu erhöhtem Konsum führt.

2.2.1 Parameterbestimmung: Verrauschte Messungen

Haben wir ein deterministisches Verhalten, so ist die Aufgabe einfach: Wir benötigen nur zwei Datenpunkte (x_1, y_1) und (x_2, y_2) und können die zwei Gleichungen mit zwei Unbekannten direkt lösen:

$$\begin{aligned} y_1 = y_0 + ax_1 &\Rightarrow (y_2 - y_1) = a(x_2 - x_1) \Rightarrow a = (y_2 - y_1) / (x_2 - x_1) \\ y_2 = y_0 + ax_2 &\Rightarrow y_0 = y_1 - ax_1 \end{aligned} \quad (2.2)$$

Analog können wir vorgehen, wenn \mathbf{x} und \mathbf{y} Vektoren sind und \mathbf{a} eine Matrix und \mathbf{y}_0 ein Vektor. Ein solches lineares Gleichungssystem mit n Komponenten (Dimension n) pro Vektor benötigt bei $m = (n^2 + n)$ Unbekannten *auch* m Gleichungen, also m Datenpunkte. Verfügen wir über so viele Messungen, so ist unser Problem gelöst.

Normalerweise reicht dies aber nicht aus, da die Realität nicht so eng deterministisch ist: Es gibt immer zufällige Abweichungen von dem deterministischen Wert. Ein typisches Gebiet dafür ist die Medizin: Messungen an Menschen, etwa Kraft-Winkel Messungen bei der Armbewegung, ergeben für denselben Menschen am selben Tag und mit dem selben Arm sehr unterschiedliche Ergebnisse. Wir können also nicht von deterministischen Daten ausgehen, sondern können versuchen, die Daten durch eine Gerade zu approximieren, siehe Abb. 2.3 .

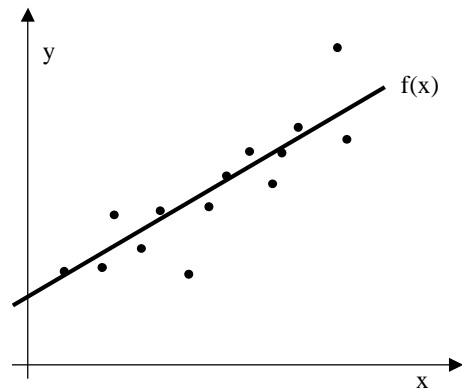


Abb. 2.3 Lineare Approximation von Daten

Wie groß sind nun die Parameter a und y_0 ? Angenommen, wir modellieren die zufälligen Abweichungen durch eine Zufallsvariable u , so dass für Gl. (2.1) die t -te Messung gilt

$$y_t = f(x_t) = y_0 + ax_t + u_t \quad (2.3)$$

Mit der Notation für den Erwartungswert

$$\bar{x} \equiv \frac{1}{T} \sum_{t=1}^T x_t \quad \text{Mittelwert} \quad = \langle x_t \rangle_t \quad \text{Erwartungswert} \quad (2.4)$$

und der Annahme, dass sich die Zufallsvariable u zu null mittelt und unkorreliert zu der Variablen x ist, erhalten wir

$$\begin{aligned} \frac{1}{T} \sum_{t=1}^T y_t &= \bar{y} = y_0 + a\bar{x}, & \langle u_t \rangle_t &= 0, \quad \langle u_t \bar{x} \rangle = \langle u_t \rangle \langle \bar{x} \rangle_t = 0 \\ \Rightarrow y_0 &= \bar{y} - a\bar{x} \end{aligned} \quad (2.5)$$

Außerdem gilt mit Gl. (2.3) und (2.4)

$$\begin{aligned} y_t - \bar{y} &= a(x_t - \bar{x}) + u_t \\ \Rightarrow (y_t - \bar{y})(x_t - \bar{x}) &= a(x_t - \bar{x})(x_t - \bar{x}) + u_t(x_t - \bar{x}) \\ \langle (y_t - \bar{y})(x_t - \bar{x}) \rangle_t &= a \langle (x_t - \bar{x})(x_t - \bar{x}) \rangle_t \end{aligned}$$

$$\text{Mit } \begin{cases} \text{cov}(x,y) \equiv \langle (y_t - \bar{y})(x_t - \bar{x}) \rangle_t & \text{Kovarianz} \\ \text{cov}(x,x) \equiv \text{var}(x) = \langle (x_t - \bar{x})^2 \rangle_t & \text{Varianz} \end{cases}$$

$$\text{ist } a = \frac{\text{cov}(x, y)}{\text{var}(x)} = \frac{\langle y_t x_t - 2y_t \bar{x} + \bar{y} \bar{x} \rangle_t}{\langle x_t^2 - 2x_t \bar{x} + \bar{x}^2 \rangle_t} = \frac{\langle y_t x_t \rangle_t - \bar{y} \bar{x}}{\langle x_t^2 \rangle_t - \bar{x}^2} \quad (2.6)$$

Dabei gehen wir von dem Fehlermodell aus, dass die T Daten den „echten“, unbeobachteten deterministischen Wert als Mittelwert besitzen.

2.2.2 Parameterbestimmung: Minimierung einer Zielfunktion

Mit der Methode von Gauß können wir die obigen Ergebnisse (2.6) auch anders herleiten. Die Methode besteht darin, durch geeignete Wahl des Parameters a die mittlere Abweichung möglichst klein zu machen. Da die T Abweichungen ($y_t - f(a, x_t)$) der t -ten Beobachtung $y_t = y_0 + ax_t + u_t$ von der Schätzung $f(a, x_t) = y_0 + ax_t$ sowohl positiv als auch negativ sein können, nehmen wir als Approximationsziel anstelle der einfachen Abweichungen das Quadrat der T Abweichungen im Mittel möglichst klein zu machen

$$R(a) = \langle (y_t - f(a))^2 \rangle_t \stackrel{!}{=} \text{minimal} \quad (2.7)$$

Wie erreichen wir dieses Ziel? Die Zielfunktion $R(a)$ wird minimal, wenn die partielle Ableitung nach a null wird:

$$\frac{\partial}{\partial a} R(a) = \frac{\partial}{\partial a} \langle (f(a) - y_t)^2 \rangle_t = \langle 2(f(a) - y_t) f'(a) \rangle_t = 0 \quad \text{mit } f'(a) = x_t$$

$$\Rightarrow \frac{1}{2} \frac{\partial}{\partial a} R(a) = \langle (y_0 + ax_t - y_t)x_t \rangle_t = a \langle x_t^2 \rangle + y_0 \bar{x} - \langle y_t x_t \rangle = 0$$

Mit Gl.(2.5) gilt also

$$0 = a \langle x_t^2 \rangle + (\bar{y} - a \bar{x}) \bar{x} - \langle y_t x_t \rangle = a \langle x_t^2 \rangle + \bar{y} \bar{x} - a \bar{x}^2 - \langle y_t x_t \rangle$$

$$a (\langle x_t^2 \rangle - \bar{x}^2) = \langle y_t x_t \rangle_t - \bar{y} \bar{x} \quad \text{oder} \quad a = \frac{\langle y_t x_t \rangle_t - \bar{y} \bar{x}}{\langle x_t^2 \rangle_t - \bar{x}^2} \quad (2.8)$$

Wir sehen also, dass die Gauß'sche Methode der kleinsten Quadrate die selbe Formel wie in Gl.(2.6) liefert; wir sagen, die Lösung in Gl.(2.6) „hat die Gauß'sche Proportion“.

2.2.3 Parameterbestimmung: Multiple Regression

Wir können diese Methode auf mehrere Variable („multiple Regression“) erweitern. Angenommen, wir haben T Beobachtungen zu Zeitpunkten $t = 1, \dots, T$ der Form

$$y_t = f(x_1, x_2, \dots, x_k) = a_0 + a_1 x_1(t) + a_2 x_2(t) + \dots + a_k x_k(t) + u_t. \quad (2.9)$$

Dann ist dies auch als Skalarprodukt formulierbar eines Parametervektors in Spaltenform mit dem Zeilenvektor der Variablen

$$y_t = (1, x_1(t), \dots, x_k(t)) (a_0, a_1, \dots, a_k)^T + u_t = \mathbf{x}\mathbf{a}^T + u_t. \quad (2.10)$$

Alle $t = 1, \dots, T$ Beobachtungen lassen sich dann zusammenfassen zu der linearen Form

$$\mathbf{y} = (y_1, \dots, y_T)^T = \mathbf{X}\mathbf{a}^T + \mathbf{u}. \quad (2.11)$$

mit der Matrix \mathbf{X} aus T Zeilen und $k+1$ Spalten. Um eine Identifizierung der $k+1$ Parameter a_i zu ermöglichen, muss dabei der Rang der Matrix $\text{rang}(\mathbf{X}) = k+1$ sein. Das Minimum des erwarteten quadratischen Fehlers $R(\mathbf{a})$ bezüglich der Schätzung $\hat{\mathbf{y}} = f(\mathbf{a}) = \mathbf{x}\mathbf{a}^T$ wird erreicht, wenn die partielle Ableitung bezüglich jedes Parameters a_i null ist:

$$\frac{\partial}{\partial a_i} R(\mathbf{a}) = \frac{\partial}{\partial a_i} \frac{1}{T} \left(\sum_{t=1}^T (y_t - \hat{y})^2 \right) = \frac{1}{T} \sum_{t=1}^T 2(y_t - \hat{y}) x_{ti} = 0 \quad (2.12)$$

Da dies für alle $i = 1, \dots, k+1$ gilt, erhalten wir insgesamt $k+1$ Gleichungen mit $k+1$ Unbekannten. Die Lösung ergibt sich nach einigen Umformungen und Zusammenfassen der a_i zu dem Lösungsvektor $\mathbf{a} = (a_0, \dots, a_k)^T$ der Parameter

$$\mathbf{a} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} \quad \Leftrightarrow \mathbf{y} = \mathbf{X}\mathbf{a} \quad (2.13)$$

2.2.4 Parameterbestimmung: TLMSE

Die Minimierung des quadratischen Fehlers bei der Approximation durch ein lineares System ist eng verwandt mit dem Problem, eine unbekannte Funktion $f(\mathbf{x})$ in einem Intervall durch eine Gerade (allgemein: Hyperebene) $F(\mathbf{x})$ mit dem kleinsten quadratischen Fehler zu approximieren. Betrachten wir dazu in Abb. 2.3 einen verrauschten Messpunkt der unbekanntes Funktion. Mit dem üblichen Kriterium des quadratischen Fehlers wird in diesem Fall versucht, die erwartete Abweichung $\langle (f(\mathbf{x}) - y)^2 \rangle$ so klein wie möglich zu machen. Dies ist aber nicht optimal. Besser ist es, anstelle nur des vertikalen Abstand der Datenpunkte den erwarteten tatsächlichen Abstand der Punkte senkrecht zu der Geraden (Hyperebene) zu minimieren (*Total Least Mean Squared Error* TLMSE). Dazu betrachten wir die geometrischen Verhältnisse in Abb. 2.4 genauer.

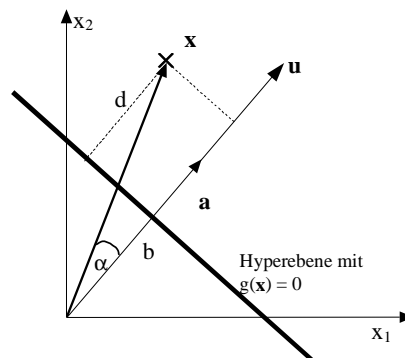


Abb. 2.4 Lineare Approximation einer verrauschten Funktion

Für einen beliebigen Datenpunkt $\mathbf{x} = (x_1, \dots, x_n, y)^T$ ist die Projektion auf einem beliebigen Vektor \mathbf{u} , der senkrecht auf der Geraden (Hyperebenen) steht, mit $\mathbf{x}^T \mathbf{u}$ gegeben. Normieren wir den Spaltenvektor \mathbf{u} auf die Länge eins, so ist die Projektion $\mathbf{x}^T \mathbf{u} / |\mathbf{u}| = \mathbf{x}^T \mathbf{a}$ auf den normierten Einheitsvektor $\mathbf{a} = \mathbf{u} / |\mathbf{u}|$ gerade die Strecke $b + d$ in der Zeichnung. Für alle \mathbf{x} , die auf der Geraden (Hyperebenen) liegen, ist diese Projektion genau der Abstand b vom Nullpunkt zur Geraden. Für alle Punkte der Geraden gilt somit die Geradengleichung

$$g(\mathbf{x}) = \mathbf{x}^T \mathbf{a} - b = 0 \quad |\mathbf{a}| = 1 \quad \text{Hesse'sche Normalform} \quad (2.14)$$

der Hyperebene

Der Abstand d eines Datenpunktes \mathbf{x} von der Geraden ist gerade der Anteil der Projektion von \mathbf{x} auf den normierten Abstandsvektor (Vektor der Flächennormalen) $\mathbf{a} = \mathbf{u} / |\mathbf{u}|$, der über b hinausgeht

$$d = \mathbf{x}^T \frac{\mathbf{u}}{|\mathbf{u}|} - b = g(\mathbf{x}) = \mathbf{x}^T \mathbf{a} - b \quad (2.15)$$

Damit lautet das Fehlerkriterium

$$\min_{\mathbf{a}, b} R(\mathbf{a}, b) = \min_{\mathbf{a}, b} \langle d^2 \rangle = \min_{\mathbf{a}, b} \langle g(\mathbf{x})^2 \rangle = \min_{\mathbf{a}, b} \langle (\mathbf{x}^T \mathbf{a} - b)^2 \rangle \quad |\mathbf{a}| = 1 \quad (2.16)$$

Das Minimum der Zielfunktion $R(\mathbf{a}, b) = \langle d^2 \rangle$ bezüglich des Abstandes b ist erreicht, wenn

$$\frac{\partial R(\mathbf{a}, b)}{\partial b} = 2 \langle \mathbf{x}^T \mathbf{a} - b \rangle = 0 \quad \text{oder} \quad b = \langle \mathbf{x}^T \mathbf{a} \rangle \quad (2.17)$$

so dass das Ziel zu

$$\min_{\mathbf{a}, \mathbf{b}} R(\mathbf{a}, \mathbf{b}) = \min_{\mathbf{a}, \mathbf{b}} \left\langle (\mathbf{x}^T \mathbf{a} - \mathbf{b})^2 \right\rangle = \min_{\mathbf{a}, \mathbf{b}} \left\langle \left((\mathbf{x}^T - \langle \mathbf{x}^T \rangle) \mathbf{a} \right)^2 \right\rangle \quad |\mathbf{a}| = 1 \quad (2.18)$$

wird. Das Minimum bezüglich der *Neigung* \mathbf{a} der Hyperebene ist bei

$$\min_{\mathbf{a}} R(\mathbf{a}) = \min_{\mathbf{a}} \left\langle \mathbf{a}^T (\mathbf{x} - \langle \mathbf{x} \rangle) (\mathbf{x}^T - \langle \mathbf{x}^T \rangle) \mathbf{a} \right\rangle = \min_{\mathbf{a}} \mathbf{a}^T \mathbf{C} \mathbf{a} \quad \text{mit} \quad (2.19)$$

$$\text{der Kovarianzmatrix } \mathbf{C} = \left\langle (\mathbf{x} - \langle \mathbf{x} \rangle)(\mathbf{x} - \langle \mathbf{x} \rangle)^T \right\rangle$$

gegeben. Die Extremwerte dieser quadratischen Form unter der Nebenbedingung $|\mathbf{a}| = 1$ bzw. $\mathbf{a}^2 - 1 = 0$ treten auf, wenn als notwendige Bedingungen die Ableitungen der Lagrange'schen Funktion null werden.

$$L(\mathbf{a}, \mu) = R(\mathbf{a}) + \mu(\mathbf{a}^2 - 1), \quad \frac{\partial}{\partial \mathbf{a}} L(\mathbf{a}, \mu) = \mathbf{0}, \quad \frac{\partial}{\partial \mu} L(\mathbf{a}, \mu) = 0 \quad (2.20)$$

Dies ist gegeben bei

$$\frac{\partial}{\partial \mathbf{a}} L(\mathbf{a}, \mu) = \frac{\partial}{\partial \mathbf{a}} R(\mathbf{a}) + \frac{\partial}{\partial \mathbf{a}} \mu(\mathbf{a}^2 - 1) = \frac{\partial}{\partial \mathbf{a}} \mathbf{a}^T \mathbf{C} \mathbf{a} + \mu 2\mathbf{a}$$

$$= 2\mathbf{C}\mathbf{a} + \mu 2\mathbf{a} = \mathbf{0} \Leftrightarrow \mathbf{C}\mathbf{a} = \lambda \mathbf{a} \quad \text{mit } \lambda = -\mu$$

Die Bedingung

$$\mathbf{C}\mathbf{a} = \lambda_i \mathbf{a} \quad (2.21)$$

bedeutet, dass die Extremwerte gerade dann angenommen werden, wenn die Parametervektoren \mathbf{a} die Eigenvektoren der Kovarianzmatrix \mathbf{C} bilden. Dabei stehen für $k+1$ Variable x_i auch $k+1$ Eigenwerte λ_i zur Verfügung.

Das Minimum wird erreicht bei dem Eigenvektor \mathbf{a}_i ,

$$\min_{\mathbf{a}} R(\mathbf{a}) = \min_{\mathbf{a}} \mathbf{a}^T \mathbf{C} \mathbf{a} = \min_i \mathbf{a}_i^T \lambda_i \mathbf{a}_i = \min_i \lambda_i \quad (2.22)$$

dessen zugehöriger Eigenwert λ_i den kleinsten Wert hat.

Für die Lösung der Minimalisierung des Fehlers für die Parameter der Hyperebene wählen wir also den *Eigenvektor mit dem kleinsten Eigenwert*, da dann als Extremum das Minimum vorliegt. Geometrisch können wir uns dies in Abb. 2.4 verdeutlichen. Würden alle Messwerte \mathbf{x} auf der Geraden liegen, so hat der Eigenvektor mit dem größten Eigenwert die Richtung der Geraden und geht durch den Mittelwert auf der Geraden. Der zweite Eigenvektor ist senkrecht dazu und damit in Richtung von \mathbf{a} , wie gewünscht.

Wie erhalten wir die Eigenvektoren und Eigenwerte? Grundsätzlich gilt

$$(\mathbf{C} - \lambda \mathbf{I})\mathbf{a} = \mathbf{0} \quad \Rightarrow |\mathbf{C} - \lambda \mathbf{I}| |\mathbf{a}| = |\mathbf{0}| = 0 \quad \Rightarrow |\mathbf{C} - \lambda \mathbf{I}| = 0 \quad (2.23)$$

Die Determinante der $n \times n$ Matrix $(\mathbf{C} - \lambda_i \mathbf{I})$ ist ein n -dimensionales Polynom in λ , das charakteristische Polynom der Matrix. Die n Nullstellen des Polynoms sind die n Lösungen für λ in Gl.(2.21). Haben wir die λ , so führt das Einsetzen in Gl. (2.21) zu n Gleichungen für n unbekannte Parameter in \mathbf{a} und damit zu den gesuchten Eigenvektoren.

Beispiel

Sei eine Kovarianzmatrix mit den Einträgen

$$\mathbf{C} = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}, \quad \mathbf{C} - \lambda \mathbf{I} = \begin{pmatrix} c_{11} - \lambda & c_{12} \\ c_{21} & c_{22} - \lambda \end{pmatrix} \quad (2.24)$$

gegeben. Dann ist die Determinante

$$\begin{aligned} \det(\mathbf{C} - \lambda \mathbf{I}) &= (c_{11} - \lambda)(c_{22} - \lambda) - c_{21}c_{12} \\ &= \lambda^2 - (c_{11} + c_{22})\lambda + c_{11}c_{22} - c_{21}c_{12} = \lambda^2 + p\lambda + q = 0 \end{aligned} \quad (2.25)$$

und die Lösung der quadratischen Gleichung, die Nullstellen des Polynoms, sind die Eigenwerte λ_i . Die Eigenvektoren können wir nun durch Gl. (2.23) berechnen. In diesem Fall ergeben sich zwei Gleichungen mit zwei Unbekannten a_1, a_2

$$(\mathbf{C} - \lambda \mathbf{I})\mathbf{a} = \mathbf{0} \quad \text{bzw.} \quad \begin{cases} (c_{11} - \lambda)a_1 + c_{12}a_2 = 0 \\ c_{21}a_1 + (c_{22} - \lambda)a_2 = 0 \end{cases} \quad (2.26)$$

die leicht nach a_1 und a_2 aufgelöst werden können. Für unterschiedliche λ ergeben sich so unterschiedliche Eigenvektoren \mathbf{a} .

Diese analytische Methode ist allerdings nur für kleine Matrizen möglich. Für größere Matrizen gibt es numerische Standard-Verfahren, die in fast allen gängigen Statistik- und Analyseprogrammen existieren.

Eine einfache Version, den Eigenvektor mit dem kleinsten Eigenwert zu ermitteln, ist die folgende stochastische Approximation, siehe Anhang A:

$$\tilde{\mathbf{a}}(t+1) = \mathbf{a}(t) - \mathbf{xx}^T \mathbf{a}(t) \quad \text{und} \quad \mathbf{a}(t+1) = \tilde{\mathbf{a}}(t+1) / |\tilde{\mathbf{a}}(t+1)| \quad (2.27)$$

für alle Muster \mathbf{x} . Liegen alle Muster \mathbf{x} vor, so kann man den Erwartungswert

$$\langle \mathbf{xx}^T \rangle = \mathbf{C}_{xx} \quad \text{Autokorrelationsmatrix}$$

bilden und die Gleichung konvergiert schneller:

$$\tilde{\mathbf{a}}(t+1) = \mathbf{a}(t) - \mathbf{C}_{xx} \mathbf{a}(t) \quad \text{und} \quad \mathbf{a}(t+1) = \tilde{\mathbf{a}}(t+1) / |\tilde{\mathbf{a}}(t+1)| \quad (2.28)$$

Den dazu gehörenden Eigenwert bekommen wir durch Einsetzen in Gl.(2.21) und Multiplikation mit der Kovarianzmatrix.

2.3 Die Grenzen des linearen Modells

Das bisher eingeführte lineare Modell ist wegen seiner Einfachheit bestechend, birgt aber viel mehr Problem, als man auf den ersten Blick vermutet. So muss man sich fragen,

- ob die beobachteten Werte wirklich einem linearen Zusammenhang entstammen,
- ob die Parameter konstant oder zeitabhängig sind und, noch grundsätzlicher,
- ob bestimmte vermutete Einflüsse wirklich vorhanden oder die Parameter null sind und nur stattdessen rein zufällige Abweichungen vorliegen.

Derartige Fragen sind bei allen Arten von Modellierungen möglich und sollen deshalb für den einfachen linearen Fall genauer untersucht werden.

2.3.1 Linearisierung

Eine wichtige Ursache der Probleme bei dem Verfahren der linearen Modellierung liegt in dem nicht näher bezeichneten Kontext der Modellierung. Im Unterschied zu vielen naturwissenschaftlichen Modellierungen sind beispielsweise ökonomische Modelle immer nur ein Aspekt der Wirklichkeit, der nur unter bestimmten Umständen zutrifft. Ändern sich die Umstände, so können die lineare Modellierung bzw. ihre Parameterwerte falsch werden; sie sind deshalb immer nur für einen Zeitabschnitt gültig.

Beispiel

In unserem ökonomischen Beispiel des vorigen Abschnitts nahmen wir eine direkte Proportionalität zwischen Einkommen und Konsum an. Dies ist eine Annahme, die durchaus auf einen großen Teil der Bevölkerung zutreffen kann, aber bei Großverdienern nicht mehr stimmt: Wer schon alles hat, kauft nichts mehr; dem persönlichen Konsum sind Schranken gesetzt. Ob die lineare Formulierung $y = f(x) = y_0 + ax$ für den Zusammenhang $f(\cdot)$ ausreicht, oder ob eine abgeflachte Form $f(x) = y_0 + ax^{1/2}$ oder eine sigmoide Form $f(x) = [1 + e^{-ax}]^{-1}$ besser passt, kann man nach einer Anpassung der Parameter an die Daten testen. Die Frage, ob das vermutete Modell richtig ist oder nicht, ob also die zufallsbehafteten Daten zu stark von der Prognose abweichen, lässt sich mit einem sog. Hypothesentest herausfinden.

Eine wichtige Motivation, trotz vielfach beobachteter nichtlinearer Zusammenhänge das lineare Modell zu verwenden, liegt in der Möglichkeit, die Zusammenhänge zu **linearisieren**.

Beispiele:

- Angenommen, der Zusammenhang existiert mit $f(x) = y_0 + ax^{1/2} + u$. Dann lässt sich durch die Wahl (Transformation der Eingangsvariablen) mittels $\tilde{x} := x^{1/2}$ der Zusammenhang auch linear formulieren zu $f(x) = y_0 + a\tilde{x} + u$.
- Angenommen, wir haben exponentielle Zusammenhänge, etwa $y = \alpha x_1^\alpha x_2^\beta e^u$ („Cobb-Douglas-Produktion“). Dann zeigt eine Logarithmierung der beobachteten Werte y den Zusammenhang $\log(y) = \tilde{y} = a + \alpha \log x_1 + \beta \log x_2 + u = a + \alpha \tilde{x}_1 + \beta \tilde{x}_2 + u$ wieder in einer linearen Formulierung.
- Angenommen, wir haben Kontextbedingungen, die nicht-linear eingehen, etwa qualitative Variable wie Geschlecht oder Religion. Dann lässt sich durch Einführen einer Dummy-Variablen, etwa $\{x_3 \mid x_3 = 1 \text{ wenn männlich, sonst } x_3 = 0\}$, der männlichen Einfluss auf y modellieren.

2.3.2 Parameterschätzung

Eine weitere Frage ist die nach der rationalen Grundlage einer gegebenen Modellierung. Sind wirklich alle Variablen notwendig, um die Beobachtungen zu beschreiben? Dies ist die wichtige Frage nach der Modellstruktur: Man sollte nur die minimale Anzahl von Variablen wählen, um die Beobachtungen zu erklären. Diese Frage transformiert sich zu der Frage: Ist der Wert der Modellparameter tatsächlich ungleich null, oder kann man beobachtete Abweichungen auf die sowieso vorhandenen zufälligen Abweichungen schieben?

Bei jeder Beobachtung für y können wir eine Schätzung \hat{a}_i der Parameter a_i abgeben, wobei wir annehmen, dass schließlich

$$\langle \hat{a}_i(t) \rangle_t = a_i \quad \text{erwartungswerttreu (unbiased)} \quad (2.29)$$

gilt. In diesem Fall ist \hat{a}_i eine Zufallsvariable, die natürlich auch eine Verteilung hat. Erwartungstreue ist zwar hinreichend, aber nicht notwendig; bei zeitveränderlichen Verteilungen reicht es uns auch, wenn \hat{a}_i *konsistent* ist, also im Grenzwert zu a_i geht:

$$\lim_{t \rightarrow \infty} \text{Prob}(|\hat{a}_i(t) - a_i| < \varepsilon) = 1 \quad \forall \varepsilon > 0 \quad \text{konsistent} \quad (2.30)$$

Dies kann man symbolisch abkürzen mit

$$\text{plim } \hat{a}_i = a_i \quad (2.31)$$

In Abb. 2.5 ist der Unterschied zwischen *unbiased* und *konsistent* bei den Dichtefunktionen eines Parameters visualisiert. Bei beiden Mechanismen konvergiert eine Dichtefunktion $p(\hat{a}_i)$ mit anwachsender Zahl von Beobachtungen T immer präziser zum „wahren“ Wert von a_i .

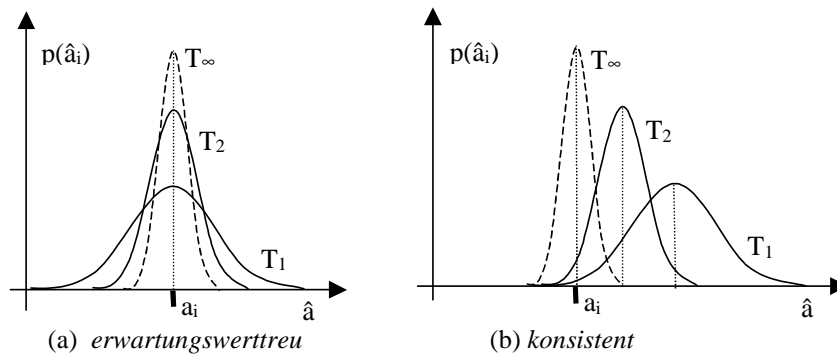


Abb. 2.5 Unterschiede in der Parameterverteilung

Wie können wir die Konsistenz nachprüfen? Üblicherweise reicht es aus, die Bedingungen

$$\lim_{T \rightarrow \infty} \langle \hat{a}_i(t) \rangle_t = a_i \quad \text{und} \quad \lim_{T \rightarrow \infty} \text{var}(\hat{a}_i) = 0 \quad \text{bei } t = 1, \dots, T \quad (2.32)$$

nachzuweisen. Leider sind sie aber nur hinreichend, nicht notwendig: Die Varianz kann auch unendlich sein und trotzdem konsistente Parameter erzeugen. Die im letzten Abschnitt vorgestellte Approximation mit dem kleinsten quadratischen Fehler erzeugt bei fortschreitender Anzahl T zwar Verteilungen der Parameter mit abnehmender Varianz (sie ist „effizient“), aber eine endliche Varianz wird vorausgesetzt.

Verfügt man noch über zusätzliche Informationen, so kann man stattdessen auch eine Ridge-Regression (s. Greene 2003) durchführen, die erheblich schneller konvergiert als die beste quadratische Fehlerabschätzung.

2.3.3 Zur Rolle des Störterms

Bei der Modellierung in Gl.(2.3) bzw. Gl.(2.9) nahmen wir an, dass der Erwartungswert $\langle u_t \rangle = 0$ ist. Angenommen, dies ist bei unserer Modellierung nicht der Fall: Was machen wir dann? Im einfachsten Fall ist es eine Konstante $\langle u_t \rangle = u_0$, so dass wir Gl. (2.9) umformen können zu

$$y_t = (a_0 + u_0) + a_1 x_1(t) + a_2 x_2(t) + \dots + a_k x_k(t) + (u_t - u_0). \quad (2.33)$$

Die so korrigierte Zufallsvariable \tilde{u}_t hat wieder der Erwartungswert null:

$$\langle \tilde{u}_t \rangle = \langle u_t - u_0 \rangle = 0 \quad (2.34)$$

Ein weiteres Problem tritt auf, wenn der Störterm mit einer der Variablen x_k korreliert sein sollte: Dies müssen wir durch einen Test ausschließen, da sonst die Modellierung falsch ist. Dabei darf der Störterm auch nicht mit sich selbst korreliert sein (Autokorrelation), so dass auch gelten muss

$$\text{cov}(u_t, u_{t'} | \mathbf{x}_t) = 0 \quad \forall t, t' = 1, \dots, T \quad t \neq t' \quad (2.35)$$

Meist wird dabei angenommen, dass die stochastische Variable eine Überlagerungssumme aller Störeinflüsse ist. Nach dem zentralen Grenzwertsatz strebt eine solche Summe aus unabhängigen stochastischen Variablen, egal welcher Verteilung, die Gauß'sche Normalverteilung an. Man nimmt deshalb an, dass u_t normalverteilt ist.

2.3.4 Multikollinearität

Für die Lösung der linearen Form aus Gl. (2.11) war Voraussetzung, dass der Rang der Matrix \mathbf{X} gleich $k+1$ ist, also die Variablen x_i linear unabhängig voneinander sind. Angenommen, dies ist bei korrelierten Variablen $x_i = \alpha x_j$ nicht erfüllt, dann gilt diese Voraussetzung nicht mehr und das Gleichungssystem hat keine Lösung. Zwar ist dies in der Praxis selten der Fall, aber um uns Schwierigkeiten zu machen reicht es, wenn die Variablen eine endliche lineare Korrelation miteinander haben. Diesen Sachverhalt nennen wir **Multikollinearität**. Zwar erhalten wir mit unserer Methode der kleinsten quadratischen Fehler immer noch eine korrekte Lösung, aber in diesem Fall verbreitert sich die Varianz jeder Variablen (und damit der Schätzfehler) um die Korrelationsanteile; die für eine gute Schätzung der Parameter benötigten Datenmenge nimmt zu. Anders ausgedrückt: Bei gleicher Datenmenge erhalten wir einen größeren Schätzfehler.

Wie erkennen wir, ob eine Multikollinearität zwischen der Variablen x_i und x_j vorliegt? Dazu bilden wir zuerst den Korrelationskoeffizienten r_{ij}

$$r_{ij} = \frac{\text{cov}(x_i, x_j)}{\sqrt{\text{var}(x_i)} \sqrt{\text{var}(x_j)}} = \frac{\text{cov}(x_i, x_j)}{\sigma_i \sigma_j} \quad (2.36)$$

Er gibt an, wie stark die beiden Variablen miteinander korreliert sind und kann Werte zwischen -1 und $+1$ annehmen. Sind die beiden Variablen nicht korreliert, so ist er null. Im Falle der Multikollinearität muss er nicht nur ungleich null sein, sondern eher an eins liegen. Wie groß nun genau?

Der benötigte Schwellwert, ab wann man von Multikollinearität spricht, sollte nicht willkürlich gewählt werden, sondern von den Daten abhängen. Betrachten wir dazu die Daten. Angenommen, wir können jede Beobachtung $y_t = \hat{y}_t + \hat{u}_t$ in einen deterministischen Teil \hat{y}_t und einen Zufallsteil \hat{u}_t , das Residuum, aufteilen. Dann folgt mit der Annahme, dass \hat{u}_t unkorreliert zu \hat{y}_t bzw. \bar{y} ist und $\langle u_t \rangle = 0$

$$\begin{aligned}\langle (y_t - \bar{y})^2 \rangle &= \langle (\hat{y}_t + \hat{u}_t - \bar{y})^2 \rangle = \langle (\hat{y}_t - \bar{y})^2 + 2(\hat{y}_t - \bar{y})\hat{u}_t + \hat{u}_t^2 \rangle \\ &= \langle (\hat{y}_t - \bar{y})^2 \rangle + 2\langle \hat{u}_t \rangle \langle (\hat{y}_t - \bar{y}) \rangle + \langle \hat{u}_t^2 \rangle = \langle (\hat{y}_t - \bar{y})^2 \rangle + \langle \hat{u}_t^2 \rangle\end{aligned}$$

Also ist

$$1 = \frac{\langle (\hat{y}_t - \bar{y})^2 \rangle}{\langle (y_t - \bar{y})^2 \rangle} + \frac{\langle \hat{u}_t^2 \rangle}{\langle (y_t - \bar{y})^2 \rangle} \geq \frac{\langle (\hat{y}_t - \bar{y})^2 \rangle}{\langle (y_t - \bar{y})^2 \rangle} \equiv R^2 \quad (2.37)$$

wobei wir uns mit R das *Bestimmtheitsmaß* des linearen Modells definieren. Haben wir keine Abweichungen, so ist $R = 1$; haben wir nur unsystematische Abweichungen, also kein lineares Modell, so ist $R = 0$. Das Maß hat also durchaus plausible Eigenschaften und erlaubt eine Aussage, wie gut der Ansatz „lineare Funktion“ die vorliegenden Daten erklärt.

Nun können wir es auf unser ursprüngliches Problem anwenden: Wir nehmen eine Multikollinearität genau dann an, wenn

$$r_{ij}^2 \geq R^2 \quad \text{Definition Multikollinearität} \quad (2.38)$$

gilt, die lineare Abhängigkeit der Variablen untereinander also größer ist als die lineare Abhängigkeit, die durch alle Variablen gemacht wird.

2.3.5 Testen von Einflüssen

Angenommen, unsere Theorie sagt eine Überlagerung von exogenen Einflüssen x_i auf die endogene Variable y in der Form Gl. (2.11) voraus. Auch wenn wir annehmen, dass es sich tatsächlich um lineare Einflüsse handelt, woher wissen wir, dass bestimmte Einflüsse x_i wichtig sind, also ihre Parameter a_i ungleich null sind? Wenn aus den beobachteten Daten ersichtlich ist, dass ein fraglicher Parameter a_r eigentlich mit null approximiert werden kann, so gewinnen wir damit über die Vereinfachung der Gleichung und die Genauigkeit der anderen ermittelten Parameter hinaus auch den prinzipiellen Erkenntnisgewinn, dass der Einfluss x_r nur marginal ist und deshalb auch in der Theorie getrost weggelassen kann. Dies vereinfacht damit auch die Theorie!

Wie können wir nun testen, ob ein Parameter $a_r \neq 0$ ist? In der Statistik wird dies als „Test einer Hypothese“ bezeichnet. Üblicherweise gibt ein solcher Test als Antwort auf die Frage nicht einfach „Ja“ oder „Nein“ zurück, sondern die Wahrscheinlichkeit, mit der die Hypothese zutrifft. Dazu bilden wir die **Nullhypothese** H_0 :

$$H_0: a_r = 0 \quad \text{Nullhypothese} \quad (2.39)$$

Ist die Nullhypothese für a_r erfüllt, so bedeutet dies, dass die Variable x_r keinen Einfluss hat und damit weggelassen werden kann.

Der Test auf eine Nullhypothese lässt sich übrigens auch verwenden, wenn wir testen wollen, ob a_r einen bestimmten, festen Erwartungswert s hat. Dazu definieren wir uns einfach eine Variable $\hat{a}_r = (a_r - s)$, für die wieder die Nullhypothese gelten muss. Ist die Nullhypothese für \hat{a}_r erfüllt, so bedeutet dies, dass a_r den Erwartungswert $s \neq 0$ hat und damit **nicht** weggelassen werden kann.

Wie funktioniert nun der Test? Wenn wir die Verteilungsdichte $p(\hat{a}_r)$ der beobachteten Werte \hat{a}_r für a_r erfassen, so erhalten wir eine Verteilung mit der beobachteten Varianz σ_r^2 . Die Werte von \hat{a}_r sollten dabei nicht *signifikant* („zu stark“) von null abweichen. Was bedeutet „zu stark“ konkret? Nehmen wir an, dass es sich um eine Normalverteilung handelt, so sollten die beobachteten \hat{a}_r mit einer Wahrscheinlichkeit $1 - \alpha$ kleiner als der Wert p_α sein, z.B. $1 - \alpha = 95\%$, siehe Abb. 2.6. Dies entspricht dem Integral (Fläche unter der Kurve) der Dichte von $-\infty$ bis p_α . Ist dagegen die relative Anzahl der überschreitenden Werte größer als α , so lässt sich die Nullhypothese nicht aufrecht erhalten und wird abgelehnt.

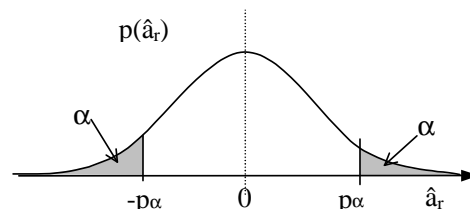


Abb. 2.6 Rechts- und linksseitiger Hypothesentest

Wir sprechen in diesem Fall von einem *rechtsseitigen* Hypothesentest und akzeptieren die Nullhypothese, wenn

$$\text{Prob}(\hat{a}_r > p_\alpha) \leq \alpha \quad \text{z.B. } \alpha = 5\% \Rightarrow p_\alpha = 1,645 \quad (2.40)$$

gegeben ist. Möchten wir sichergehen, dass die beobachteten Werte auf beiden Seiten nicht zu stark abweichen, also im nicht schraffierten Bereich der Dichte in Abb. 2.6 bleiben, so fordern wir zusätzlich noch den *linksseitigen* Hypothesentest:

$$\text{Prob}(\hat{a}_r < -p_\alpha) \leq \alpha \quad (2.41)$$

Die Kombination beider Tests wird als *beidseitiger* Hypothesentest bezeichnet. Die Nullhypothese wird in diesem Fall akzeptiert, wenn

$$\text{Prob}(|\hat{a}_r| > p_\alpha) \leq \alpha \quad \text{z.B. } \alpha = 5\% \Rightarrow p_\alpha = 1,96 \quad (2.42)$$

gilt. Die praktische Durchführung des Tests verwendet für die Bestimmung der p_α bei gegebenem α die in Tabellen erfassten Werte der Normalverteilung, also des Integrals der Gauß-Funktion, genannt „Gaußsche Fehlerfunktion Φ “. Dabei wird eine auf eins normierte Varianz σ^2 der Parameter vorausgesetzt. Für den praktischen Test müssen wir also den aus den Messwerten errechneten Parameterwert

mit dem auf die gemessene Abweichung σ skalierten Tabellenwert $\sigma \cdot p_\alpha$ vergleichen.

Die Varianzen s_0^2 und s_1^2 der Parameter a_0 und a_1 erhalten wir entweder aus linearen Approximationen der gemessenen Daten, oder direkt aus den T gemessenen Daten mithilfe der geschätzten, korrigierten Stichprobenvarianz s_u^2 des Rauschens

$$s_0^2 = \frac{s_u^2 \langle x_t^2 \rangle}{\sum_{t=1}^T (x_t - \bar{x})^2}, \quad s_1^2 = \frac{s_u^2}{\sum_{t=1}^T (x_t - \bar{x})^2} \quad \text{mit} \quad s_u^2 = \frac{1}{T-1} \sum_{t=1}^T u_t^2 \quad (2.43)$$

Für multiple Regression sind die Varianzen s_i^2 der Parameter a_i die Hauptdiagonalelemente der Kovarianzmatrix

$$\mathbf{C}_a = s_u^2 (\mathbf{X}^T \mathbf{X})^{-1} \quad (2.44)$$

mit der Matrix \mathbf{X} aus den $t = 1, \dots, T$ Datenzeilen $(1, x_{t1}, \dots, x_{tk})$.

Liegt keine Normalverteilung vor, so kann man den Variablenbereich von \hat{a}_r in Intervalle aufteilen, ein Histogramm aufstellen, es auf eins normieren und so die Verteilungsdichte approximieren. Die weiteren Überlegungen gelten dann analog.

Aufgabe 2.1

Stellen Sie bei 10 von Ihren Bekannten die Schuhgröße, Handlänge, das Gewicht und die Körpergröße fest.

- Ist die Körpergröße ein lineare Funktion von Schuhgröße, Handlänge und Gewicht?
- Existieren Multikollinearitäten?
- Lassen sich Variable weglassen?

2.4 Nichtlineare Modellierung mit Polynomen

Sind die systematischen Abweichungen zu groß, weil etwa die Mittelwerte eine deutliche Abweichung von einer Geraden andeuten, so liegt es nahe, eine nichtlineare Modellierung durchzuführen. Was steht uns dazu zur Verfügung? Der intuitivste Ansatz besteht darin, systematisch nichtlineare Funktionen an den Daten auszuprobieren. Da sich die meisten relevanten reellen Funktionen durch Polynome darstellen lassen, etwa in Form einer Taylorentwicklung

$$f(x) = f(x_0) + (x-x_0)f^{(1)}(x_0) + \frac{(x-x_0)^2}{2!}f^{(2)}(x_0) + \frac{(x-x_0)^3}{3!}f^{(3)}(x_0) + \dots \quad (2.45)$$

$$= A + B(x-x_0) + C(x-x_0)^2 + D(x-x_0)^3 + \dots$$

so können wir versuchen, auch die Daten durch Polynome höherer Ordnung zu approximieren. Die Parameter erhalten wir im deterministischen Fall wieder durch ein System von n Gleichungen für n unbekannte Parameter.

Beispiel

Angenommen, wir haben 3 Punkte (x_1, y_1) , (x_2, y_2) , (x_3, y_3) des Polynoms $f(x) = a + bx + cx^2$. Wie lauten dann die Parameter des Polynoms? Wir wissen, dass

$$\begin{aligned} y_1 &= a + bx_1 + cx_1^2 & \Rightarrow y_2 - y_1 &= b(x_2 - x_1) + c(x_2^2 - x_1^2) \\ y_2 &= a + bx_2 + cx_2^2 & \Rightarrow y_3 - y_2 &= b(x_3 - x_2) + c(x_3^2 - x_2^2) \\ y_3 &= a + bx_3 + cx_3^2 \end{aligned} \quad (2.46)$$

Also ist

$$\frac{y_2 - y_1}{x_2 - x_1} = b + c \frac{x_2^2 - x_1^2}{x_2 - x_1} \quad \text{und} \quad \frac{y_3 - y_2}{x_3 - x_2} = b + c \frac{x_3^2 - x_2^2}{x_3 - x_2} \quad (2.47)$$

Bilden wir die Differenz beider Gleichungen in Gl.(2.47), so können wir die erhaltene Gleichung nach c auflösen. Eingesetzt in eine der Gleichungen von Gl.(2.47) gibt uns b , und eingesetzt in eine der Originalgleichungen von Gl.(2.46) gibt uns den letzten Parameter a .

Verallgemeinert kann man mit n Parametern a_1, \dots, a_n und n Funktionswerten y_1, \dots, y_n ein lineares Gleichungssystem aus den Spaltenvektoren $\mathbf{y} = (y_1, \dots, y_n)^T$, $\mathbf{a} = (a_1, \dots, a_n)^T$ und der Matrix \mathbf{X} aus den numerisch ermittelten Potenzen bilden

$$\mathbf{y} = \mathbf{X} \mathbf{a} \quad \mathbf{X} = \begin{bmatrix} \mathbf{X}_{ij} \end{bmatrix} = \begin{bmatrix} x_i^j \end{bmatrix} \quad (2.48)$$

Und mit einem Schema lösen, das auf dieser rekursiven Ermittlung der Parameter beruht, dem *Gauß'schen Eliminationsverfahren*. Dabei wird durch eine Reihe von

Schritten die Matrix \mathbf{X} auf die Form einer oberen Dreiecksmatrix gebracht. Inhaltlich wird dabei die n -te Ableitung gebildet und damit der n -te Parameter ermittelt, dann die $(n-1)$ -te Ableitung für den $(n-1)$ -ten Parameter usw.

Haben wir keine deterministischen Beobachtungen, so müssen wir wieder die Parameter der vermuteten Funktion der Menge der Punkte so anpassen, dass der mittlere quadratische Fehler minimal wird. Das Verfahren wird als **Regression n -ter Ordnung**, das entsprechende Polynom als **Ausgleichsparabel n -ter Ordnung** bezeichnet.

2.5 Nichtlineare Modellierung mit Neuronalen Netzen

Ein interessantes Hilfsmittel, um unbekannte Funktionen auf der Basis gestörter Messpunkte zu modellieren, ist die Approximation mithilfe künstlicher neuronaler Netze.

Solche Netze bestehen aus formalen Neuronen, die analog zu den biologischen Neuronen als Einheiten aufgefasst werden, die viele Eingangsvariable (Signaleingänge) und nur einen Ausgang haben, siehe Abb. 2.7 .

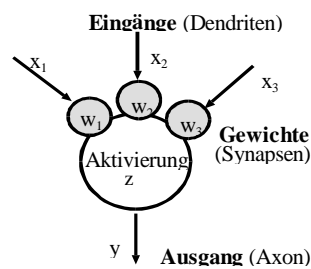


Abb. 2.7 Formales Neuron

Fassen wir die Eingabeaktivitäten $x_1 \dots x_n$ zum Eingabevektor $\mathbf{x} = (x_1, \dots, x_n)^T$ (Muster der Eingabe) und die Gewichte $w_1 \dots w_n$ zum Gewichtsvektor $\mathbf{w} = (w_1, \dots, w_n)^T$ zusammen, so lässt sich die resultierende Aktivität z beispielsweise als Summe der gewichteten Eingaben im Modellneuron (*Sigma-unit*) und damit formal als Skalarprodukt (*inneres Produkt*) beider Spaltenvektoren schreiben:

$$z(\mathbf{w}, \mathbf{x}) = \sum_j w_j x_j = \mathbf{w}^T \mathbf{x} \quad \text{Aktivitätsfunktion} \quad (2.49)$$

Sehr oft muss die Aktivität erst eine Schwelle (*bias*) s überschreiten, bevor sie sich beim Ausgang (Axon) auswirkt. Dies lässt sich durch die Minderung der Aktivität um den Schwellwert modellieren:

$$z(\mathbf{w}, \mathbf{x}) = \mathbf{w}^T \mathbf{x} - s$$

Den zusätzlichen Term s , der hier gegenüber Gl. (2.49) auftritt, kann man allerdings mit einem Trick in der Notation wieder verschwinden lassen, indem man eine Erweiterung der Vektoren um eine Zusatzkomponente

$$\mathbf{x} \rightarrow \mathbf{x} = (x_1, \dots, x_n, 1)^T \quad \text{und} \quad \mathbf{w} \rightarrow \mathbf{w} = (w_1, \dots, w_n, -s)^T \quad (2.50)$$

vor nimmt. Das Skalarprodukt ist somit wieder

$$z(\mathbf{w}, \mathbf{x}) = \sum_j w_j x_j - s = (w_1, \dots, w_n, -s)(x_1, \dots, x_n, 1)^T = \mathbf{w}^T \mathbf{x} \quad (2.51)$$

Die Aktivität y am Neuronenausgang wird durch die Ausgabefunktion (*activation function*) $S(\cdot)$, abhängig von der internen Aktivität z , beschrieben:

$$y = S(z) \quad \text{Ausgabefunktion} \quad (2.52)$$

Mehrere formale Neuronen kann man zu einem Netz zusammenfügen. Alle Neuronen mit der gleichen Aufgabe fassen wir zu einer „Schicht“ zusammen. Betrachten wir dazu zunächst nur Netze aus zwei Schichten, einer Schicht aus Neuronen mit nichtlinearen Ausgabefunktionen, sowie einer zweiten Schicht von einem linearen Neuron pro AusgabevARIABLE $y = \hat{f}(\mathbf{x})$ abgeschlossen wird. In Abb. 2.8 ist dies dargestellt.

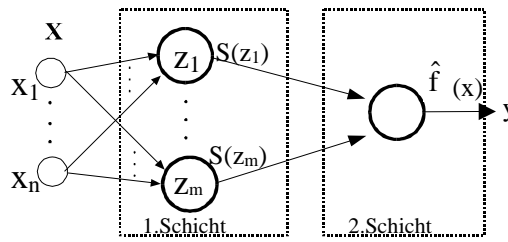


Abb. 2.8 Ein zweischichtiges Netz

Da die Ausgabe S_j der Neuronen der ersten Schicht nicht direkt beobachtet werden kann, nennt man sie auch "versteckte Einheiten" (*hidden units*). Die Eingabeschicht (*input units*) besteht dabei nur aus Signaleingangspunkten, so dass sie keine "echte", lernfähige Schicht darstellt.

Formal lässt sich die Netzfunktion schreiben als

$$\hat{f}(\mathbf{x}) = \sum_{j=1}^m w_j^{(2)} S(\mathbf{w}_j, \mathbf{x}) \quad (2.53)$$

wobei w_j und $F(\mathbf{x})$ aus \mathfrak{R} , \mathbf{x} aus \mathfrak{R}^n sind. Ein solche Architektur reicht für die meisten Aufgaben aus.

Die Aufgabe des Modellierers besteht nun darin, Algorithmen für das Festsetzen und Verbessern („Lernen“) der Gewichte $w_j^{(2)}$ und w_j zu finden. Aus der Fülle möglicher Algorithmen sollen an dieser Stelle nur zwei herausgegriffen werden. Sie werden am häufigsten benutzt und haben sich in vielen Anwendungen bewährt.

2.6 Backpropagation-Netze

Ziel des Lernens soll sein, den Ausführungsfehler (die Differenz zwischen der gewünschten Ausgabe und der tatsächlichen Ausgabe des Netzwerks) möglichst klein zu machen. Dazu benutzen wir den bei der Netzwerkausgabe ermittelten Fehler, die Gewichte der Netzwerkschicht für die gewünschte Funktion besser einzustellen. Diese Idee wird beim berühmten *error back-propagation* Algorithmus auf Mehrschichten-Architekturen (*multilayer feedforward*) angewendet. Der Algorithmus stellt einen wichtigen Standardalgorithmus dar und wurde deshalb bisher als "das Arbeitspferd der neuronalen Netze" bezeichnet. Damit lernen wir einen Algorithmus kennen, den wir immer dann einsetzen können, wenn wir ein neuronales Netz mit Beispielen trainieren wollen, um eine unbekannte Funktion $y = f(\mathbf{x})$ möglichst gut zu approximieren.

2.6.1 Die Funktionsarchitektur

Obwohl der Algorithmus beliebig viele Schichten gestattet, geht man üblicherweise nur von zwei Schichten aus, die aber, wie wir in Abschnitt 2.11 sehen werden, für die Approximation einer beliebigen Funktion prinzipiell ausreichen. Die Eingänge beider Schichten sind, wie in Abb. 2.9 gezeichnet, jeweils vollständig miteinander vernetzt.

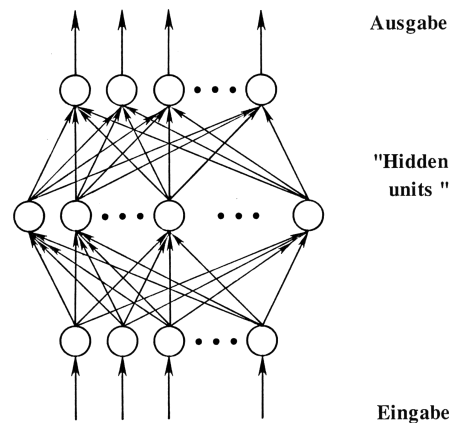


Abb. 2.9 Grundstruktur des Backpropagation-Netzwerks

Die Aktivität des Backpropagation-Netzwerks ist mit Gl. (2.53) gegeben, wobei hier als nichtlineare Ausgabefunktionen meist sigmoidale Funktionen genommen werden, etwa die Fermi-Funktion $S_F(z)$

$$S_F(z) = \frac{1}{1 + e^{-z}}. \quad (2.54)$$

Die Ausgabe wird dabei in mehreren Schritten gewonnen: Zuerst wird die Ausgabe $S_i(z)$ bei jedem einzelnen Neuron gebildet, und dann wird die Gesamtaktivität aller Neuronen der *hidden units* als Eingabe zur zweiten Schicht durchgereicht und die gewichtete Summe gebildet mit der Ausgabe y . Die Aktivität wird also pipeline-artig durch das Netz propagiert, wobei auf strikte Synchronität der Aktivitäten geachtet werden muss.

2.6.2 Lernen im Netzwerk

Die Gewichte der *hidden units* und der Ausgabeneuronen werden beim Backpropagation-Algorithmus durch Training mit Eingabemustern \mathbf{x} und den dazu gewünschten Ausgabemustern \mathbf{L} (Lehrervorgabe) solange verbessert, bis das Netzwerk die gewünschte Leistung (die gewünschte Ausgabe bei einer Eingabe) mit genügender Genauigkeit erbringt. Dazu wird nach dem Durchlaufen (*propagation*) des Eingangssignals \mathbf{x} durch die Netzschichten der Fehler $\delta := \mathbf{y}(\mathbf{x}) - \mathbf{L}(\mathbf{x})$ des Ausgangssignals \mathbf{y} bezüglich der gewünschten Ausgabe \mathbf{L} ermittelt und durch alle Schichten zurückgeführt (*error back-propagation*), wie dies in Abb. 2.10 schematisch gezeigt ist.

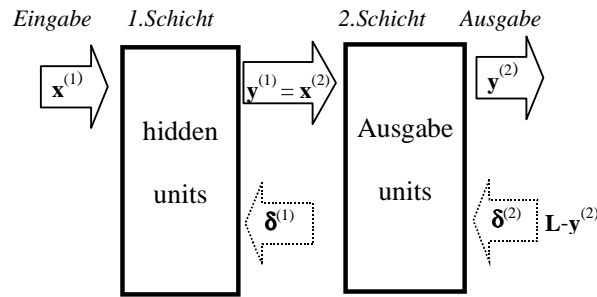


Abb. 2.10 Fehlerrückführung im Mehrschichtensystem

Als Lernziel geben wir uns vor, dass der Algorithmus den erwarteten Fehler für alle Trainingseingaben (*Trainingsmuster*) minimieren soll. Die **Zielfunktion** R lautet also mit dem Proportionalitätsfaktor $1/2$

$$R(\mathbf{w}) = \sum_{\mathbf{x}} R_{\mathbf{x}} = 1/2 \sum_{\mathbf{x}} (\mathbf{y}(\mathbf{x}) - \mathbf{L}(\mathbf{x}))^2 \quad (2.55)$$

über alle Trainingsmuster \mathbf{x} .

Die Frage des Lernens bedeutet hier, wie wir iterativ die optimalen Parameter $(w_1^*, \dots, w_n^*) = \mathbf{w}^*$ für das Minimum der Zielfunktion finden. Angenommen, die zu minimierende Zielfunktion $R(\mathbf{w})$ habe ein lokales, relatives Minimum bei \mathbf{w}^* . Ausgehend von einem initialen \mathbf{w} in der Nähe dieses lokalen Minimums versuchen wir, schrittweise iterativ uns dem optimalen \mathbf{w}^* anzunähern, s. Abb. 2.11.

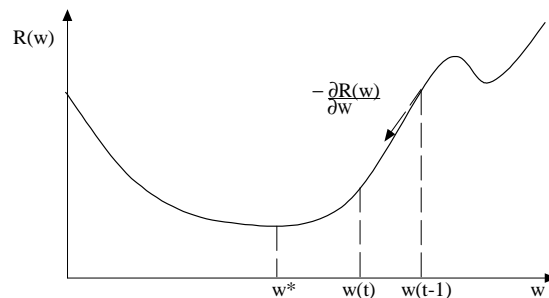


Abb. 2.11 Die Gradientensuche

Die Ableitungen nach allen Komponenten können wir in einem Vektor $(\partial R/\partial w_1, \dots, \partial R/\partial w_n)^T$ zusammenfassen und bezeichnen diesen als *Gradienten* $\text{grad } R(\mathbf{w})$ der Funktion $R(\mathbf{w})$. Dies lässt sich auch mit dem *Nabla-Operator* $\nabla_{\mathbf{w}}$ $:= (\partial/\partial w_1, \dots, \partial/\partial w_n)^T$ als " $\nabla_{\mathbf{w}} R(\mathbf{w})$ " notieren. Der so definierte Vektor der Richtungsableitungen zeigt in der Nähe des Minimums in die Richtung des stärksten Anstiegs der Funktion; der negative Gradient also in Richtung des Minimums. Mit

dieser Überlegung soll die Differenz von $\mathbf{w}(t-1)$ beim $(t-1)$ -ten Schritt zu $\mathbf{w}(t)$ vom nächsten Schritt t proportional zum negativen Gradienten sein:

$$\Delta \mathbf{w} := (\mathbf{w}(t) - \mathbf{w}(t-1)) \sim -\nabla_{\mathbf{w}} R(\mathbf{w}(t-1)) \quad (2.56)$$

oder mit der schrittabhängigen Proportionalitätskonstanten γ

$$\mathbf{w}(t) = \mathbf{w}(t-1) - \gamma(t) \nabla_{\mathbf{w}} R(\mathbf{w}(t-1)) \quad \textit{Gradienten-Lernregel} \quad (2.57)$$

Eine solche Iterationsgleichung zur Optimierung von Parametern einer Zielfunktion wird auch als *Lernregel*, die Proportionalitätskonstante $\gamma(t)$ dabei als *Lernrate* bezeichnet. Im allgemeinen Fall ist γ eine Matrix, die eine Amplituden- und Richtungskorrektur vornimmt; im Normalfall ist eine skalare Funktion ausreichend.

Der Gradientenalgorithmus der Zielfunktion (2.55) für die Iteration des Gewichts w_{ij} von Neuron j zu Neuron i ist

$$w_{ij}(t) = w_{ij}(t-1) - \gamma(t) \sum_{\mathbf{x}} \partial R_{\mathbf{x}} / \partial w_{ij} \quad (2.58)$$

mit der Lernrate $\gamma(t)$. Präsentieren wir hintereinander die einzelnen Trainingsmuster, so ist die Gewichtsänderung für das einzelne Trainingsmuster \mathbf{x}

$$\Delta w_{ij}(\mathbf{x}) := w_{ij}(t) - w_{ij}(t-1) = -\gamma \frac{\partial R_{\mathbf{x}}}{\partial w_{ij}} = -\gamma \frac{\partial R_{\mathbf{x}}}{\partial z_i} \frac{\partial z_i}{\partial w_{ij}} \quad (2.59)$$

Mit der Notation für die partielle Ableitung der Ausgabe

$$\frac{\partial y_i}{\partial z_i} = \frac{\partial S(z_i)}{\partial z_i} =: S'(z_i) \quad (2.60)$$

$$\text{ist mit } \delta_i := \frac{-\partial R_{\mathbf{x}}}{\partial z_i} = \frac{-\partial R_{\mathbf{x}}}{\partial y_i} \frac{\partial y_i}{\partial z_i} = \frac{-\partial R_{\mathbf{x}}}{\partial y_i} S'(z_i), \quad \frac{\partial z_i}{\partial w_{ij}} = \frac{\partial}{\partial w_{ij}} \sum_k w_{ik} x_k = x_j \quad (2.61)$$

die Gewichtsänderung

$$\Delta w_{ij}(\mathbf{x}) = \gamma \delta_i x_j \quad \textit{Delta-Regel} \quad (2.62)$$

Für die Neuronen der zweiten Schicht, deren Ausgabe wir beobachten können, gilt also

$$\frac{\partial R_{\mathbf{x}}}{\partial y_i} = \frac{\partial}{\partial y_i} 1/2(y_i - L_i)^2 = (y_i - L_i) \quad (2.63)$$

der beobachtete Fehler.

Bezeichnen wir zur Unterscheidung die Variablen der Neuronen der ersten Schicht mit dem Index ⁽¹⁾ und die der zweiten Schicht mit ⁽²⁾, so ist für die Ausgabeschicht mit der Definition aus Gl. (2.61)

$$\delta_k^{(2)} = -(y_k^{(2)} - L_k^{(2)})S'(z_k^{(2)}) \quad (2.64)$$

Für die anderen Schichten, beispielsweise für die erste Schicht ("hidden units"), gilt ein komplizierteres Delta. Betrachten wir dazu ein Neuron j , das in der folgenden Schicht andere Neuronen $k=1..m$ beeinflusst, siehe Abb. 2.12.

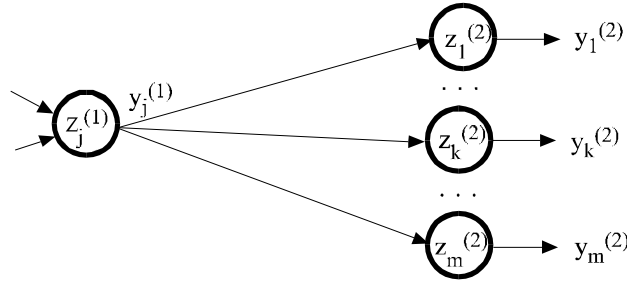


Abb. 2.12 Propagierung der Aktivität von Neuron i zu Neuronen k

Der Fehler R_x ist dabei die Summe der Einzelfehler, die jedes der beeinflussten Ausgabeneuronen macht

$$R_x = (\mathbf{y}(\mathbf{x}) - \mathbf{L}(\mathbf{x}))^2 = \sum_{k=1}^m (y_k - L_k)^2 = \sum_k R_{xk} \quad (2.65)$$

Der beobachtete Fehler im i -ten Neuron der ersten Schicht lässt sich mit (2.61) rekursiv durch die Folgefehler der zweiten Schicht darstellen:

$$\frac{\partial R_x}{\partial y_i^{(1)}} = \sum_k \frac{\partial R_{xk}}{\partial z_k^{(2)}} \frac{\partial z_k^{(2)}}{\partial y_i^{(1)}} = - \sum_k \delta_k^{(2)} \frac{\partial z_k^{(2)}}{\partial y_i^{(1)}} \quad (2.66)$$

Durch

$$\frac{\partial z_k^{(2)}}{\partial y_i^{(1)}} = \frac{\partial}{\partial y_i^{(1)}} \sum_j w_{kj}^{(2)} x_j^{(1)} = w_{ki}^{(2)} \quad (2.67)$$

wird mit Gl. (2.61), (2.66), (2.67) und (2.60)

$$\delta_i^{(1)} = \frac{-\partial R_x}{\partial y_i^{(1)}} \frac{\partial y_i^{(1)}}{\partial z_i^{(1)}} = \left(\sum_{k=1}^m \delta_k^{(2)} w_{ki}^{(2)} \right) S'(z_i^{(1)}) \quad (2.68)$$

Setzen wir nun noch eine sigmoidale Fermi-Ausgabefunktion S_F voraus mit der Ableitung

$$S'(z) = \frac{\partial}{\partial z} \frac{1}{1+e^{-z}} = \frac{+1-1+e^{-z}}{(1+e^{-z})^2} = (1-S(z))S(z) \quad (2.69)$$

so sind die Fehlerinkremente $\delta_i^{(2)}$ und $\delta_i^{(1)}$ bestimmt und wir haben damit die zwei Delta-Lernregeln, eine für die Ausgabeschicht

$$\Delta w_{ij}^{(2)}(x) = \gamma \delta_i^{(2)} x_j^{(2)} = -\gamma (y_i^{(2)} - L_i)(1 - S(z_i^{(2)})) S(z_i^{(2)}) x_j^{(2)} \quad (2.70)$$

und eine für die Schicht der "hidden units"

$$\Delta w_{ij}^{(1)}(x) = \gamma \left(\sum_k \delta_k^{(2)} w_{ki}^{(2)} \right) (1 - S(z_i^{(1)})) S(z_i^{(1)}) x_j^{(1)} \quad (2.71)$$

vollständig hergeleitet.

Für die allgemeinen Rekursionsformeln des Fehlerinkrements für eine beliebige Zahl von Schichten reicht es, in (2.68) die Ersetzung der $^{(2)}$ durch $^{(n)}$ und der $^{(1)}$ durch $^{(n-1)}$ vorzunehmen, so dass

$$\delta_i^{(n-1)} = - \left(\sum_k \delta_k^{(n)} w_{ki}^{(n)} \right) S'(z_i^{(n-1)}) \quad (2.72)$$

resultiert.

Mit den Lernregeln Gl.(2.70) und (2.71) errechnet der Algorithmus für jedes Trainingsmuster die notwendige Korrektur der Gewichte Δw und speichert sie zunächst nur als Summe. Erst nach dem letzten Trainingsmuster erfolgt dann tatsächlich die Korrektur der Gewichte mit

$$\mathbf{w}(t+1) = \mathbf{w}(t) + \Delta \mathbf{w} \quad (2.73)$$

Da die einzelnen Korrekturbeiträge auf der Basis der alten Gewichte ohne Kenntnis der bereits errechneten Korrekturen errechnet wurden, wird dies als *OFF-Line-Lernen* bezeichnet. Wollen wir im Unterschied dazu die Gewichte sofort nach jedem Trainingsmuster korrigieren (*ON-Line-Lernen*), so können wir statt Gl.(2.58) die stochastische Version verwenden

$$w_{ij}(t) = w_{ij}(t-1) - \gamma \partial R_x / \partial w_{ij} \quad (2.74)$$

Von der stochastischen Approximation wird uns eine gute Konvergenz garantiert, wenn wir anstelle einer konstanten Lernrate $\gamma = \text{const}$ eine andere Lernrate, beispielsweise $\gamma(t) = 1/t$, einsetzen. Allerdings gilt dies nur, wenn δ eine stochastische Variable mit zeitunabhängiger Verteilung ist. Das ist aber bei uns nicht gegeben: durch die Gewichtskorrektur verschiebt sich trotz konstanter Verteilung der \mathbf{x} die Verteilung der y und damit die des Fehlers δ . Es ist also deshalb sinnvoll, beispielsweise anfangs eine konstante Lernrate zu verwenden, bis keine großen Änderungen mehr stattfinden, und dann auf $\gamma(t) = 1/t$ überzugehen, um die Gewichtsänderungen zu null konvergieren zu lassen.

Was können wir von Backpropagation-Netzen erwarten? Der ursprüngliche, nicht-lineare Backpropagation-Algorithmus weist verschiedene Nachteile auf:

- ◆ Die Konvergenz ist relativ *langsam*, besonders bei mehreren Schichten
- ◆ Das System kann in einem *lokalen Optimum* "stecken" bleiben
- ◆ Trotz guter Trainingsleistung zeigt der Test *schlechte Ergebnisse*

Besonders der letzte Punkt verdient wegen seines prinzipiellen Hintergrunds gesonderte Aufmerksamkeit.

2.7 Training, Validierung, Testen

Wird ein System trainiert, so kann es bei guter Trainingsleistung „überangepasst“ (*overfitting*) werden, wenn es die Daten „speichert“ anstatt zu generalisieren. Wie muss man sich das vorstellen? In der folgenden Abb. 2.13 ist dies an der Approximation einer Funktion visualisiert. Die Funktion (durchgezogene schwarze Linie) ist nicht direkt bekannt, sondern nur einzelne, mit zufälligen Abweichungen behaftete Funktionswerte (schwarze Punkte). Wird das Netz nur mit diesen Punkten trainiert, so ist die Abweichung der Netzfunktion (gestrichelte Linie) davon bei Überanpassung sehr gering; der Trainingsfehler ist minimal. Beachten wir aber noch weitere, mit zufälligen Abweichungen versehene Beobachtungswerte (Kreise) der unbekannteten Funktion, so nimmt der auf der Menge der Testmuster (*Validierungsmenge*) beobachtete Fehler wieder zu, wenn die vermutete Funktion von der tatsächlichen Funktion abweicht und sich an die zufälligen Schwankungen anpasst.

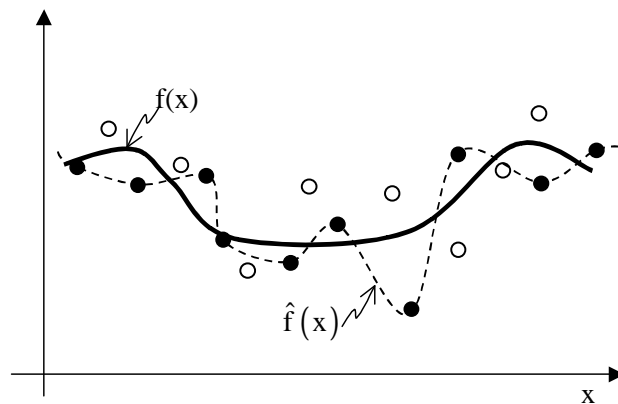


Abb. 2.13 Generalisierung und Überanpassung bei einer Funktionsapproximation

Wie kann man feststellen, ob im konkreten Fall eine Überanpassung vorliegt? Dazu testet man die Leistung des Netzes nach jeder Lernperiode ΔT mit Daten, die vom Training unabhängig sein sollen. Erhöht sich der Fehler nach einer Trai-

ningsepoche wieder, so liegt eine Überanpassung vor; das Training sollte gestoppt werden („*stopped training*“, „*early stopping*“), siehe Abb. 2.14. Hier ist die Fehlerentwicklung eines Netzes während des Trainings gezeigt, verdeutlicht an dem Fehler, den ein Trainingsmuster beim Lernschritt t macht. Der Fehler, den das trainierte Netz auf der gesamten Validierungsmenge erbringt, ist in Abständen ΔT jeweils als schwarzer Punkt ebenfalls eingetragen. Misst man ihn nicht in großen Schritten, sondern nach jedem Lernschritt (was sehr aufwendig ist), so ergibt sich die durchgezogene Kurve. Obwohl der Trainingsfehler deutlich abnimmt, kann der Fehler dagegen auf einer unabhängigen Validierungsmenge zunehmen, sobald eine Überanpassung durchgeführt wird. Der Zeitpunkt, an dem der Fehler der Validierungsmenge am geringsten ist, bietet sich dazu an, das Training abubrechen. Um die tatsächliche Güte des trainierten Netzes zu bestimmen sollte ein einheitlicher Test für alle Versionen des Netzes existieren, der auch von der Validierung unabhängig ist: der Test mit einer unabhängigen Mustermenge, der Testmenge.

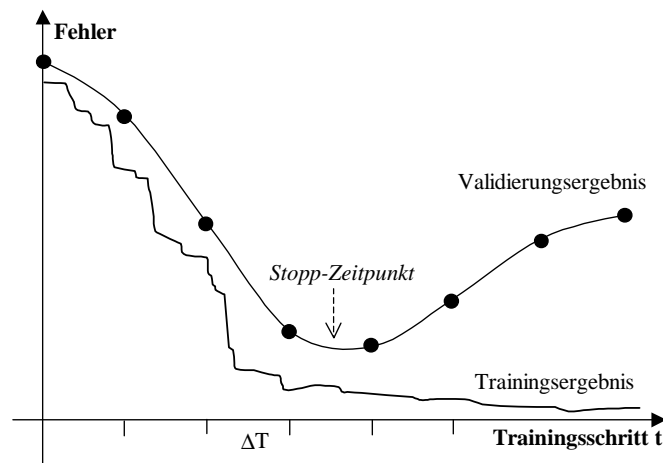


Abb. 2.14 Überanpassung beim Lernen

In diesem Fall wird die Gesamtmenge an Mustern in drei Mengen eingeteilt: die Trainingsmenge, die Validierungsmenge und eine weitere Testmenge. Leider hat man für die Aufteilung in drei Mengen nicht immer genügend Muster zur Verfügung: So ist es sicher problematisch, eine Gruppe von 15 Patienten in drei Gruppen je 5 Patienten aufzuteilen: Die Ergebnisse sind einfach zu unzuverlässig; die Daten jedes Patienten sind für das Training notwendig.

Wie können wir trotz solch widriger Bedingungen trotzdem noch unsere Qualitätskontrolle von getrenntem Training und Test beibehalten? In dieser Notlage hilft uns ein Verfahren: die Kreuzvalidierung (*cross validation*). Dazu teilen wir die Gesamtmenge von N Mustern auf in k Untermengen. Von den k Untermengen wird eine ausgewählt als Testmenge, die anderen $k-1$ werden zum Training verwendet. Das Ergebnis des abschließenden Tests wird notiert. Nun wird eine andere Menge als Testmenge ausgewählt, wieder trainiert und getestet. Insgesamt erhalten wir nach k Auswahlen k Ergebnisse. Die Mittelung der Ergebnisse gibt uns die erwartete Güte eines Trainings, also die erwartete Leistung des Systems auf unbekanntem Daten.

Der Fehler auf der Validierungsmenge bzw. Testmenge kann von der zufälligen Zusammensetzung der Validierungsmenge abhängen. Aus diesem Grund kann man jede der k Untermengen so auswählen, dass die Verteilung der Muster innerhalb der Untermenge jeweils ähnlich ist, so dass auch die Ergebnisse nicht so stark variieren. Eine solche Validierung heisst „stratifizierte Kreuzvalidierung“.

Es gibt viele Arbeiten in der Literatur für jeden der anderen erwähnten Kritikpunkte, um mit einem geänderten Algorithmus die Ergebnisse zu verbessern. Wir wollen dies hier nicht näher betrachten, sondern eine Alternative dazu betrachten: die RBF-Netze.

2.8 RBF-Netze

Die Netzarchitektur aus zwei Schichten lässt sich formal durch Gl. (2.53) beschreiben

$$\hat{f}(\mathbf{x}) = \sum_{j=1}^m w_j^{(2)} S(\mathbf{w}_j, \mathbf{x})$$

Dabei nehmen die Funktionen $S(\mathbf{w}_j, \mathbf{x})$ die Rolle von Basisfunktionen an, deren lineare Überlagerung die Approximation bilden. Solche Basisfunktionen sind in der Mathematik gut bekannt und erforscht. In unserem Fall interessieren wir uns für eine besondere Sorte von Basisfunktionen: solche, die im Unterschied zu den klassischen sigmoidalen Basisfunktionen der Backpropagation-Netze nur lokal sensibel sind, deren Ausgabe bei steigendem Abstand („Radius“) von einem Punkt \mathbf{w}_j sich ändert

$$S(\mathbf{w}_j, \mathbf{x}) = S(|\mathbf{w}_j - \mathbf{x}|) \quad \text{Radiale Basisfunktionen} \quad (2.75)$$

Eine typische Kategorie derartiger radialer Basisfunktionen sind die *Glockenfunktionen* (*bell shaped functions*). Sie sind im \pm Unendlichen jeweils null und haben an einer Stelle a einen endlichen, positiven reellen Wert. Formal lässt sich dies folgendermaßen definieren:

Eine Funktion S_G mit den Eigenschaften

- $S_G(x) > 0$ für reelles x , $S_G(-\infty) = S_G(\infty) = 0$,
- $\int_{-\infty}^{\infty} S_G(x) dx < \infty$ und $\neq 0$
- Es ex. ein reelles a mit $S_G(a) \neq 0$ und $S_G(z)$ nicht anwachsend $\forall z \in [a, \infty)$
nicht abfallend $\forall z \in (-\infty, a)$

nennen wir eine *Glockenfunktion*.

Für Glockenkurven nach obiger Definition kann man zeigen, dass ein Netzwerk, das aus linear überlagerten, normierten Glockenfunktionen

$$\hat{f}(\mathbf{x}) = \sum_i w_i \tilde{S}_i(\mathbf{x}, \mathbf{c}_i) \quad (2.76)$$

mit $\tilde{S}_i(\mathbf{x}, \mathbf{c}_i) = S_i(\mathbf{x}, \mathbf{c}_i) / (\sum_k S_k(\mathbf{x}, \mathbf{c}_k))$

besteht, jede beliebige Funktion beliebig dicht approximieren kann [Cardaliaguet 1992]. Wenn die Gewichte dabei beschränkt bleiben, werden auch die Ableitungen gleichzeitig beliebig dicht approximiert.

Dies spielt in Netzen für Kontrollaufgaben eine wichtige Rolle. Das dazugehörige Netzwerk ist in Abb. 2.15 gezeigt, wobei die RBF-Einheiten mit dem Symbol " Ω " gekennzeichnet sind. Das ganze Netzwerk ist ein Zwei-Schichten-Netz; die erste Schicht enthält als Parameter die Zentren \mathbf{c}_i und die Breite σ der Basisfunktionen, und die zweite Schicht die Gewichte w_i .

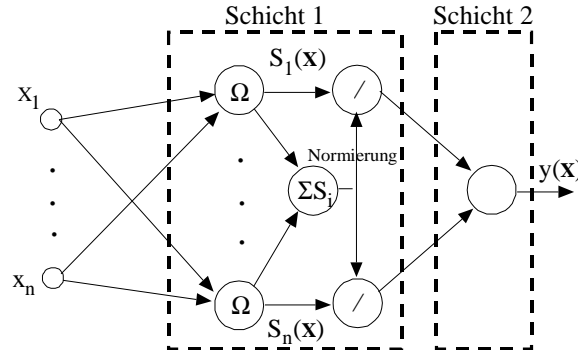


Abb. 2.15 Das Netzwerk normierter Basisfunktionen

Nachdem wir eine allgemeine Definition von Glockenfunktionen und ihre wichtige Approximationseigenschaft kennen gelernt haben, stellt sich für uns

immer noch die Frage: Wie wählen wir die Zentren und Ausdehnungen der Glockenfunktionen? Wie verbessern wir die Gewichte der zweiten Schicht?

2.8.1 Approximation mit Glockenfunktionen

Angenommen, wir modellieren für einen n -dimensionalen Argumentwert \mathbf{x}^i den Funktionswert $f(\mathbf{x})$ als Mittelwert einer Variablen y mit einer unbekanntem Dichtefunktion $p(\mathbf{x}, y)$. Dann ist dies der bedingte Erwartungswert der Zufallsvariablen y , wenn \mathbf{x}^i gegeben ist

$$f(\mathbf{x}) = \langle y | \mathbf{x}^i \rangle = \int y p(y | \mathbf{x}^i) dy = \frac{\int y p(y, \mathbf{x}^i) dy}{p(\mathbf{x}^i)} = \frac{\int y p(y, \mathbf{x}^i) dy}{\int p(y, \mathbf{x}^i) dy} \quad (2.77)$$

Fassen wir beide Zufallsvariablen zu einer gemeinsamen Variablen $\mathbf{z} = (\mathbf{x}, y)$ zusammen, so ist eine gute Approximation (*Schätzer*) $\hat{p}(\mathbf{z})$ für $p(\mathbf{z})$, wie wir oben sahen, die Überlagerung der Gauß'schen Glockenfunktionen

$$\hat{p}(\mathbf{x}, y) = \hat{p}(\mathbf{z}) = \frac{1}{M} \sum_{i=1}^M \frac{1}{[(2\pi)^{1/2} \sigma]^{n+1}} \exp(-(\mathbf{z}-\mathbf{z}^i)^T(\mathbf{z}-\mathbf{z}^i)/2\sigma^2) \quad (2.78)$$

Mit den beobachteten Werten $y^i := y(\mathbf{x}^i)$ und der Abkürzung

$$d_i^2 := (\mathbf{x}-\mathbf{x}^i)^T(\mathbf{x}-\mathbf{x}^i)$$

$$\text{ist} \quad (\mathbf{z}-\mathbf{z}^i)^T(\mathbf{z}-\mathbf{z}^i) = d_i^2 + (y-y^i)^2 \quad (2.79)$$

so daß nach dem Einsetzen von Gl.(2.79) in Gl.(2.78), dem Hinausziehen der Summen aus den Integralen und unter Beachtung von der Normierung $\int p(y) dy = 1$ und dem Erwartungswert $\int y p(y, y^i) dy = y^i$ der geschätzte Approximationswert lautet

$$\hat{f}(\mathbf{x}^i) = \sum_i y^i \frac{\exp(-d_i^2/2\sigma^2)}{\sum_k \exp(-d_k^2/2\sigma^2)} = \sum_i y^i \frac{S_i(\mathbf{x})}{\bar{S}} \quad (2.80)$$

Die Gl. (2.80) läßt sich wieder als zweischichtiges Netzwerk deuten [Specht,1991], [Schioler et al., 1992]. Die erste Schicht zur Approximation der unbekanntem Funktion f enthält Neuronen mit Glockenkurven als Ausgabefunktionen $S_G(\mathbf{x}, \sigma)$, die, mit bekannten, beobachteten Werten y^i gewichtet, in der zweiten Schicht aufsummiert werden. Die Ausgabefunktionen sind als *bedingte Wahrscheinlichkeitswerte* zusätzlich noch mit \bar{S} normiert, so daß $\sum_i S_i(\mathbf{x}) = 1$ gilt. Eine Klassifikation, die auf der maximalen bedingten Wahrscheinlichkeit (*Bayes-Klassifikation*) beruht, ist mit einem solchen Netz leicht zu implementieren.

Es sind auch Netzwerke mit dem einfachen Abstandsbetrag (L_1 -Abstandsmaß) anstatt des quadratischen Euklid'schen Abstands (L_2 -Abstandsmaß) denkbar.

Diese Art von Neuronen ersparen bei numerischer Berechnung das aufwendige Quadrieren der Komponenten in der Exponentialfunktion [Specht, 1991].

Ist eine ein-eindeutige Abbildung zwischen Eingabe und Ausgabe gegeben, so lässt sich genauso gut auch die Umkehrfunktion lernen – ein wichtiger Vorteil für Kontrollaufgaben [Specht, 1991].

Die Normierung S_i/S der Ausgabeaktivität in Gl. (2.80) ist dabei ziemlich wichtig. In Abb. 2.16(a) ist die Ausgabe $y = w_1S_1(\mathbf{x}) + w_2S_2(\mathbf{x}) + w_3S_3(\mathbf{x})$ eines RBF-Netzes aus drei RBF-Neuronen in der ersten Schicht und einem linearen Neuron in der zweiten Schicht gezeigt.

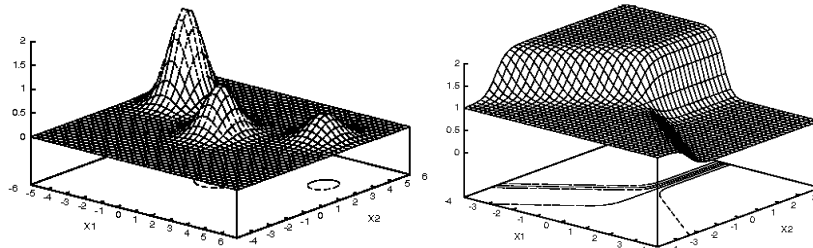


Abb. 2.16 Einfache (a) und normierte Netzaktivität (b) aus drei RBF Neuronen

Normieren wir zusätzlich die Ausgabe $S_i = S_G(\mathbf{x})$ zu $\tilde{S}_i = S_i/(S_1+S_2+S_3)$, so ergibt sich das Verhalten in Abb. 2.16(b). Hier ist deutlich zu sehen, dass die normierte, in der zweiten Schicht mit w_i gewichtete Ausgabe eines RBF Neurons über den gesamten "Zuständigkeitsbereich" konstant ist. Die Grenze zum benachbarten Zuständigkeitsbereich definiert sich ziemlich scharf durch den Aktivitätsquotienten (die relative Aktivität), die an der Grenze schnell auf Null abfällt. Damit ist es im Unterschied zu (a) relativ einfach, eine Hyperfläche zu approximieren.

Die Ansätze des vorigen Abschnitts sind nur bei einer kleinen Anzahl von Trainingsmustern sinnvoll. Bei einer größeren Zahl von Mustern kann man von den Einzelmustern absehen und nur noch wenige, aber typische Basisfunktionen verwenden, die nur durch wenige Parameter festgelegt sind und damit eine Datenreduktion (*Generalisierung*) bedeuten. Hier besitzen die Basisfunktionen nicht nur unterschiedliche Mittelwerte c^i , sondern auch unterschiedliche Varianzen σ_i , die iterativ gelernt werden müssen. Hauptanliegen dieses Abschnittes ist es, dafür geeignete Verfahren zu zeigen.

2.8.2 Normierung der Eingabevariablen

Bei vielen Problemen aus der realen Welt ist von vornherein nicht klar, welche die adäquate Skalierung der Variablen ist. Sollte man eine Länge in *Metern* oder in *cm* messen, einen Druck in *Pascal* oder *Hektopascal*? Und wenn wir uns auf eine Einheit festgelegt haben, welches ist dann die passende andere der zweiten Vari-

ablen? Und der dritten, vierten ,usw.? Beachten wir dies nicht, so werden die Punkthaufen im Musterraum zu Elypsen verzerrt; für ein schnelles Lernen der Parameter müssen die Lernschritte dann in den unterschiedlichen Dimensionen der Variablen unterschiedlich groß sein. Diese Komplikation können wir vermeiden, indem wir alle Variablen so normieren, dass ihr Mittelwert den Wert null und ihre Varianz den Wert eins annimmt. Die Zusammenhänge der Variablen werden dadurch nicht berührt, nur ihre Darstellung.

Jede Variable wird also vor der eigentlichen Verarbeitung mit der Transformation $d = (\mathbf{x}-\mathbf{c})/\sigma$ um ihren Mittelwert c und ihre mittlere Abweichung σ korrigiert. Im multidimensionalen, allgemeinen Fall ist die Länge d^2 der um den Punkt c zentrierten und auf die Varianz eins skalierten Zufallsvariablen \mathbf{x}

$$d^2 = (\mathbf{x} - \mathbf{c})^T \mathbf{C}^{-1} (\mathbf{x} - \mathbf{c}) \quad \text{Mahalanobis Abstand} \quad (2.81)$$

mit \mathbf{C}^{-1} als der Inversen der Kovarianzmatrix $\mathbf{C} = \langle (\mathbf{x} - \mathbf{c})(\mathbf{x} - \mathbf{c})^T \rangle$ gegeben. Geometrisch lässt sich dies so interpretieren, dass die Variable \mathbf{x} einer Verschiebung um \mathbf{c} und einer Drehung und Skalierung durch eine Matrix (hier \mathbf{C}^{-1}) unterworfen wird. Stimmen die Koordinatenachsen mit den Hauptachsen der Punktwolke $\{\mathbf{x}\}$ überein, so ist $\mathbf{C}^{-1} = \Lambda^{-1}$, eine Diagonalmatrix. Die Hauptdiagonale besteht aus den Eigenwerten $\lambda_i^{-1} = \sigma_{ii}^{-2}$: Es findet dann keine Drehung mehr statt, sondern nur noch eine Skalierung mit $1/\sigma_{ii}^2$.

Der mit Gl. (2.81) korrigierte Abstand $(\mathbf{x} - \mathbf{c})$ zwischen dem Punkt \mathbf{x} und dem Zentrum \mathbf{c} wird als *Mahalanobis* Entfernung $d_M(\mathbf{x}, \mathbf{c})$ bezeichnet und stellt einen statistisch korrigierten Euklid'schen Abstand dar. Die Gaußsche Glockenfunktion wird in diesem Fall

$$S_G(\mathbf{x}) = A \exp(-(\mathbf{x} - \mathbf{c})^T \mathbf{C}^{-1} (\mathbf{x} - \mathbf{c})) \quad (2.82)$$

wobei die Normierungsbedingung für eine n -dimensionale Normalverteilung den Wert $A = [(2\pi)^n \det \mathbf{C}]^{-1/2}$ bestimmt.

Die Transformation eines Musterhaufens auf eine zentrierte Verteilung mit gleicher Varianz in allen Richtungen ist in Abb. 2.17 illustriert. Hierbei sind alle Punkte \mathbf{x} einer zwei-dimensionalen Gaußverteilung, die den gleichen Funktionswert $S_G(\mathbf{x})$ ergeben ("Höhenlinien"), durch eine Linie verbunden. Wie man sich leicht klarmachen kann, ist dies im Allgemeinen eine Elypse, die nach der Transformation in einen Kreis mit dem Einheitsradius übergeht. Die Abbildung zeigt aus der Vogelperspektive, wie die Transformation Gl.(2.81) sich auch formal durch eine allgemeine Verschiebung, Drehung und Skalierung erreichen lässt.

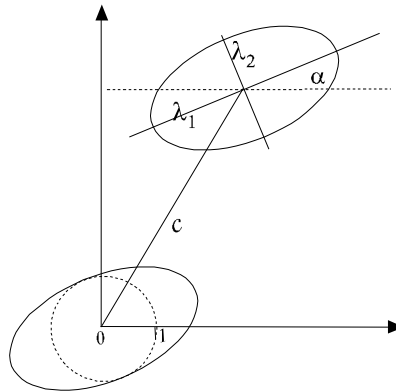


Abb. 2.17 Transformation auf Normalverteilung

Die Skalierung \mathbf{S} und die Drehung \mathbf{D} lassen sich als eine Matrizenmultiplikation formulieren, ebenso die Verschiebung \mathbf{V} , wenn wir die Eingabe \mathbf{x} um eine konstante Komponente $\mathbf{x} \rightarrow (\mathbf{x}^T, 1)^T$ erweitern. Im zwei-dimensionalen Fall ist dies also

$$\mathbf{x} \rightarrow \mathbf{z} = \mathbf{SDVx} \quad \text{mit} \quad \mathbf{S} = \begin{pmatrix} s_1 & 0 \\ 0 & s_2 \end{pmatrix}, \mathbf{D} = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}, \quad (2.83)$$

$$\mathbf{V} = \begin{pmatrix} 1 & 0 & -c_1 \\ 0 & 1 & -c_2 \end{pmatrix}$$

oder $\mathbf{x} \rightarrow \mathbf{z} = \mathbf{Mx}$ mit

$$\mathbf{M} = \begin{pmatrix} s_1 \cos \alpha & -s_1 \sin \alpha & -c_1 s_1 (\cos \alpha - \sin \alpha) \\ s_2 \sin \alpha & s_2 \cos \alpha & -c_2 s_2 (\cos \alpha + \sin \alpha) \end{pmatrix}$$

und die Transformation Gl.(2.81) wird zu

$$d^2 = \mathbf{z}^2 = \mathbf{z}^T \mathbf{z} = \mathbf{x}^T \mathbf{M}^T \mathbf{M} \mathbf{x} \quad (2.84)$$

Hier kann der Mahalanobis-Abstand mit \mathbf{Mx} nicht nur als die Auswirkung der Transformation \mathbf{M} gesehen werden, sondern es ist die geometrische Bedeutung der einzelnen Koeffizienten M_{ij} der Matrix \mathbf{M} direkt klar. Beispielsweise ist mit $\cos^2 \alpha + \sin^2 \alpha = 1$ der Skalierungskoeffizient $s_1 > 0$ durch $M_{11}^2 + M_{12}^2 = s_1^2$ gegeben. Dadurch sind über $\cos \alpha$, $\sin \alpha$, der Drehwinkel α , und damit auch die Verschiebung c_1 , bestimmt.

Mit der Aufbereitung der Eingabe auf Mittelwert null und Varianz eins haben wir nun die Voraussetzungen geschaffen, um mit einem Lernverfahren alle Eingabe-

bedimensionen einheitlich zu behandeln und daraus die gewünschten Parameter zu bestimmen.

2.9 RBF-Lernverfahren

Ein RBF-Netzwerk besteht meist aus zwei oder mehr Schichten. Jede Schicht hat ihre Parameter, die ihre Funktion festlegen. Die Lernverfahren, mit denen diese Parameter bestimmt werden können, lassen sich in zwei Ansätze unterteilen:

- 1) dem Anpassen zuerst der ersten Schicht, also Lage \mathbf{c}_i und Varianz \mathbf{C} der RBF-Neuronen, und dann Anpassen der Gewichte \mathbf{w}_j nächsten Schicht bei konstanter Lage,
- 2) dem gleichzeitigen Anpassen aller Parameter aller Schichten bei gegebenen Anzahlen der RBF-Neuronen in der ersten Schicht und der Neuronen in den weiteren Schichten.

Die getrennte Optimierung der Parameter der ersten und zweiten Schicht hat verschiedene Vorteile. So lassen sich zum einen verschiedene, voneinander unabhängige Methoden zur Optimierung der ersten und der zweiten Schicht einsetzen, zum anderen ist die Konvergenz dieser Verfahren für die Anpassung einer einzelnen Schicht deutlich besser als bei der gleichzeitigen Anpassung beider Schichten. Dies hängt damit zusammen, dass der Suchraum bei beiden Schichten (die Anzahl aller möglichen Parameterwertkombinationen) exponentiell größer ist als der Suchraum bei Beschränkung auf die Parameter nur einer Schicht; die Dimensionszahl beider Suchräume addieren sich.

2.9.1 Anpassung der Parameter der ersten Schicht

Ein gängiger Ansatz für die erste Schicht besteht darin, die Wahrscheinlichkeitsdichten der Eingabemuster durch die Lage der RBF-Neuronen zu approximieren. Ein Lernalgorithmus für die Approximation mit Glockenfunktionen besteht grundsätzlich aus zwei Schritten:

- einer initialen Verteilung (Anzahl, Lage und Form) der Glockenfunktionen
- der iterativen Adaption der Parameter an die Trainingsdaten

Während die Adaption der Parameter relativ leicht durchzuführen ist, ist die initiale Phase für das Lernen entscheidend und der häufigste Grund, wenn das System nach dem Training nicht die gewünschte Leistung zeigt, sondern in lokalen Minima "hängen" bleibt oder nur langsame bzw. keine Konvergenz aufweist.

Wichtigstes Ziel der Initialisierung ist eine Aufteilung der "Datenwolke" in unterschiedliche Haufen (*cluster*), die jeweils von einer RBF "besetzt" werden. Als

Initialisierung ist also eine parallele (OFF-LINE) oder sequentielle (ON-LINE) Clusteranalyse nötig. Für eine effiziente Initialisierung müssen wir nun unterscheiden, ob die Trainingsdaten bekannt oder unbekannt sind.

- **Bekannte Trainingsdaten**

In diesem Fall können wir eine grobe, statistische Vorverarbeitung der Gesamtmenge der Trainingsdaten durchführen. Dazu wird zuerst mit einem einfachen Clustersuchalgorithmus (z.B. k-mean) eine Unterteilung der Daten in Haufen (*cluster*) vorgenommen. Bei Klassifizierungsproblemen wird dies durch die Annahme der Existenz von Muster-Haufen gleicher Klassenzugehörigkeit gerechtfertigt. Dann wird in wenigen Rechner-Minuten die Streuung (Eigenwerte) der Cluster bestimmt und damit die Parameter \mathbf{M} bzw. \mathbf{c} und \mathbf{C} der dazu gehörenden Glockenfunktionen initialisiert.

Grundalgorithmus k-mean-Clusterung

Bei der *k-mean-Clusterung* werden Cluster aus einer Menge A von vorhandenen Daten gebildet und schrittweise verfeinert. Dabei werden folgende Schritte durchgeführt:

- Wähle m zufällige Muster $\mathbf{x}_k \in A$ als Clusterzentren \mathbf{c}_k , bilde Cluster $C_k = \{ \mathbf{c}_k \}$.

REPEAT

- Ordne alle $\mathbf{x}_i \in A$ zu den nächstgelegenen Clusterzentren zu: Suche \mathbf{c}_z so, dass $|\mathbf{x}_i - \mathbf{c}_z| = \min_k |\mathbf{x}_i - \mathbf{c}_k|$, und füge \mathbf{x}_i zu C_z zu.
- Entferne alle Cluster mit $|C_i| \leq 1$
- Bilde für jedes Cluster k ein neues Zentrum $\mathbf{c}_k = \langle \mathbf{x} \rangle$ als Mittelwert aller Muster in C_k

UNTIL Iterationszahl > Max

- **Unbekannte Trainingsdaten**

Der einfachste Ansatz einer sequentiellen Clusteranalyse besteht darin, die initiale Verteilung der Glockenfunktionen so zu wählen, dass der Raum der Eingabemuster $\{\mathbf{x}\}$ entweder systematisch in periodischen Abständen durch die Glockenfunktionen überdeckt wird. In Abb. 2.18 ist dieser Gedanke verdeutlicht.

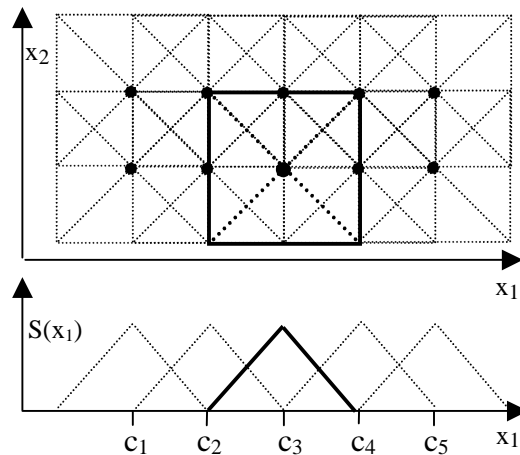


Abb. 2.18 Überdeckung des Eingaberaums durch genormte dreieckige Glockenfunktionen

Jede Dreiecks-Glockenfunktion verdeutlicht die Lage und die Ausdehnung des rezeptiven Feldes eines potentiellen RBF-Neurons. Am Anfang gibt es kein Neuron im System. Tritt nun ein Eingabemuster \mathbf{x} auf, das keinem existierendem Neuron zuzuordnen ist, so wird die Glockenfunktion, die maximalen Wert $S(\mathbf{x})$ hat, gewählt, und an der Stelle ein RBF-Neuron platziert. Auf diese Weise werden nur die tatsächlich benötigten Neuronen auf einer Liste gehalten; alle nicht aktivierten Glockenfunktionen werden nicht berücksichtigt. Die Gesamtapproximation ergibt sich dann als gewichtete Überlagerung aller Glockenfunktionen, wobei die Gewichte die Häufigkeit der Aktivierung der einzelnen Funktion angibt, s. [Alexandridis 2003].

Dieser Ansatz vermeidet zusätzliche Anpassarbeiten wie das Justieren der Breite und Lage der RBF. Stattdessen gibt es nur Standard-RBF, die eine Standard-Größe und Standard-Ort besitzen.

Ein anderer Ansatz besteht darin, zunächst wenigen Glockenfunktionen großer Reichweite (großer Überdeckung, d.h. großem σ) anzufangen und verkleinert dann die Radien und erhöht die Anzahl der Funktionen. So kann sich das System mit diesem sukzessiven Netzaufbau die Wahrscheinlichkeitsdichten approximativ "einfangen". In Abb. 2.19 ist die iterative Verkleinerung der Radien und Erhöhung der Neuronenzahl für eine handgeschriebene Zahl "3" gezeigt.

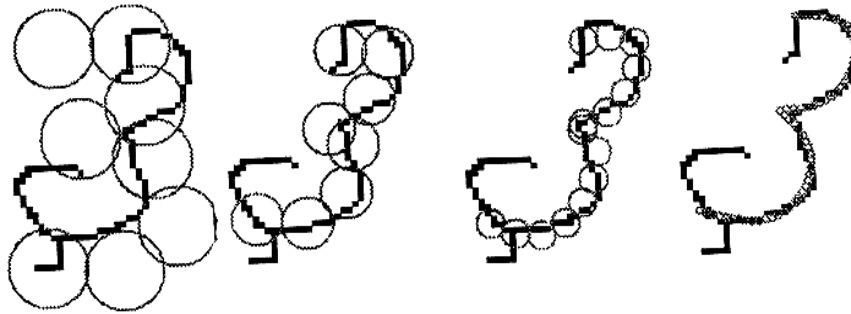


Abb. 2.19 Annäherung von Anzahl und Überdeckung von Glockenfunktionen an die Intensitätsverteilung eines Bildes (nach [Hinton1992]).

In beiden Situationen, bei bekannten und unbekanntem Daten, lässt sich die Komplexität des Netzwerks und die Trainingszeiten deutlich verringern, wenn man anstelle von einer festen Anzahl von Neuronen das Netzwerk durch die sukzessive Erzeugung neuer Glockenfunktionen (Hinzufügen von Neuronen) verbessert. Eine typische Methode ist die der *sequentiellen Regression*.

Überdeckung durch Fehlerminimierung

Dazu wird versucht, nur mit wenigen Neuronen anzufangen (ideal mit einem) und dann inkrementell weitere hinzuzufügen. Kriterium für die Lage der neuen RBF-Neuronen bei diesem Verfahren ist der Punkt, an dem der größte Fehler gemacht wird. Vermutet man an Orten mit größerem mittleren Fehler auch mehr Information (z.B. bei der Adaption von RBF-Bereichen an Bildkonturen, s. Abb. 2.19), so ist die Erzeugung und damit die Konzentration neuer Neuronen gerade an diesen Häufigkeitspunkten sinnvoll.

Bei bekannten Daten kann man nach jedem Trainingslauf mit allen Mustern den Ort des größten Fehlers des Netzes feststellen und dort ein Neuron platzieren. Sein Gewicht in der zweiten Schicht wird entweder so gewählt, dass es nur den Fehler an dieser Stelle kompensiert, oder aber durch eine Minimierung des quadratischen Fehlers mit allen Gewichten der zweiten Schicht. Das Training bricht ab, wenn der maximale Fehler klein genug ist.

Gibt man einen festen Fehlerwert als Schranke vor, bei dessen Überschreiten jeweils ein Neuron eingefügt wird, so kann man bei jedem Trainingslauf auch mehrere Neuronen zum Netz hinzufügen. Dazu kann man auch die Strategie einer Fehlerschranke mit der einer Abstandsschranke (Überschreiten des Abstands $|\mathbf{x} - \mathbf{c}_m|$ der Eingabe \mathbf{x} vom Zentrum \mathbf{c}_m des nächsten Neurons) kombinieren, um eine einigermaßen konsistente Überdeckung des Eingaberaumes zu erreichen. Diese

Art von Strategien empfiehlt sich auch bei sequentieller Approximation unbekannter Daten, wobei mit fortlaufender Zeit auch die Schranken erniedrigt und somit die Ansprüche erhöht werden können.

Ein überwachter Clusteralgorithmus zum Erzeugen und Anpassen von Glockenkurven ist beispielsweise auch von (Lee und Kil, 1991) vorgestellt worden. Hier werden solange neue Glockenfunktionen erzeugt, bis alle Eingabemuster innerhalb eines sog. "effektiven Radius" fallen und somit durch eine Glockenfunktion "abgedeckt" werden. Tritt keine Verkleinerung des Fehlers mehr ein ("Sättigung") und ist der Fehler noch zu hoch, so werden die Weiten σ der Glockenfunktionen verkleinert und Training und Erzeugung neuer Funktionen erneut durchgeführt, bis der vorgegebene Maximalfehler unterschritten wird.

2.9.2 Anpassen der Parameter der zweiten Schicht

Die Approximation mit Radialen Basisfunktionen verbindet verschiedene Mechanismen miteinander. In der ersten Schicht werden Muster aus dem Eingaberaum auf Reaktionen von RBF-Neuronen abgebildet. Durch ihre Eigenschaften wird trotz geringen Variationen die Eingabe auf gleiche Ausgabe abgebildet; dies bedeutet eine Tolerierung von Abweichungen. Beispielsweise wird eine Linie im zwei-dimensionalen Eingaberaum bei radialsymmetrischer RBF immer auf die gleiche Ausgabe abgebildet, egal in welchem Winkel sie durch den Eingaberaum des RBF-Neurons verläuft, was große Vorteile, z.B. bei der Erkennung von handgeschriebenen Buchstaben, bedeutet. Die erste Schicht lässt sich dabei als deformationsinvariante Mustervorverarbeitung betrachten, die von der zweiten Schicht, einem adaptiven Filter, weiterverwendet wird.

Für die Verbesserung der Gewichte w_i dieser zweiten, linearen Schicht lässt sich nun, wie für einen adaptiven, linearen Filter üblich, beispielsweise die Backpropagation-Lernregel (2.70) verwenden, sich für lineare Neuronen zu

$$\Delta w_{ij}^{(2)}(\mathbf{x}) = -\gamma (y_i^{(2)} - L_i) x_j^{(2)} \quad (2.85)$$

verkürzt.

Eine andere Möglichkeit ist die Minimierung des TLSME aus Gl.(2.19). Hierbei lernt ein Neuron mit einem Gradientenabstieg, d.h. einer negativen Hebb-Regel (*Anti-Hebb-Regel*) bei normierten Gewichten, den Eigenvektor mit dem kleinsten Eigenwert für das Minimum der Varianz. Für den stochastischen Fall ist dies

$$\tilde{\mathbf{w}}(t+1) = \mathbf{w}(t) - \gamma_1 \mathbf{x} \mathbf{y} \quad \text{und} \quad \mathbf{w}(t+1) = \frac{\tilde{\mathbf{w}}(t+1)}{|\tilde{\mathbf{w}}(t+1)|} \quad (2.86)$$

wobei mit $\mathbf{x} := (x_1, \dots, x_n, L)$ und $\mathbf{w} := (w_1, \dots, w_n, w_{n+1})$ für die Neuronenausgabe $y = S(\mathbf{x}) := g(\mathbf{x}) = \mathbf{x}^T \mathbf{w} - w_0$ geschrieben wird. Die konstante Nullpunktverschiebung w_0 lässt sich ebenfalls als Mittelwert lernen

$$w_0(t+1) = w_0(t) + \gamma_2 (\mathbf{x}^T \mathbf{w} - w_0(t)) \quad (2.87)$$

Da die Ausrichtung von \mathbf{w} stark von w_0 abhängig ist, sollte in der Lernphase die Zentrierung der Eingabe (Nullpunktverschiebung) mit $\gamma_1 \ll \gamma_2$ wesentlich schneller erfolgen als das Lernen des Eigenvektors \mathbf{w} .

Für das Lernen von Klassenentscheidungen können wir auch das Verfahren des "*Competitive Learning*" verwenden. Hier gibt es für jede Klasse ein Ausgabeneuron in der zweiten Schicht. Abhängig vom Erfolg wird der nur Gewichtsvektor desjenigen Neurons der zweiten Schicht verbessert, der bei Eingabe eines Musters die größte Aktivität bewirkt. Damit werden die Neuronen der zweiten Schicht auf Ereignisse "spezialisiert"; sind es binäre Neuronen, so lernen sie die Klassenprototypen den Ereignisklassen zuzuordnen.

2.10 Anwendungsbeispiel: Reaktormodell

Für die Modellierung einer Kopolymerisierung (Bildung eines Kunststoffes) aus MMA (Methyl-Methanacrylat) und VAc (Vinylacetat) wurde von (Bomberger et al. 2001) ein Simulationsmodell aus 12 Zuständen und 27 verschiedenen Reaktionen aufgestellt, aus dem als Referenz insgesamt 360 Daten der Menge M erstellt wurden. Zwar könnte man bei einem simulierten System mehr Daten gewinnen, aber bei realen Systemen sind normalerweise nicht wesentlich mehr Daten verfügbar. In Abb. 2.20 ist ein Schema des Prozesses gezeigt.

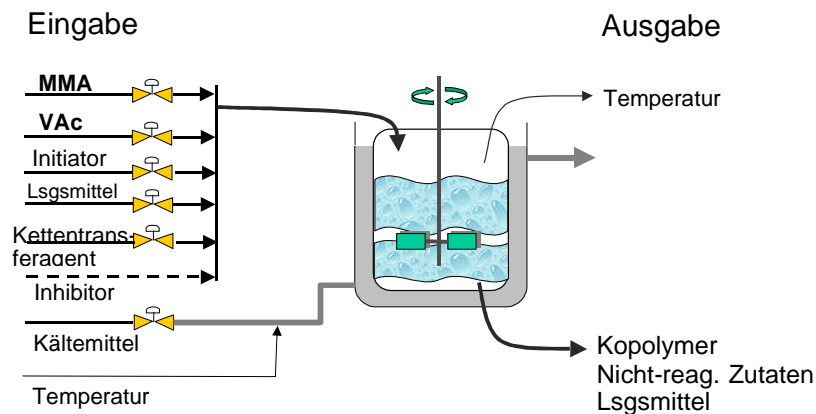


Abb. 2.20 Die Reaktorsimulation

Die Daten bestehen aus 8 Eingabevariablen (MMA, VAc gelöst in Benzen, den Initiator AIBN Azobisisobutyronitril, den Kettentransferagenten Acetaldehyd, die

Kühlmittelzuflusstemperatur, die Aussentemperatur und den Inhibitor *m*-Dinitrobenzen als Verunreinigung von MMA und VAc) und 4 Ausgabevariablen (das resultierende Kopolymer, die nicht reagierten Eingabezutaten, das Lösungsmittel und die Reaktorinnentemperatur), die unterschiedlich gewonnen werden können. Generell werden die Eingabedaten zu diskreten, äquidistanten Zeitpunkten erzeugt und die dazu gehörenden Ausgabewerte aus der Simulation gewonnen. Dabei stellt sich die Frage, ob und welcher Eingabewert zu einem Zeitpunkt generiert wird. Zwei verschiedene Methoden wurden implementiert:

- RUS: Für jeden Zeitpunkt gibt es eine Wahrscheinlichkeit < 1 , den Eingabewert zu ändern. Der neue Wert wird aus einer uniformen Zufallsverteilung zwischen Maximal- und Minimalwert gewonnen.
- RDS: Der Eingabewert wird zu jedem in Frage kommenden Zeitpunkt erzeugt, wobei der Wert aus einer endlichen Menge aus diskreten Eingabewerten gewonnen wird.

Die erste Strategie hält bei einigen Zeitpunkten die Eingabe konstant, so dass sie sich nicht zu schnell ändert. Die zweite Strategie wählt nur aus bereits existierenden Werten aus.

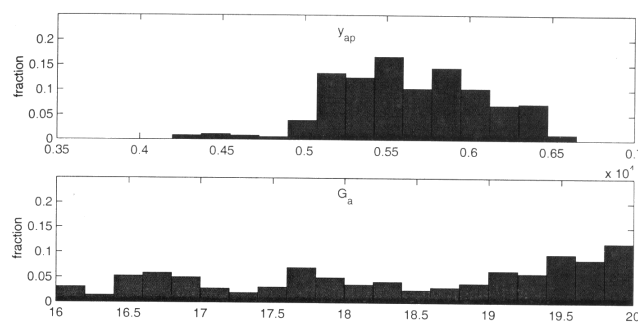


Abb. 2.21 RUS Strategie: Ausgabe (oben) und Eingabe (unten)

In der Abbildung ist die aus der RUS-Strategie erstellte Eingabe und Ausgabe als Histogramm dargestellt. Vergleicht man dies mit den Resultaten der RDS-Strategie in Abb. 2.22, so sieht man, dass die zweite Strategie eine breitere Verteilung der Ausgabe bewirkt.

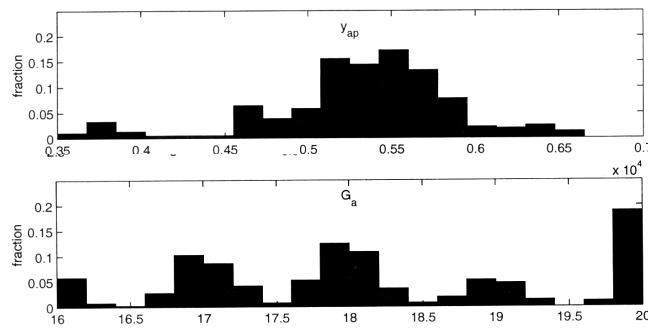


Abb. 2.22 RDS Strategie: Ausgabe (oben) und Eingabe (unten)

Das Training des RBF-Netzes geschieht separat für beide Schichten. Die Gewichte der zweiten Schicht werden jeweils so ermittelt, dass sie den kleinsten quadratischen Fehler erzeugen. Bei der ersten Schicht wurden zwei verschiedene Strategien ausprobiert:

1. Die Zentren der ersten Schicht werden mittels *k-mean-Clusterung* ermittelt.
2. Die Zentren der ersten Schicht werden durch *sequentielle Regression* ermittelt.

Ergebnisse

Es wurden drei verschiedene Mengen erzeugt: Eine Trainingsmenge nach der RUS-Strategie, eine Trainingsmenge nach der RDS-Strategie und eine Testmenge. Die Ergebnisse zeigen, dass breitere RBF-Funktionen (mit größerer Streuung β) auch bessere Ausgaben erzeugen.

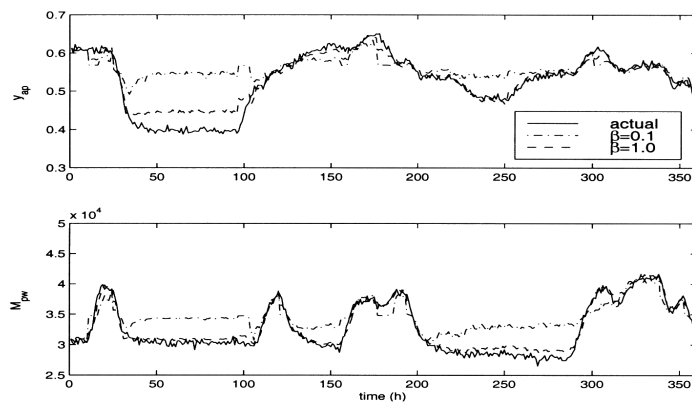


Abb. 2.23 Ergebnisse der RBF-Approximation mit unterschiedlich breiten Zentren

Man sieht, dass die gestrichelten Ausgaben von schmalere RBF deutliche Abweichungen von der vorgegebenen, durchgezogen gezeichneten Ausgaben haben. Bei einem analogen Training mit der zweiten Datenmenge ergeben sich bessere Ergebnisse.

Die Erklärung dafür liegt in den Unterschieden zwischen den Verteilungen der Ein- bzw. Ausgabevariable: Im ersten Fall gibt es bei der Ausgabe y nur sehr geringe Eingabewerte unterhalb des Wertes 0,45. Dies führt dazu, dass durch fehlende RBF-Zentren in den „leeren“ Bereichen die Interpolation bei den Testdaten nur sehr schlecht möglich ist.

Ein andere Frage ist die, welche der beiden Initialisierungsmethoden für die erste Schicht bessere Ergebnisse bei diesem Anwendungsbeispiel zeigt. In Abb. 2.24 sind die Ausgabeabweichungen für beide Methoden gezeigt.

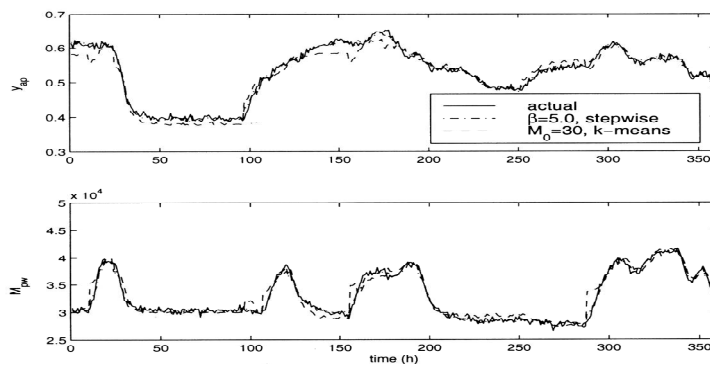


Abb. 2.24 Ergebnisse der RBF-Approximation mit k-mean-Clusterung vs. sequentielle Regression.

Es stellt sich heraus, dass die sequentielle Methode der Regression etwas besser abschneidet als die k-mean-Clusterung, da sie stärker ergebnisorientiert auf den mittleren Fehler abzielt und sich nicht allein auf die Eingabecluster orientiert.

2.11 Approximationsleistung neuronaler Netze

Was können wir nun grundsätzlich von Neuronalen Netzen erwarten? Können wir sie immer dazu verwenden, unbekannte, nichtlineare Funktionen zu approximieren, oder haben sie Grenzen? Wie viele Schichten benötigen wir ?

Betrachten wir eine Klasse von Funktionen, genannt *Sigma-Funktionen* $\Sigma^n(S)$

$$\Sigma^n(S) := \{ \hat{f} \mid \hat{f} : \mathfrak{R}^n \rightarrow \mathfrak{R} \text{ mit } \hat{f}(\mathbf{x}) = \sum_{j=1}^m w_j^{(2)} S(z_j(\mathbf{x})) \} \quad (2.88)$$

wobei $w_j^{(2)}$ aus \mathfrak{R} , \mathbf{x} aus \mathfrak{R}^n

Die Sigma-Funktionen sind genau die Menge von Funktionen, die mit einem zweischichtigem, neuronalen Netz wie in Abb. 2.8 erzeugt werden. Betrachten wir nun als Aktivitätsfunktion z_i des i -ten Neurons der ersten Schicht eine *affine Funktion*, d.h. sie kann Größenänderungen, Drehungen und Verschiebungen eines Vektors \mathbf{x} bewirken

$$z_j \in z^n := \{ z \mid z(\mathbf{x}) = \mathbf{w}^{(1)T} \mathbf{x} + b \} \quad \text{affine Funktionen } \mathfrak{R}^n \rightarrow \mathfrak{R} \quad (2.89)$$

wobei $b \in \mathfrak{R}$ die negative Schwelle und $\mathbf{w}^{(1)}$ der Gewichtsvektor eines Neurons der ersten Schicht ist. Die Ausgabefunktion $S(\mathbf{x})$ ("Quetschfunktion") kann dabei beliebig sein, vorausgesetzt, sie erfüllt die Bedingungen

$$\lim_{x \rightarrow \infty} S(x) = 1, \quad \lim_{x \rightarrow -\infty} S(x) = 0, \quad (2.90)$$

Die zweite Schicht hat wieder eine lineare Aktivitätsfunktion mit den Gewichten $\mathbf{w}^{(2)}$; die Ausgabefunktion ist die Identität $\hat{f}(\mathbf{x}) = S(z^{(2)}) = z^{(2)}$.

Als Maß für die Abweichung zwischen der approximierenden Netzausgabe $F(\mathbf{x})$ und der zu lernenden Funktion $f(\mathbf{x})$ nehmen wir die $L_s(p)$ -Norm

$$\|f(\mathbf{x}) - \hat{f}(\mathbf{x})\| = \left(\int |f(\mathbf{x}) - \hat{f}(\mathbf{x})|^s dP(\mathbf{x}) \right)^{1/s} \quad (2.91)$$

was z.B. für zufällige \mathbf{x} mit der Wahrscheinlichkeitsverteilung $P(\mathbf{x})$ der erwarteten Abweichung entlang des ganzen Kurvenzuges von $f(\mathbf{x})$ entspricht; im Fall $s = 2$ ist dies die Wurzel aus unserem erwarteten quadratischen Fehler. Im Gegensatz dazu entspricht der *uniforme* Abstand

$$\|f(\mathbf{x}) - \hat{f}(\mathbf{x})\| = \sup_x |f(\mathbf{x}) - \hat{f}(\mathbf{x})| \quad (2.92)$$

dem maximalen Fehler, der bei der Approximation überhaupt auftreten kann, unabhängig von seiner Häufigkeit $P(\mathbf{x})$.

Dann gilt:

- 1) Für die Funktionswerte *jeder beliebigen Funktion* $f(\mathbf{x}) : \mathfrak{R}^n \rightarrow \mathfrak{R}$ von N Mustern $\mathbf{x}^1 \dots \mathbf{x}^N$ gibt es eine Sigma-Funktion \hat{f} derart, dass für alle Muster \mathbf{x}^i mit $i = 1..N$ gilt

$$\hat{f}(\mathbf{x}^i) = f(\mathbf{x}^i) \quad (2.93)$$

Natürlich gilt dies auch für eine ganze Schicht, d.h. für $\hat{\mathbf{f}} = (\hat{f}_1, \dots, \hat{f}_m)$, den Ausgabevektor aus Sigma-Funktionen.

- 2) Jede beliebige, stetige Funktion $f(\mathbf{x})$ in einem kompakten Intervall ("kompakte Teilmenge des \mathfrak{R}^n ") kann beliebig dicht (*uniform dicht* im Sinne der L_∞ -Norm in der Menge C^n aller stetigen Funktionen und ρ_p -*dicht* in der Menge der Borel meßbaren Funktionen) durch eine Sigma-Funktion $\hat{f}(\mathbf{x})$ approximiert werden.

Die obigen Aussagen gelten übrigens auch für den allgemeineren Fall, wenn die zweite Schicht aus *Sigma-Pi-Units* besteht. Für die Quetschfunktion reicht sogar die Eigenschaft "stetig, begrenzt und nicht-konstant" aus. Damit ist auch die Klasse der *radialen Basisfunktionen* wie die Gaußsche Glockenkurve in den möglichen Ausgabefunktionen enthalten.

3 Wissensbasierte Modellierung

Eine Modellierung, die nur nach dem Verhalten eines Systems geht, ermöglicht eine Vorhersage nur in den Grenzen des Kontextes, in dem vorher Daten gemessen wurden. Wir können nun den Vorhersagebereich stark ausdehnen, wenn wir vorhandenes Wissen nutzen und das Verhalten des Systems mithilfe von internen Mechanismen erklären können. In diesem Fall benötigen wir auch weniger Messungen, um das Modell an die Realität, die gemessenen Daten, anzupassen.

Es gibt nun verschiedene Modellierungsmöglichkeiten für zeitkontinuierliche Systeme. Eine der bekanntesten ist die Modellierung durch dynamische Systeme.

3.1 Einführung in dynamische Systeme

Der Ansatz der dynamischen Systeme hat eine lange Tradition. Populär wurde der Ansatz, als der „Club of Rome“, eine 1968 gegründete Vereinigung von Naturwissenschaftlern, Ökonomen, Planungs- und Bildungsexperten, ihren Eindruck der bedrohlichen Entwicklung der Welt mit einer Studie genauer prüfen lassen wollte. Auf der Suche nach einer geeigneten quantitativen Methodologie stießen sie auf die von Jay. W. Forrester am MIT entwickelte Methode der „Systemdynamik“ und initiierten ein Modellierungs- und Simulationsprojekt. Das Projekt „Weltsimulation“ wurde von seinem Schüler Dennis Meadows und Mitarbeitern am MIT durchgeführt und die Ergebnisse im März 1972 unter dem Titel „The Limits to Growth“ („Die Grenzen des Wachstums“) veröffentlicht. Der 600 Seiten starke technische Bericht dazu wurde in 35 Sprachen übersetzt und 9 Mill. Mal weltweit verkauft. Ein Nachfolgebericht von Meadows erschien 1992.

Die Ergebnisse der Studie besagten, dass wir dringend unsere Technik und Wirtschaftsweisen entwickeln müssen, um einen Kollaps der Menschheit in wenigen Jahrzehnten zu verhindern. Dies wurde natürlich sehr kontrovers diskutiert und führten zu weiteren Entwicklungen, bei denen das globale Modell durch eine Anzahl miteinander gekoppelter lokaler Modelle ersetzt wurde.

Bevor wir uns grundsätzlich mit dem Für und Wider dieser Methode befassen, möchte ich Ihnen gern die Modellierungsmethode selbst vorstellen und dazu einige Beispiele einführen nach dem Vorbild von Bossel (1994).

3.1.1 Modellierung mit Regeln

Beginnen wir mit einem einfachen System, etwa dem Wirkungszusammenhang zwischen Eisflächenbildung und Sonneneinstrahlung in der Antarktis. Dazu können wir uns am Anfang einige Zusammenhänge lose notieren:

- (1) Die Sonneneinstrahlung unterteilt sich in einen Teil, der absorbiert wird, und einen Teil, der zurückgestrahlt wird. Wenn sich die Eisfläche vergrößert, erhöht sich die *Albedo*, das Rückstrahlvermögen (Verhältnis Rückstrahlungsenergie zu Einstrahlungsenergie); wenn sie kleiner wird, verringert sich die *Albedo*.
- (2) Ein Anstieg der Temperatur bewirkt eine Verkleinerung der Eisfläche, ein Absinken eine Vergrößerung.
- (3) Erhöht sich die *Albedo*, so wird weniger Einstrahlungsenergie absorbiert. Umkehrt wird mehr Energie absorbiert, wenn sich die *Albedo* verkleinert.
- (4) Wird weniger Energie absorbiert, so sinkt die Temperatur; bei erhöhter Absorption steigt die Temperatur.

Aus diesen umgangssprachlichen Feststellungen kann man in einem ersten Schritt alle wichtigen Zustandsgrößen herausziehen und eine Liste bilden („Glossar“, ähnlich wie in der Softwareentwicklung zum Bilden einer Spezifikation). In unserem Fall sind dies die Größen „Eisflächengröße“, „Temperatur“, „*Albedo*“ und „Absorption“. Die vier Feststellungen kann man nun in Form von formalen Regeln formulieren, deren Argumente von der Zeit t abhängen:

- WENN wächst_Eisfläche(t) DANN wächst_*Albedo*(t)
- WENN sinkt_Eisfläche(t) DANN sinkt_*Albedo*(t)
- WENN wächst_Temperatur(t) DANN sinkt_Eisfläche(t)
- WENN sinkt_Temperatur(t) DANN wächst_Eisfläche(t)
- WENN wächst_*Albedo*(t) DANN sinkt_Absorption(t)
- WENN sinkt_*Albedo*(t) DANN wächst_Absorption(t)
- WENN sinkt_Absorption(t) DANN sinkt_Temperatur(t)
- WENN wächst_Absorption(t) DANN wächst_Temperatur(t)

Hätten wir noch zusätzliche Einflüsse, so würde die eine Bedingung nach WENN durch weitere Bedingungen mit UND verbunden.

Dabei beobachten wir, dass bei den Regeln immer Paare existieren, die eine positive bzw. negative Wirkung charakterisieren. Wir können diese Paare zusammenfassen, indem wir einen *Wirkungsgraphen* konstruieren, in dem wir alle Grundentitäten als Knoten und die Einflüsse als gerichtete Kanten (Pfeile) darstellen, siehe Abb. 3.1. Hierbei sind gegenläufige Einflüsse mit dem Minuszeichen versehen, gleichlaufende mit dem Pluszeichen.

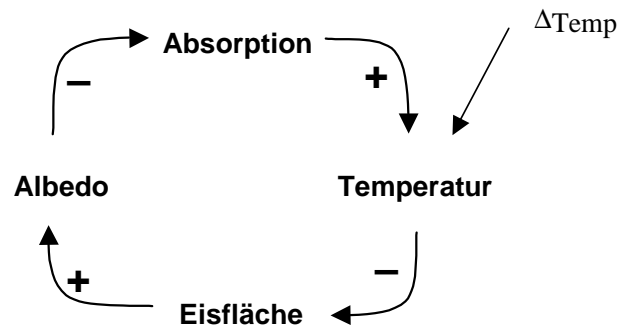


Abb. 3.1 Der Wirkungsgraph des Antarktis-Energiemodells

Prüfen wir die logische Wirkungskette sowohl bei den Regeln als auch im Graphen, dann bemerken wir, dass durch zweimalige Negation im Kreislauf der gesamte Kreislauf positiv ist: Eine Erhöhung der Temperatur (angedeutet durch den Einfluss $\Delta Temp$ in der Abbildung) führt zu einer Erhöhung der Temperatur, eine Erniedrigung zu einer Erniedrigung. Tendenziell ist dies richtig: eine gerade Zahl von Minuszeichen auf einem Zyklus (geschlossenen Weg) bedeutet eine Instabilität; eine ungerade Zahl eine negative Rückkopplung und damit eine Stabilität. Würden wir die obigen logischen Regeln hintereinander abarbeiten, so stellt sich sofort ein System mit maximaler bzw. minimaler Temperatur ein. Dies ist aber in der Realität nicht so. Warum?

Der Grund liegt darin, dass die Einflüsse mit Gewichtungsfaktoren versehen sind. Sehen wir beispielsweise einen Einflussfaktor von jeweils nur 0,1 bei jedem Begriff vor, so erniedrigt sich die Änderung einmal im Kreis um $0,1 \cdot 0,1 \cdot 0,1 \cdot 0,1$; sie klingt schnell ab. Um dies bei der Modellierung mit Regeln zu berücksichtigen, müsste sie derart verändert werden, dass die Regeln nur mit einem gewissen Faktor wirken. Dies erreichen wir mit einem quantifizierten Wirkungsgraphen, in dem die Einflüsse bzw. Regeln jeweils mit einem Einflussfaktor versehen sind.

Für die Erläuterung, wie wir einen solchen quantifizierten Wirkungsgraphen aufbauen, wählen wir uns als größeres Beispiel ein sehr ehrgeiziges Projekt: Wir wollen die Funktion der ganzen Welt verstehen. Dazu beginnen wir wieder mit den Informationen, die wir darüber haben, den umgangssprachlichen Texten, und ziehen das Wesentliche heraus: Wir erstellen zunächst ein Textmodell. Unser Ziel dabei soll sein, mit einer möglichst geringen Zahl von Variablen qualitative richtige Aussagen über Entwicklung und Dynamik der Bevölkerung und Wirtschaft zu machen. Falls sich daraus instabiles Verhalten andeutet, so sollen auch grundsätzliche Möglichkeiten der Stabilisierung erwogen werden.

3.1.2 Modellierung des Textes

Angenommen, wir haben folgenden ökologischen Text (s. Bossel 1994), den wir zur Übersicht mit Bezeichnern (a), (b), ... annotiert haben:

- (a) „Wir beobachten heute eine zunehmende *Belastung der natürlichen Ressourcen und der natürlichen Umwelt*. Die Gründe hierfür sind einerseits eine ständige Zunahme der *Bevölkerung* und damit auch des *Verbrauchs der Rohstoffe* und der damit verbundenen *Abgabe von Abfallstoffen* aller Art am Ende ihrer Nutzung *an die Umwelt*.
- (b) Eine wichtige Bestimmungsgröße dieser *Ressourcen- und Umweltbelastung* ist der *spezifische Verbrauch an Rohstoffen und Energie pro Kopf*. Dieser spezifische Verbrauch steigt noch tendenziell mit der wachsenden *Umweltbelastung* durch den zusätzlichen Aufwand für Umweltschutz und die schwierigeren Abbaubedingungen der Rohstoffe.
- (c) Mit wachsendem spezifischem *Konsum* durch materiellen Wohlstand verbessern sich aber auch die *Versorgungsmöglichkeiten*, was einen entsprechenden Einfluss auf die *Bevölkerungsentwicklung* hat.
- (d) Aufgrund der wachsenden *Umweltbelastungen* mit Schadstoffen wie auch der schwindenden *natürlichen Ressourcenbasis* ergeben sich aber auch Rückwirkungen auf die *Gesundheit und Lebenserwartung der Bevölkerung*.
- (e) Die *Umweltbelastungen* und die Eingriffe in die *natürliche Ressourcenbasis* führen zu wachsenden *gesellschaftlichen Kosten*, die wiederum ein zunehmendes *gesellschaftliches Handeln* erwarten lassen, um schädlichen Entwicklungen zu begegnen.“

Wir wollen nun die wichtigsten Größen aus dem Text extrahieren. Dazu stellen wir eine Liste von Worten (*Glossar*) auf, die jeweils einen Zustand darstellen. Die Semantik soll die Wechselwirkungen zwischen diesen Zuständen ausdrücken. Problematisch dabei ist, dass Größen existieren können, die nicht explizit erwähnt werden, sondern nur als Kontext allen bewusst sind. Diesen Fall wollen wir hier nicht weiter betrachten und isolieren stattdessen folgende fünf Einflüsse bzw. Begriffe:

- (1) Umwelt- und Ressourcenbelastung
- (2) Bevölkerung
- (3) Spezifischer materieller Verbrauch pro Kopf
- (4) Konsum
- (5) Versorgungsmöglichkeiten
- (6) Gesellschaftliche Kosten
- (7) Gesellschaftliches Handeln

Dabei haben wir Begriffe wie „*Belastung der natürlichen Ressourcen und natürlichen Umwelt*“ sowie „*Verbrauchs der Rohstoffe und ... Abgabe von Abfallstoffen an die Umwelt*“ zu dem Begriff „Umwelt- und Ressourcenbelastung“ zusammengefasst. Auch die Bezeichnung „*spezifische Verbrauch an Rohstoffen und Energie*

pro Kopf“ kann mit „Spezifischer materieller Verbrauch pro Kopf“ abgekürzt und verallgemeinert werden.

Uns ist klar, dass wir mit diesen wenigen Begriffen die Welt nicht erschöpfend erklären können, sondern nur wenige, wichtige Aspekte herausnehmen, um unseren Zweck der Modellierung zu erfüllen.

Eine geeignete Modellierung entsteht, wenn wir mit den Grundbegriffen die Wirkungsbeziehungen aufstellen. Dazu betrachten wir zuerst Abschnitt (a):

- (1) WENN die Bevölkerungszahl wächst, DANN wächst auch die Umwelt- und Ressourcenbelastung.

Aus dem zweiten Satz in Abschnitt (b) folgt

- (2) WENN die Umwelt- und Ressourcenbelastung wächst, DANN wächst auch der spezifische materielle Verbrauch pro Kopf.

Weiterhin steht darin, dass der spezifische Verbrauch aber auch eine „Bestimmungsgröße“ ist:

- (3) WENN der spezifische materielle Verbrauch pro Kopf wächst, DANN wächst auch die Umwelt- und Ressourcenbelastung.

Mit Abschnitt (c) folgt

- (4) WENN der Konsum wächst, DANN wächst die Versorgung.
- (5) WENN die Versorgung wächst, DANN wächst auch die Bevölkerung.

und aus Abschnitt (d) folgt

- (6) WENN die Umwelt- und Ressourcenbelastung wächst, DANN vermindert sich die Bevölkerungszahl.

sowie aus Abschnitt (e) die Aussagen

- (7) WENN die Umwelt- und Ressourcenbelastung wächst, DANN wachsen auch die gesellschaftlichen Kosten.
- (8) WENN die gesellschaftlichen Kosten wachsen, DANN wird auch das gesellschaftliche Handeln zunehmen.

Die „schädlichen Entwicklungen“ interpretieren wir als Umschreibung von wachsender Bevölkerung und wachsendem spezifischem materiellem Verbrauch pro Kopf:

- (9) WENN gesellschaftliches Handeln erfolgt, DANN wird es bei zu starkem Bevölkerungswachstum durch geeignete Maßnahmen dieses reduzieren.
- (10) WENN gesellschaftliches Handeln erfolgt, DANN wird es bei zu hohem spezifischem materiellem Verbrauch pro Kopf diesen reduzieren.

Für ein einfaches Modell müssen wir als erstes semantisch ähnliche Begriffe zusammenfassen. In diesem Fall sind die Begriffe „Konsum“ und „materieller Verbrauch pro Kopf“ sehr ähnlich, so dass wir sie zu „Konsum pro Kopf“ oder kurz „Konsum“ zusammenfassen.

Natürlich können wir diese Regeln wieder formal hinschreiben, analog zu den Regeln der Eisbildung, aber stattdessen wollen wir sofort den Wirkungsgraphen dazu aufstellen. Dazu stellen wir wieder die Grundbegriffe als Knoten und die Wirkungen zwischen den Grundelementen als gerichtete Kanten dar, die zusätzlich noch eine Gewichtung in Form eines Vorzeichens erhalten. In Abb. 3.2 ist dies gezeigt.

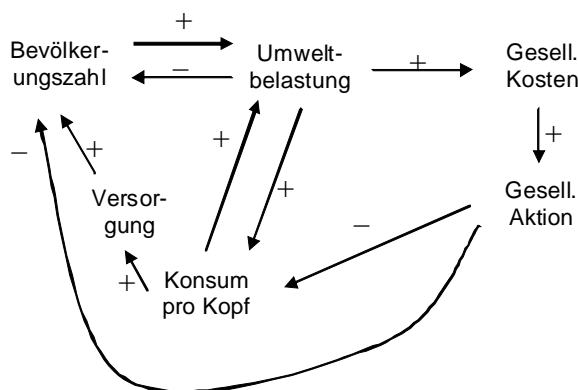


Abb. 3.2 Wirkungsgraph des Weltmodells

Bei der Aufstellung der Beziehungen beachten wir drei wichtige Randbedingungen, um Fehler zu vermeiden:

1. Wir modellieren **nur direkte Einflüsse**, also den Einfluss von A auf B und von B auf C, die durch Pfeile (gerichtete Kanten) im Wirkungsgraphen visualisiert werden. Der indirekte Einfluss, der sich dadurch von A auf C ergibt, darf nicht durch einen Pfeil von A nach C visualisiert werden, da er nicht direkt beobachtet wird.
2. Dabei wird jede Wirkung so betrachtet, als ob das restliche System sich nicht verändern würde, also **alle anderen Einflüsse „eingefroren“** seien. Dies entspricht einer Art Sensitivitätsanalyse, bei der eine kleine Störung an einer Stelle existiert und im (quasi stabilen) System direkte Änderungen hervorruft.
3. Der entstehende Wirkungsgraph gilt **nur für einen bestimmten Kontext**, etwa die Ausgangssituation. Ändert sich der Kontext, können neue Wirkun-

gen entstehen und alte wegfallen: Der Wirkungsgraph kann sich stark ändern.

Für die Analyse des so entstandenen Wirkungsgraphen können wir uns einige Begriffe definieren. Laufen wir auf von einem Knoten entlang der Pfeile durch den Graphen, so definieren wir uns damit eine **Folge** von Knoten. Beachten wir, dass wir dabei jeden Knoten nur einmal durchlaufen, so erhalten wir einen **Pfad**. Endet der Pfad beim Startknoten, so ist dies ein **Kreis** oder **Zyklus**. Passieren wir mehrmals denselben Knoten, so ist dies kein Kreis mehr, sondern nur noch eine einfache **Schleife**.

Bevor wir das Wirkungsmodell näher analysieren, ist es sinnvoll, möglichst alle Knoten zu entfernen, die nichts zum Systemverhalten beitragen. Beispielsweise könnten wir noch einen Knoten „Resignation“ einzeichnen, der bei hohen gesellschaftlichen Kosten größer wird. Da er aber keine Wirkung auf andere Knoten in diesem Modell hat (**Senke**), können wir ihn auch einfach weglassen. Auch andere Kontextvariable, die nicht beeinflusst werden (**Quellen**), können wir weglassen, etwa einen nicht erreichbaren Knoten „Katastrophe“, mit der die Bevölkerungszahl in einem Schlag heruntergesetzt werden kann. In Abb. 3.2 sind außerdem noch zwei Knoten „Versorgung“ und „Gesell. Handeln“ vorhanden, die als Übergangsknoten entbehrlich sind (**unkritische Elemente**), da deren Auswirkungen zum direkten Nachbarknoten mit dem Vorgänger- oder Nachfolgeknoten zu einem einzigen Knoten zusammengefasst werden können. Wir fassen deshalb „Versorgung“ mit „Konsum“ sowie „Gesell. Handeln“ mit „Gesell. Aktion“ zusammen. Sinnvollerweise quantifizieren wir außerdem die Wirkungen durch Gewichte w_i und erhalten so einen gewichteten Wirkungsgraphen, siehe Abb. 3.3.

Tragen wir die einzelnen Zustände, die Knoten, sowie die Gewichte der Wechselwirkungen zwischen den Knoten in eine 4×4 Matrix **A** ein, so erhalten wir für unser Beispiel die Wirkungsmatrix in Abb. 3.3. Hier sind in jeder Spalte die von einem Knoten ausgehenden Wirkungen bzw. ihre Gewichte eingetragen; in den Zeilen stehen die auf einen Knoten wirkenden Einflüsse mit ihren Gewichten. Die Wirkungsmatrix wird auch als **Systemmatrix** bezeichnet.

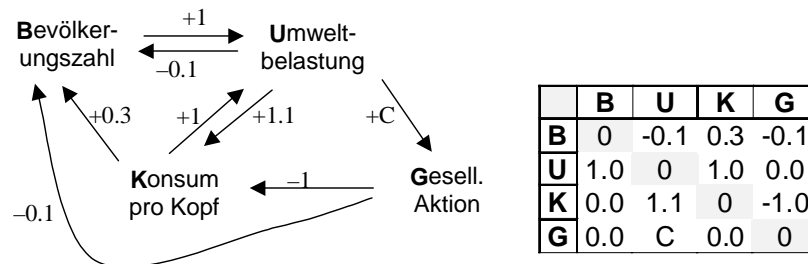


Abb. 3.3 Gewichteter Wirkungsgraph des Weltmodells und seine Wirkungsmatrix

In der Matrix ist die Auswirkung der Umweltbelastung auf das gesellschaftliche Handeln mit einem Parameter C gekennzeichnet; er ermöglicht eine Kontrolle des Systems.

Welche Systemgrößen sind wichtig, welche nicht? Die *aktiven* Knoten x_i im Graph sind diejenigen, die alle anderen stark beeinflussen, selbst aber wenig beeinflusst werden. Den Einfluss auf andere Knoten von x_i können wir durch die Spaltensumme S_i der Absolutwerte der Systemmatrix A , die Beeinflussung von x_i durch die Zeilensumme Z_i der Absolutwerte messen.

$$S_i = \sum_j |A_{ji}|, \quad Z_i = \sum_j |A_{ij}| \quad (3.1)$$

Der Quotient S_i / Z_i gibt dann ein Maß für die Aktivität des Knotens. Entsprechend können wir auch die *kritischen* Knoten im Graphen bestimmen: Bei ihnen sind sowohl der Einfluss als auch die Beeinflussung, also das Produkt $S_i \cdot Z_i$ besonders hoch.

Beispiel Im Wirkgraphen in Abb. 3.3 ist mit $C = 0,3$ der aktivste Knoten G mit dem Quotienten 3,67, der passivste der Knoten K mit 0,62. Der kritischste Knoten ist U mit dem Produkt 3,00 und der unkritischste der Knoten G mit 0,33.

Weiterhin müssen wir uns über den Sinn des Wirkungsgraphen klar werden. Was soll damit ausgesagt werden? Soll die *absolute* Größe x_i eines Grundbegriffs (etwa die absolute Bevölkerungszahl) oder die *relative* Veränderung des Grundgröße (etwa eine Zunahme um 10%) bei einer Veränderung der Eingangsgrößen als Wirkung auftreten?

Ist die Gewichtung für die absoluten Größen (Zustände) gemeint, so beschreiben wir damit eine Zustandsentwicklung

$$\text{Zustandsdynamik} : z_i(t+1) = z_i(t) + \sum_{\text{alle Einflüsse}} w_j z_j \quad (3.2)$$

Beispiel Bei einer Bevölkerungszahl B und einer Umweltbelastung von U erhöht sich die Umweltbelastung U durch Rohstoffabbau nach einem Zeitschritt Δt um B auf $U(t+\Delta t) = U(t) + B$.

Typische Systeme für eine Zustandsdynamik sind alle Regelungssysteme, etwa von Staubecken oder Industrieprozessen, siehe Kapitel 2.10.

Für eine Störungsanalyse nehmen wir dagegen an, dass das gesamte System stabil ist und wir nur kleine Störimpulse darauf geben, um zu sehen, wie das System reagiert. Typische Systeme dafür sind Verstärker, etwa für Musik, die in einem festen Gleichgewicht sind, durch die Eingangssignale aus diesem Gleichgewicht nur ausgelenkt werden und danach wieder dahin zurückkehren (periodische Signale als Eingabe). Bezieht sich also die Wirkung auf kleine Änderungen (Pulse) und ihre Auswirkungen, so beschreiben wir die Pulsfortpflanzung als eine Störung

eines augenblicklichen Zustands; der neue Zustand ergibt sich dann aus den alten Zuständen z_i zuzüglich der Änderungen Δz_j der Knoten j , die Auswirkungen haben auf Knoten i

$$\text{Pulsfortpflanzung: } z_i(t+1) = z_i(t) + \sum_{\text{alle Einflüsse}} w_j \Delta z_j \quad (3.3)$$

Beispiel Erhöht sich die Bevölkerungszahl um 10%, so erhöht sich auch die Umweltbelastung durch Rohstoffabbau um 10%.

An den beiden Beispielen sehen wir, dass die Auswirkungen der Koeffizienten der Systemmatrix sehr davon abhängen, in welchem Modell sie interpretiert werden. Die diskreten Zahlenwerte der Systemmatrix hängen also von der Interpretation der Wechselwirkungen ab; unterschiedliche Interpretationen haben unterschiedliche Systemmatrizen zur Folge.

Die obigen Formulierungen (3.2) und (3.3) können wir verallgemeinern. Angenommen, wir fassen alle Systemvariablen z_i zu einem Vektor $\mathbf{z} = (z_1, z_2, \dots, z_n)^T$ zusammen, so können wir mit Hilfe der Systemmatrix \mathbf{A} alle Auswirkungen auf eine Variable z_i als Skalarprodukt von \mathbf{z} mit der i -ten Zeile der Matrix \mathbf{A} beschreiben. Dies gilt für alle Variablen, so dass wir pro Zeitschritt folgende Formulierung für das Zustandssystem (3.2) erhalten:

$$\mathbf{z}(t+1) - \mathbf{z}(t) = \Delta \mathbf{z}(t) / \Delta t = \mathbf{A} \mathbf{z}(t) \quad \text{mit } \Delta t = 1 \quad (3.4)$$

Dies gilt für einen Zeitschritt $\Delta t = 1$. Machen wir die Zeitschritte sehr klein, so können wir für den Grenzübergang von Gl.(3.4) schreiben

$$\lim_{\Delta t \rightarrow 0} \frac{\Delta \mathbf{z}}{\Delta t} = \frac{d\mathbf{z}}{dt} = \mathbf{A} \mathbf{z}(t) \quad \text{lineares System} \quad (3.5)$$

Für die Pulsfortpflanzung von (3.3) erhalten wir

$$\Delta \mathbf{z}(t+1) = \mathbf{A} \Delta \mathbf{z}(t) \quad \text{lineares System} \quad (3.6)$$

3.1.3 Stabilität

Im Folgenden wollen wir nun die Stabilität des Gesamtsystems damit untersuchen. Wann ist ein solches System stabil? Betrachten wir dazu die Wirkungspfade in Abb. 3.3, speziell die Zyklen. Beispielsweise erkennen wir, dass es insgesamt 6 Zyklen gibt. Es sind die Folgen der Grundgrößen, abgekürzt mit ihrem Anfangsbuchstaben

1. B,U,B
2. B,U,G,B
3. B,U,K,B
4. B,U,G,K,B

5. U,K,U
6. U,G,K,U

Laufen wir über einen solchen Pfad, so können wir die Auswirkungen einer Eingabe oder Störung über mehrere Knoten hinweg akkumulieren. So ist für den Weg 2 für die Interpretation der Systemmatrix als Störungspropagierung

$$\Delta U(t) = 1.0 \Delta B(t), \quad \Delta G(t) = C \Delta U(t), \quad \Delta B(t+1) = -0.1 \Delta G(t)$$

Also erhalten wir

$$\Delta B(t+1) = -0.1 \cdot C \cdot 1.0 \cdot \Delta B(t) = w_g \cdot \Delta B(t) \quad (3.7)$$

Halten wir alles andere konstant, so ist durch diese Rückkopplung das System nur dann stabil, wenn die Störungen nicht größer werden, also

$$|w_g| \leq 1 \quad (3.8)$$

gilt. Dabei reicht es nicht aus, wenn nur $w_g < 1$ gelten würde, da auch bei einem starken negativen Wert von w_g die Störung verstärkt werden würde, wenn auch mit wechselndem Vorzeichen nach jedem Zyklus.

Ein weiteres Beispiel für ein solches System aus der täglichen Erfahrung ist ein akustisches Verstärkersystem, bei dem das Mikrofon neben dem Lautsprecher aufgebaut ist. Haben wir die Verstärkung zu weit „aufgedreht“, so dass eine kleine Schwankung am Mikrofon eine stärkere Luftdruckschwankung durch den Lautsprecher bewirkt, so gerät das System in Rückkopplung und fängt (mit unangenehm hohem Ton) zu schwingen an. Erst eine Verstärkungsreduzierung auf <1 am Mikrofon unterbricht die Voraussetzung für die Schwingung und damit den Ton.

Nun existiert ja nicht nur eine Rückkopplung, sondern mehrere, so dass wir im Endeffekt eine Überlagerung aller Rückkopplungen haben, die je nach Größe der Parameter völlig andere Resultate hervorbringen können. Zur Illustration sind für unser Beispiel des Weltmodells verschiedene Systementwicklungen mit verschiedenen Parameterwerten C in Abb. 3.4 gezeigt. Da bei der Pulspropagierung nur die Zustandsänderungen weitergegeben werden, ist das qualitative Systemverhalten unabhängig von den Anfangszuständen.

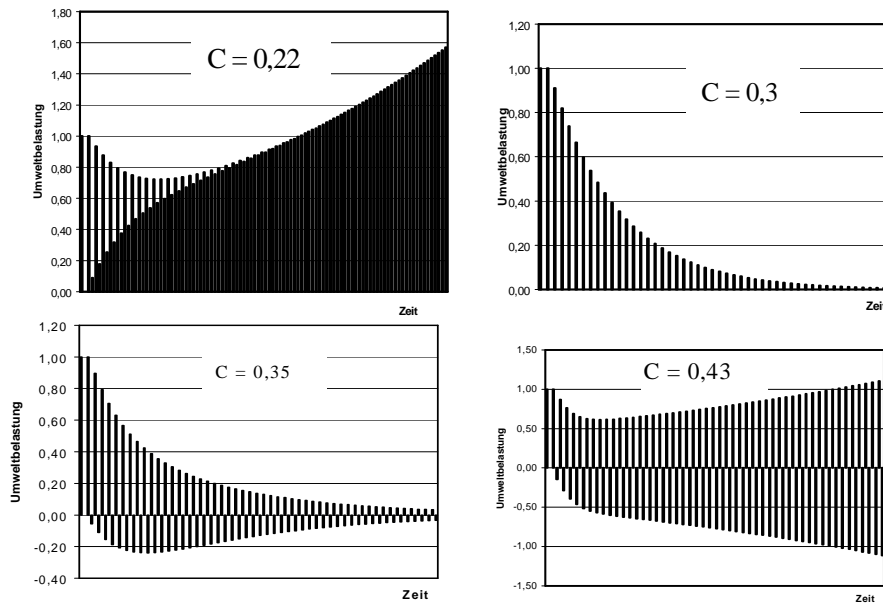


Abb. 3.4 Entwicklung der Umweltbelastung U mit der Zeit für verschiedene Werte von C

Es zeigt sich, dass das System für verschieden Werte von C verschieden reagiert. Bei kleinen Werten ist es instabil und pendelt zwischen großen und kleinen Werten, die beide mit der Zeit anwachsen. In dem Intervall $0,23 < C < 0,41$ wechselt zwar das Systemverhalten zu positiv-negativ alternierenden Größen, aber die Störungen klingen mit der Zeit ab. Wird C zu groß, so wachsen die Störungen wieder an und das System gerät außer Kontrolle. Wir sehen, zu wenig oder zu viel Kontrolle ist fatal in diesem Weltmodell.

Allgemeinerweise wissen wir aus der Systemtheorie, dass das Pulssystem stabil ist, wenn das System nach Gl. (3.6) stabil ist. Es ist klar, dass nach der linearen Transformation der Zustand stabil ist, also derselbe Zustand resultiert, wenn die Gleichung

$$\Delta \mathbf{z}(t+1) = \mathbf{A} \Delta \mathbf{z}(t) = \lambda_i \Delta \mathbf{z}(t) \quad (3.9)$$

erfüllt ist und dabei

$$|\lambda_i| < 1 \quad (3.10)$$

gilt. Dies bedeutet, dass alle Eigenwerte λ_i der Systemmatrix vom Betrag her kleiner als 1 sein müssen.

Wie ist nun die Situation bei dem Zustandsmodell? Im skalaren Fall von Gl.(3.5) haben wir eine direkte Rückkopplung, s. Abb. 3.5a).

$$\frac{dz}{dt} \equiv z' = wz \Rightarrow \frac{z'}{z} = w \quad (3.11)$$

Eine Integration beider Seiten ergibt

$$\ln(z) + a_1 = \int \frac{z'}{z} dt' = \int w dt' = wt + a_2 \quad a_0 = a_2 - a_1 \quad (3.12)$$

und somit

$$\ln(z) = wt + a_0 \Rightarrow z(t) = e^{wt+a_0} = z(0) e^{wt} \quad \text{mit } z(0) = e^{a_0} \quad (3.13)$$

Wann ist dies stabil? Die Exponentialfunktion bleibt nur gleich oder nimmt ab, wenn

$$w \leq 0 \quad (3.14)$$

gegeben ist, siehe Abb. 3.5 b).

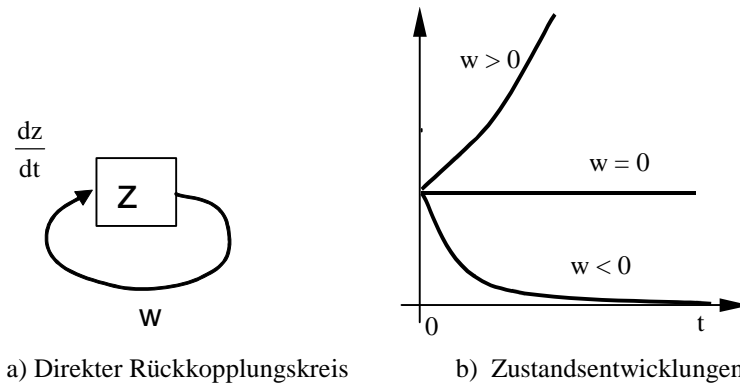


Abb. 3.5 Systemzustandsänderung durch direkte Rückkopplung

Wie ist nun die Situation in unserem komplexeren Weltmodell bei mehreren Zustandsvariablen? Dazu reduzieren wir das System um einen weiteren Zustand. Betrachten wir dazu den Wirkungsgraphen in Abb. 3.3, so bemerken wir, dass von der Größe "Umweltbelastung" ein direkter Einfluss der Größe C auf "gesell. Aktion" und dann mit einem festen Faktor ein Einfluss auf Konsum und Bevölkerungszahl besteht. Aus der Zustandsdynamik Gl.(3.2) folgt

$$\begin{aligned} B_{t+1} &= B_t - 0.1U_t + 0.3K_t - 0.1G_t, & U_{t+1} &= U_t + 1.0B_t + 1.0K_t \\ K_{t+1} &= K_t + 1.1U_t - 1.0G_t, & G_{t+1} &= G_t + C \cdot U_t \end{aligned}$$

Ist die gesellschaftliche Aktion nur von der augenblicklichen Umweltbelastung abhängig und hat kein Gedächtnis, also $G_t = 0$, so kann sie als abgeleitete Hilfsgröße eliminiert werden und wir erhalten durch Einsetzen

$$B_{t+1} = B_t - 0.1U_t + 0.3K_t - 0.1C \cdot U_t, \quad K_{t+1} = K_t + 1.1U_t - 1.0C \cdot U_t$$

und in der differenziellen Form mit der zeitlichen Ableitung

$$\begin{aligned} B' &= (-0.1 - 0.1C)U + 0.3K \\ U' &= 1.0B + 1.0K \\ K' &= (1.1 - 1.0C)U \end{aligned} = \mathbf{A} \cdot (B, U, K)^T \tag{3.15}$$

Damit haben wir eine neue 3x3 Systemmatrix **A** und einen Wirkungsgraphen gefunden. In Abb. 3.6 ist der resultierende Wirkgraph gezeigt. Die kastenförmigen Umrandungen sollen andeuten, dass der jeweilige Einfluss darauf differentiell ist, also für die Änderung der Zustandsgrößen $z_1 = B$, $z_2 = U$, $z_3 = K$ gilt

$$\frac{dz}{dt} = \mathbf{A}z \tag{3.16}$$

wie in Gl. (3.5). Die Pfeile bedeuten hier eine Addition der Einflüsse; die Zahlen bzw. Variablenamen eine vorherige Multiplikation (Gewichtung) des Einflusses mit den angegebenen Werten.

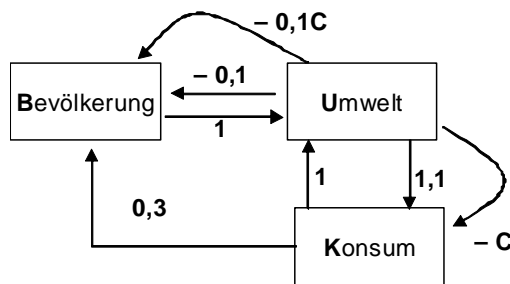


Abb. 3.6 Systemgraph der differentiellen Kopplung des Weltmodells

Die Systemmatrix **A** dieses Systems ähnelt sehr der aus Abb. 3.3, hat aber hier eine andere Bedeutung, da sie die differentielle Wirkung modelliert und nicht die absolute.

Wie können wir das Systemverhalten simulieren, wenn nur die Differentialgleichungen (3.16) davon gegeben sind? Zweifelsohne ist für das Zeichnen der Funktion $z(t)$ die explizite Angabe der Funktion $z(t)$ nötig, also ist die Funktion $z'(t)$ zu integrieren. Dies führt aber zu Problemen, wenn $z'(t)$ selbst wieder vom (unbekannten) $z(t)$ abhängig ist.

Um dieses Problem zu umgehen und die Formulierung aus Gl.(3.5) direkt zu verwenden, approximieren wir das gesuchte Integral: Nach dem Mittelwertsatz existiert ein t_0 aus $t \leq t_0 \leq t+\Delta t$, so dass

$$\frac{dz(t_0)}{dt} = \frac{z(t + \Delta t) - z(t)}{\Delta t} \quad (3.17)$$

gilt. Wählen wir ein sehr kleines Δt , so gilt dies approximativ auch für $t_0 = t$, und wir haben

$$z(t + \Delta t) \approx z(t) + \frac{dz(t)}{dt} \Delta t \quad \text{Euler-Cauchy-Integration} \quad (3.18)$$

und somit

$$z(t + \Delta t) := z(t) + \mathbf{A}z(t) \Delta t \quad (3.19)$$

Man beachte, dass dies nur für kleine Δt eine gute Approximation ist; je größer wir die Schritte machen, umso ungenauer ist die Integration. Wir müssen also möglichst viele Iterationsschritte bei diesem Verfahren durchführen.

Zur Illustration ist in Abb. 3.7 das Systemverhalten wieder am Beispiel der Umwelt- und Ressourcenbelastung gezeigt. Dabei wurde $\Delta t = 0.02$ und $B(0) = U(0) = K(0) = 1$ verwendet.

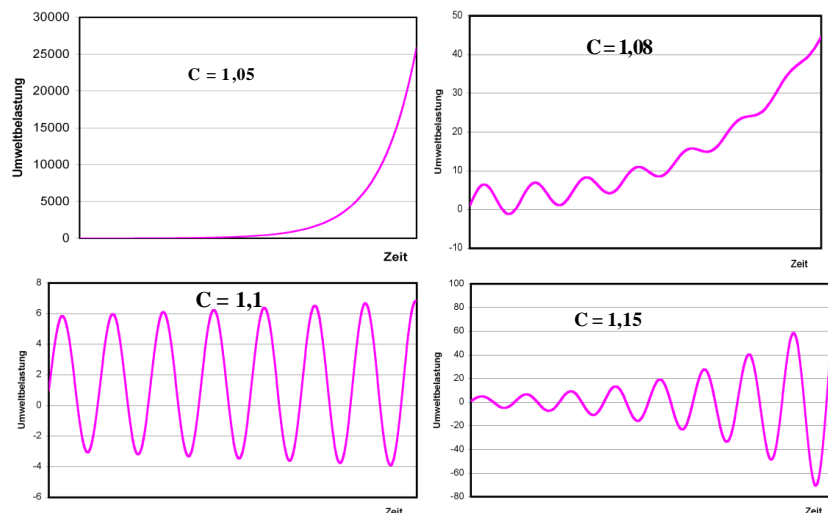


Abb. 3.7 Entwicklung der Umweltbelastung U mit der Zeit für verschiedene Werte von C

Vergleichen wir die Diagramme aus beiden Abbildungen, so fallen zum einen die Unterschiede auf: Für verschiedene Werte von C gibt es durchaus unterschiedliche dynamische Entwicklungen, was angesichts der unterschiedlichen Interpretation

der Parameter nicht verwundert. Interessant sind die Gemeinsamkeiten: Auch hier bemerken wir einen schmalen Bereich von C , in dem das System einigermaßen stabil ist. Außerhalb, für kleinere und größere Werte, wird das System instabil: entweder gleichsinnig oder alternierend (oszillierend) vergrößert sich die Umweltbelastung bis ins Unendliche.

Wann ist nun dieses System stabil? Es lässt sich zeigen, dass im allgemeinen linearen Fall von Gl. (3.5) die Stabilitätsbedingung (3.14) ersetzt wird durch die Bedingung

$$\operatorname{Re}(\lambda_i) < 0 \quad \lambda_i = \text{Eigenwerte von } \mathbf{A} \quad (3.20)$$

Im allgemeinen Fall kann die Systemmatrix \mathbf{A} komplexe Eigenwerte haben; für die Stabilität sind aber immer nur die reellen Komponenten wichtig; die imaginären Anteile sagen etwas über die Phasenlage bei Rückkopplungen aus.

Aufgabe 3.1

Man simuliere das Weltmodell als Pulspropagation und Zustandsänderung mit Hilfe eines *spread-sheet*-Programms und verifiziere die Diagramme in Abb. 3.4 und Abb. 3.7. Was passiert, wenn wir $\Delta t = 0,02$ vergrößern oder verkleinern? Sind die entstehenden Graphen gleich?

3.2 Dynamische Modellierung

Im vorigen Abschnitt haben wir eine dynamische Modellierung kennen gelernt, die auf einer linearen Systembeschreibung beruht. Dies ist ein nahe liegender Ansatz für eine erste, grobe Betrachtung. In Kapitel 2 hatten wir bereits gesehen, dass alle, auch nichtlineare Systeme als Taylorentwicklung geschrieben und dabei in einen linearen Teil und einen Restterm unterteilt werden können. Ist der Restterm gering, so ist die lineare Approximation sicher ein guter Ansatz.

Allerdings sind die meisten Systeme in der Realität mehr oder weniger nichtlinear. Wir wollen deshalb in diesem Kapitel genauer untersuchen, wie wir durch Unterteilung des Gesamtsystems in einzelne, überschaubare nichtlineare Subsysteme und ihre nichtlineare Modellierung die Gesamtkomplexität in den Griff bekommen und ein komplexes, nichtlineares System modellieren können. Dazu betrachten wir unser Beispiel des Weltmodells etwas genauer und untersuchen alle drei Kästen in Abb. 3.6. Beginnen wir mit der Bevölkerungsentwicklung.

3.2.1 Subsystem 1: Bevölkerungsentwicklung

Eine Bevölkerung existiert nicht einfach statisch so für sich, sondern sie hat eine interne Dynamik: Sie nimmt durch Geburten zu und durch Sterbefälle ab. Dabei ist sowohl die Zahl der Geburten als auch der Sterbefälle von der Bevölkerungsgröße abhängig. Dies bedeutet im einfachsten Fall einer ausgewogenen Bevölke-

zung, dass sowohl die Zunahme ΔB_B als auch Abnahme der Bevölkerung ΔB_D pro Jahr jeweils proportional zur Bevölkerungsstärke ist:

$$\frac{\Delta B_B}{\Delta t} \sim B \quad \text{oder} \quad \frac{\Delta B_B}{\Delta t} = C_B B \quad C_B = \text{Geburtenanteil in } \Delta t$$

und

$$-\frac{\Delta B_D}{\Delta t} \sim B \quad \text{oder} \quad \frac{\Delta B_D}{\Delta t} = -C_D B \quad C_D = \text{Sterbeanteil in } \Delta t$$

Der Geburten- bzw. Sterbeanteil fungiert dabei als Proportionalkonstante zwischen der Bevölkerungszunahme bzw. -abnahme. Beide Effekte überlagern sich

$$\Delta B = \Delta B_D + \Delta B_B = C_B B \Delta t - C_D B \Delta t = (C_B - C_D) B \Delta t$$

so dass im Zeitabschnitt Δt gilt

$$\frac{\Delta B}{\Delta t} = (C_B - C_D) B$$

Für sehr kleine Zeitabschnitte gehen wir auf den Grenzwert über

$$\lim_{\Delta t \rightarrow 0} \frac{\Delta B}{\Delta t} = \frac{dB}{dt} = (C_B - C_D) B \quad (3.21)$$

Dies hat nach Gl. (3.13) die Lösung

$$B(t) = B(0) e^{(C_B - C_D)t} \quad (3.22)$$

Entsprechend wächst ohne Nebenbedingungen die Bevölkerung bei $(C_B - C_D) > 0$ ins Unendliche und nimmt bei $(C_B - C_D) < 0$ ab zu null. In der Realität lässt sich im Mittel $C_B = 0,03$ und $C_D = 0,01$ annehmen.

Die eigentliche Bevölkerungsentwicklung aber wird durch die Nebenbedingungen wie Gesundheit (Umwelt), Ernährungsmöglichkeiten u.dgl. bestimmt, die außerhalb dieses Subsystems modelliert werden und in das Subsystem hineinwirken.

Den Systemgraphen (*Systemdiagramm*) dazu können wir aufzeichnen, indem wir zusätzlich zu den Kästen noch runde Kreise mit dem Multiplikationszeichen * einführen, siehe Abb. 3.8. Sie sollen Zwischengrößen visualisieren, die sich rechnerisch aus den Zustandsgrößen der Kästen durch Multiplikation ergeben. Wie vorher bedeutet dabei ein abgehender Pfeil vom Kasten, dass die Zustandsvariable wirkt, und ein eingehender Pfeil, dass die Wirkung differentiell ist. Die Größe, die neben dem Pfeil steht, wird vor der Operation im Kreis (hier: "Multiplikation") mit der Variablen multipliziert. Natürlich lässt sich dies auch implizit visualisieren durch die Notation einer Gewichtung eines Einflusses. Allgemein können wir aber in den runden Kreisen auch eine andere Funktion oder Operator vermerken, der die eingehenden Pfeile (Variablen) symmetrisch miteinander verknüpft. Beispielsweise lässt sich a-b als Addition (Operator +) auffassen, bei der ein Term mit -1 vorher multipliziert wurde.

In Abb. 3.8 ist links das resultierende Systemdiagramm unseres Subsystems und rechts die entsprechende, in zwei Teilgleichungen zerlegte Gl. (3.21) gezeigt.

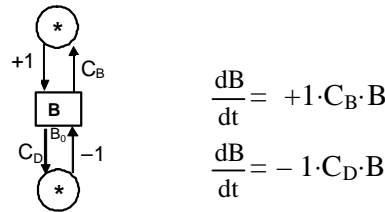


Abb. 3.8 Das Systemdiagramm des Subsystems 1 und seine Teilgleichungen.

Die Startgröße der Zustandvariable B ist dabei darunter mit B_0 notiert, um in Erinnerung zu rufen, dass die Zustandsentwicklung durch sie bestimmt ist.

3.2.2 Subsystem 2: Umweltbelastung

Auch die Umweltbelastung lässt sich durch innere Subsysteme beschreiben. Alle Abfallprodukte werden durch innere Prozesse (radioaktiver Zerfall) oder äußere Prozesse wie Wind, Wetter, Bakterien usw. abgebaut. Dabei ist der Abbau pro Zeit (*Abbaurate*) produktspezifisch und hängt von den Rahmenbedingungen ab. Ähnlich wie bei der Bevölkerung lässt sich dies als ein Prozess aus zwei unterschiedlichen Einflüssen betrachten: einen Abbau, der proportional zur vorhandenen Belastung ist, und einer konstanten, von außen vorgegebenen zusätzlichen Belastung S pro Zeiteinheit. Unsere Differentialgleichung für die innere Dynamik der Umweltbelastung U ist also

$$\frac{dU}{dt} = S - C_U U \tag{3.23}$$

Dabei müssen wir allerdings zusätzlich bedenken, dass die Belastung nicht bis Unendliche gesteigert werden kann: Wird ein Ökosystem überlastet, so kann nicht mehr, sondern maximal die gleiche Menge wie vorher abgebaut werden, falls das Ökosystem auch nicht schlechter wird oder zusammenbricht. Wir müssen also für unser System ein zweites Verhalten einführen, das bei Überlast $U \geq U^*$ auftritt.

$$\frac{dU}{dt} = S - C_U U \quad \text{bei } U < U^*, \text{ sonst } \frac{dU}{dt} = S - C_U U^* \tag{3.24}$$

In Abb. 3.9 ist das Systemdiagramm beider Gleichungen zu sehen, wobei die Vorgabe S bzw. U^* mit einem Sechseck als Einflussgröße visualisiert wird.

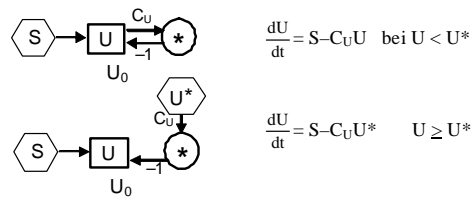


Abb. 3.9 Das Systemdiagramm beider Gleichungen

Schön wäre es, wenn wir statt zweier Gleichungen und eines geteilten Systemdiagramms nur eine benötigen würden, die aber mit einer Bedingung verknüpft sein muss um den qualitativen Umschwung zu bewirken. Dazu definieren wir uns die relative Größe $Q = U^*/U$, deren Unterschreiten von eins das Umschalten bewirken soll. Wir erhalten so die beiden Gleichungen

$$\frac{dU}{dt} = S - C_U U \quad \text{bei } Q \geq 1, \quad \text{sonst} \quad \frac{dU}{dt} = S - C_U U Q = S - C_U U \frac{U^*}{U} \quad (3.25)$$

Beim Umschalten wird also nur eine zusätzliche Größe anmultipliziert und keine ersetzt. Wie sieht das Systemdiagramm davon aus? Dazu müssen wir uns ein neues Grafikelemente definieren: Eine Bedingung sei mit einer Raute notiert. In Abb. 3.10 ist das Element und seine Interpretation visualisiert.

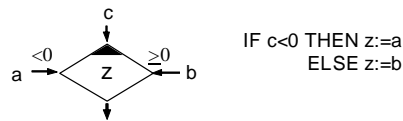


Abb. 3.10 Visualisierung der Entscheidung sowie ihre Definition

Mit dieser Definition lässt sich das Systemdiagramm aus zwei Systemen zu einem einzigen System zusammenfassen, siehe Abb. 3.9.

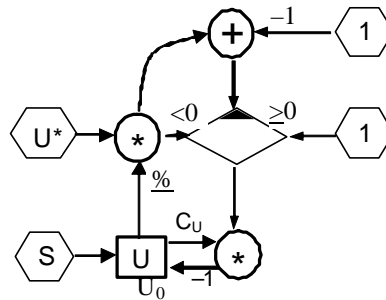


Abb. 3.11 Das Subsystem 2 und seine Zusammenfassung

Hier ist das Basissystem aus S und U durch eine zusätzliche Rechenoperationen angereichert, in der U* multipliziert wird (notiert durchs *-Zeichen) mit dem Quotienten 1/U von U, notiert durchs %-Zeichen. Dieser Quotient Q ist wirksam, wenn $Q-1 < 0$ ist, wird multipliziert mit $U \cdot C_U$ und von U abgezogen, wie es Gl. (3.22) verlangt. Ist $Q-1 \geq 0$, so wird stattdessen Eins mit $U \cdot C_U$ multipliziert und von U abgezogen, ebenso wie in Gl. (3.22) vorgeschrieben. Typische Werte sind $U_0 = 1$ und $C_U = 0,1$ [pro Jahr].

3.2.3 Subsystem 3: Konsumentwicklung

Der Konsum pro Kopf der Bevölkerung wächst in den meisten Ländern kontinuierlich an. Wählt man zur Modellierung einen reinen selbstbezüglichen Ansatz, etwa

$$\frac{dK}{dt} = C_K K,$$

so führt dies bekanntlich nur zu exponentiellem Wachstum, siehe Gl. (3.11) mit $w > 0$. Dies kann nicht sein, da jeglicher Konsum von Ressourcen entweder an die Endlichkeit der Ressourcen oder der Umweltbelastung stößt und zum Zusammenbruch des Systems führt. Gehen wir von endlichen Ressourcen aus, die einen maximalen Konsum K^* ermöglichen, so ist durch den Preisanstieg oder durch den wachsenden Schwierigkeitsgrad bei der Ressourcengewinnung (Ölreserven etc.) die Konsumrate bei der Annäherung an K^* gedämpft, etwa durch einen Faktor $(1 - K/K^*)$. Verwenden wir für unser Subsystem einen solchen Ansatz, so erhalten wir

$$\frac{dK}{dt} = C_K K \left(1 - \frac{K}{K^*}\right) = C_K K (1 - f \cdot K) \quad \text{mit } f = 1/K^* \quad (3.26)$$

Das dazu gehörende Systemdiagramm ist in Abb. 3.12 gezeigt.

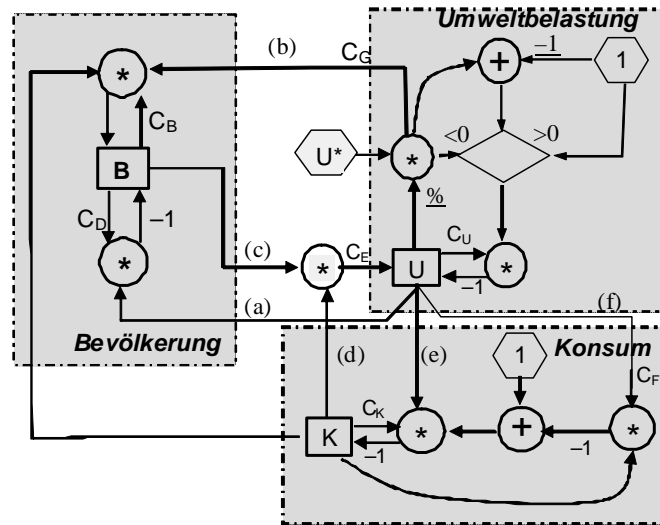


Abb. 3.13 Das aus den Subsystemen zusammengesetzte Gesamtsystem der Mini-Welt

Für die Kopplung des Subsystems „Bevölkerung“ mit dem Subsystem „Konsum“ ist der Einfluss des Konsumanstiegs durch höheres Einkommen auf die Lebensperspektive der Menschen zu beachten. Gehen wir davon aus, dass Kinder ein Sozialprestige darstellen, das „man“ sich bei höherem Lebensstandard „erlauben“ kann, so betrifft der Einfluss des Konsums direkt die Geburtenrate und geht multiplikativ dort ein.

Es bleibt nur noch, die Kopplung des Subsystems „Konsum“ mit dem Subsystem „Umweltbelastung“ zu modellieren. Hier finden wir

- (d) Die Auswirkungen des Konsums auf die Umweltbelastung ist zwar unabhängig von der Bevölkerungszahl, aber geht an der selben Stelle ein. Wir verknüpfen deshalb beide Einflüsse mit einander. Da $U \sim K$ und $U \sim B$ gelten, ist $\dot{U} \sim K \cdot B$; die Verknüpfung ist multiplikativ.
- (e) Die Umweltbelastung erhöht den Konsum, wenn mehr Medikamente (Anti-allergika), mehr medizinische Reparaturen und mehr Umweltschutzmaßnahmen nötig werden. Dies ist proportional zum Konsum, so dass der Einfluss multiplikativ mit dem vorhandenen Begrenzungsfaktor geschehen kann.
- (f) Da dadurch die Auswirkungen der Umweltschäden gemildert werden können und mehr Konsum möglich wird, hebt der Einfluss auch die obere Kapazitätsgrenze des Konsums an, so dass der Einfluss die Vorgabe f in Abb. 3.12 ersetzt.

Damit haben wir nach einer detaillierten Modellierung der einzelnen Subsysteme alles wieder zusammen gebaut. Dabei entstand aber ein gegenüber Abb. 3.6 leicht verändertes Gesamtbild, siehe Abb. 3.14.

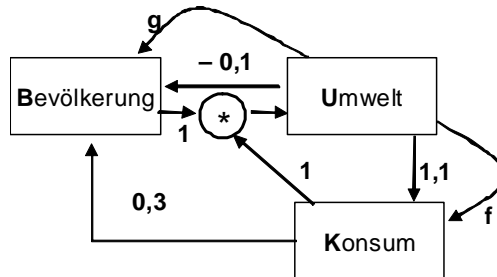


Abb. 3.14 Das veränderte Gesamtsystem

Der Unterschied der zusätzlichen Multiplikation ist aber nicht essentiell: Für kleine Änderungen ΔB und ΔK lässt sich die Multiplikation linearisieren. Bei einer Ausgangsgröße von $B = K = 1$ ist

$$(1+\Delta B)(1+\Delta K) = 1 + \Delta B + \Delta K + \dots \quad (3.27)$$

wobei das Produkt der beiden kleinen Terme vernachlässigt wird. Damit geht das Systemdiagramm aus Abb. 3.14 in das Abb. 3.6 aus über.

Die korrespondierenden Differentialgleichungen des Gesamtsystems sind auch entsprechend verändert. Betrachten wir Abb. 3.13, so wird aus Gl. (3.21)

$$\frac{dB}{dt} = C_B B K C_G \frac{U^*}{U} - C_D B U \quad (3.28)$$

Aus den beiden Gleichungen (3.25) wird bei Ersetzen von S durch den kombinierten Einfluss aus Bevölkerung und Konsum

$$\frac{dU}{dt} = C_E B K - C_U U \text{ bei } U < U^*, \text{ sonst } \frac{dU}{dt} = C_E B K - C_U U^* \quad (3.29)$$

und aus Gl. (3.26) wird

$$\frac{dK}{dt} = U C_K K (1 - U C_F \cdot K) \text{ mit } C_F = 1/K^* \quad (3.30)$$

Als Parameter wählen wir $C_U = 0.1$, $C_B = 0.03$, $C_K = 0.05$, $C_D = 0.01$, $C_E = 0.02$ sowie den Anfangsbedingungen $B_0 = 1$, $U_0 = 1$, $K_0 = 1$. Als Steuerparameter verwenden wir C_G ungefähr eins und C_F aus $[0..1]$. In Abb. 3.15 (a) ist mit $\Delta t = 0,1$ und 500 Datenpunkten die Simulation für 50 Jahre bei unbegrenztem Wachstum ($K^* = \infty$, $C_F = 0.0$) gezeigt.

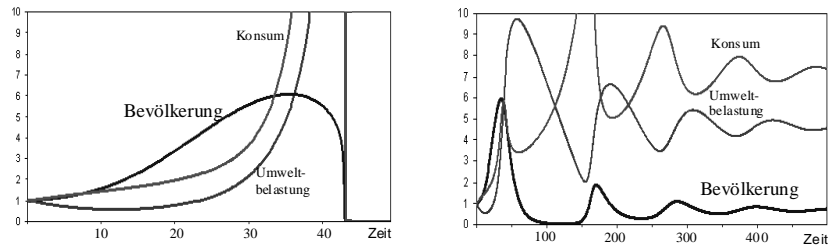


Abb. 3.15 Systementwicklung bei (a) $C_F = 0.0$ und (b) $C_F = 0.03$

Man sieht, dass die Entwicklung zunächst langsam ansteigt und dann in einem geringen Zeitraum stark wächst: Das exponentiale Wachstum macht sich bemerkbar. Entsprechend verlangsamt sich das Bevölkerungswachstum und sackt dann durch den Sättigungseffekt steil ab mitsamt Konsum.

Im Gegensatz dazu ist bei $C_F = 0.03$, also einem gedämpften Wachstum durch den negativen Einfluss, zwar ebenfalls ein Absinken der Bevölkerungszahl zu bemerken, aber sie erholt sich nach einer gewissen Zeit (ca. 120 Jahre) wieder und bewirkt ein erneutes Absacken. Dies wiederholt sich, so dass ein periodisches Schwanken aller Zustandsvariablen zu beobachten ist. In Abb. 3.15 (b) ist dies für den Zeitraum von 500 Jahren mit 5000 Datenpunkten und $\Delta t = 0,1$ gezeigt. Erst eine stärkere Konsumbremse ($C_F = 1.0$) stabilisiert unsere Mini-Welt in einem überschaubaren Zeitraum, siehe Abb. 3.16.

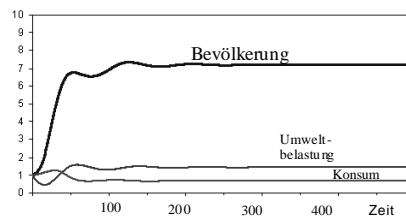


Abb. 3.16 Systementwicklung bei $C_F = 1.0$

Allerdings wird die schnelle Stabilität bei hoher Bevölkerung durch geringen Konsum erkaufte. Geringere Werte von C_F erzeugen stärkere Schwankungen, stabilisieren aber auf lange Sicht die Bevölkerung auf niedrigerer Zahl und damit auf höherem Konsumniveau.

Wir erhalten also die Erkenntnis, dass in diesem Modell technisch verschiedene Stabilitätsszenarien möglich sind, die erst durch ihre gesellschaftliche Bewertung eine Auswahl erfahren. Hierbei sind nicht die genauen absoluten Zahlenwerte wichtig (sie hängen stark von den verwendeten Werten der Parameter ab), sondern das grundsätzliche Verhalten des Gesamtsystems. Die Effekte und Nebeneffekte eines gesellschaftlich motivierten Eingriffs in das Mini-Weltmodell können so

besser abgeschätzt werden und liefern Antworten auf grundsätzliche politisch-soziale Fragen.

3.2.5 Diskussion

Bei allen Modellierungen müssen wir sehr vorsichtig sein: Viele Annahmen sind nicht unbedingt realitätsnah. Schlussfolgerungen aus dem Gesamtverhalten können deshalb falsch sein. So kann man für das konkrete Beispiel des Mini-Welt-systems auch ganz andere Kopplungen aufbauen. Beispielsweise ist in Entwicklungsländern zu beobachten, dass bei erhöhtem Einkommen, besserer medizinischer Versorgung und Altersversorgung weniger Kinder geboren werden, da die Altersversorgung durch ausreichend überlebende Kinderzahl nicht mehr nötig wird. Dies bedeutet, dass der Konsum sich nicht positiv, sondern negativ auf die Geburtenrate auswirkt.

Auch wird nicht beachtet, dass der Konsum normalerweise direkt von der Bevölkerung erzeugt wird. Durch sein Anwachsen wird daraufhin eine Umweltbelastung erzeugt, nicht umgekehrt.

Eine alternative Modellierung bei Umkehr von Kausalitäten ist aber nicht zwingend notwendig: Alle parallel ablaufenden Prozesse, deren Korrelation wir beobachten, sind nicht unbedingt kausal miteinander gekoppelt, sondern nur statistisch. Dies bedeutet, dass wir zwar das System durch eine angenommene Abhängigkeit erfolgreich modellieren können, aber diese logisch nicht unbedingt stimmen muss. Die Beschreibung stimmt solange mit der Realität überein, wie beide Zustandsgrößen miteinander stark korreliert sind und keine Zeitverschiebung sichtbar wird.

3.3 Grundelemente dynamischer Modellierung

In unserem Beispiel des Mini-Weltmodells im vorigen Abschnitt lernten wir verschiedene Elemente dynamischer Modellierung kennen. So konnten wir unterscheiden zwischen

- **Zustandsvariable** $\mathbf{z}(t) = (z_1(t), \dots, z_m(t))^T$

Diese Variablen lassen sich nicht als Funktion anderer errechnen, sondern speichern einen augenblicklichen Zustand, etwa die Höhe eines Wasserstandes, eine Bevölkerungszahl oder die Geschwindigkeit eines Autos. Sie entsprechen den Zuständen bei diskreten endlichen Automaten, haben aber nicht endlich viele Zustände, sondern unendlich viele.

- **Eingabevariable** $\mathbf{x}(t) = (x_1(t), \dots, x_n(t))^T$

Die Eingabevariablen entsprechen einem Kontext bzw. Vorgaben, die in das System eingehen und sein Verhalten mehr oder weniger bestimmen. Eine

Untermenge der Eingabevariable sind die Parameter und Konstanten, die in den Funktionen des Systems genutzt werden.

- **Ausgabevariable** $\mathbf{y}(t) = (y_1(t), \dots, y_n(t))^T$
Die Ausgabevariablen sind die Resultate der Systemaktion nach außen und werden als **Verhalten** wahrgenommen.

Der Zustand im System zu einem Zeitpunkt ergibt sich aus der Eingabe und dem augenblicklichen Zustand im System

$$\mathbf{z}(t) = \mathbf{F}(\mathbf{z}(t), \mathbf{x}(t), t) \quad \text{Zustandsgleichung} \quad (3.31)$$

Da der Folgezustand eine Zusammenfassung mehrerer Einzelzustände zu einem Vektor \mathbf{z} ist, besteht die Funktion \mathbf{F} auch aus mehreren Funktionen, zusammengefasst zu einem Vektor. Mit den Ableitungen $\dot{\mathbf{z}} = d\mathbf{F}_i(\cdot)/dt$ gilt entsprechend

$$\frac{d\mathbf{z}}{dt} = \mathbf{f}(\mathbf{z}(t), \mathbf{x}(t), t) \quad \text{differentielle Zustandsgleichung} \quad (3.32)$$

Auch die Ausgabe lässt sich derart formulieren:

$$\mathbf{y}(t) = \mathbf{g}(\mathbf{z}(t), \mathbf{x}(t), t) \quad \text{Ausgabegleichung} \quad (3.33)$$

Ein System lässt sich mit diesen drei Begriffen als Blockdiagramm visualisieren, siehe Abb. 3.17.

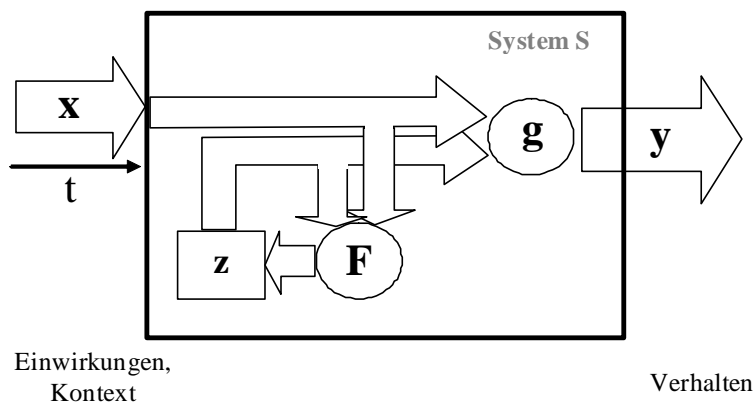


Abb. 3.17 Allgemeines Diagramm eines dynamischen Systems

Typisches Beispiel dafür ist ein lineares diskretes System, wie wir es in Abschnitt 3.2 kennen gelernt haben. Zusammen mit der Eingabe \mathbf{x} erhalten wir für den neuen Zustand \mathbf{z}

$$\mathbf{z}(t+1) = \mathbf{z}(t) + \mathbf{A}\mathbf{z}(t) + \mathbf{B}\mathbf{x}(t) = (\mathbf{I}+\mathbf{A}) \mathbf{z}(t) + \mathbf{B}\mathbf{x}(t) \quad (3.34)$$

Die Simulation eines solchen Systems findet dabei in folgenden Schritten statt:

- (1) Vorgabe des Kontext: Alle Anfangswerte $\mathbf{z}(t_0)$ sowie alle festen Parameter (Kontext)
- WHILE $t < t_{\max}$ DO
- (2) Ermittlung der Eingabe und zeitabhängigen Kontextes $\mathbf{x}(t)$
- (3) Berechnung der Veränderungsraten $\mathbf{z}'(t) = \mathbf{f}(\mathbf{z}, \mathbf{x}, t)$
- (4) Integration der Raten für den neuen Zustand $\mathbf{z}(t) = \mathbf{z}(t_0) + \int_0^t \mathbf{z}'(t') dt'$
- (5) Berechnung der Ausgabe $\mathbf{y}(t) = \mathbf{g}(\mathbf{z}, \mathbf{x}, t)$
- ENDWHILE

Die Dynamik des Systems wird dabei durch die Anfangswerte und die Zustandsentwicklung in (4) bestimmt; die Ausgabe hat dabei keinen Einfluss. Wird die Simulation unterbrochen, so genügt es, die Zustandsgrößen zu speichern; die Ausgabe ergibt sich dann automatisch. Die Zustände bestimmen also das Verhalten des Systems; die Anzahl m der Zustände wird deshalb auch als **Dimension** des Systems bezeichnet. Im Unterschied dazu ergeben alle Rechenoperationen nur abgeleitete Größen.

Im Unterschied zu instantanen, trägheitslosen, sich sofort aus den Zuständen und Ausgaben ergebenden Variablen wie die beobachtete Ausgabe $\mathbf{y}(t)$ ergeben sich die Zustände selbst nicht sofort aus den jetzigen Zuständen, sondern erst mit einer gewissen Trägheit. Dies entspricht der Realität, dass die Höhe eines Stausees nicht von einem Augenblick auf den nächsten um die Hälfte gesunken sein kann, oder ein Auto von null auf 100 Km/h in einer Nanosekunde beschleunigt haben kann. In beiden Fällen sind Materialtransportphänomene involviert, also physikalische Größen, die eine gewisse Trägheit besitzen. Die Zustandsvariablen modellieren also meist Speichergrößen des Systems wie Energie, Masse, Information, Populationen, und alle Variablen, die Trägheit oder Verzögerungen ausdrücken wie Förderbänder, Warteschlangen oder Transportverzögerungen.

Lässt sich der Zustand eines Systems durch eine Zustandsvariable charakterisieren, so zeigt diese Variable zwar typische Werte des Systems an, aber sie ist nicht eindeutig bestimmt. Es sind auch andere Variable als Zustandsvariable denkbar, die mit dieser Variablen verbunden sind.

Beispiel Speicherbecken

Der Wasserzustand in einer Wasserspeicher kann sowohl durch die Wasserstandshöhe h (bei Breite b und Länge g) als auch durch das Volumen $V = h \cdot b \cdot g$ oder durch seine Wassermasse in [kg/l] mit $M = a \cdot V = a \cdot h \cdot b \cdot g$ charakterisiert werden.

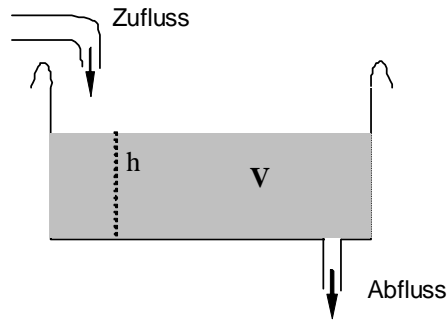


Abb. 3.18 Wasserstand und Zustandsvariable bei einer Badewanne

Wir haben zwar für die Zustandsvariable z die drei möglichen Kandidaten h , V und M , aber es reicht, nur eine einzige davon anzugeben. Alle anderen lassen sich aus der einen errechnen: Das System hat die Ordnung eins.

Betrachten wir nun das Systemverhalten. Nehmen wir an, dass bei einem konstanten Zufluss eine konstante Menge r_z an Wasser pro Zeiteinheit hineinfließt und bei festem Abflussdurchmesser auch eine konstante Menge r_a pro Zeiteinheit abfließt, so ist die Gesamtveränderung pro Zeiteinheit

$$\lim_{\Delta t \rightarrow 0} \frac{\Delta V}{\Delta t} = \frac{dV}{dt} = r_z - r_a \quad (3.35)$$

Integration beider Seiten der Gleichung gibt uns das Systemverhalten

$$V(t) = V_0 + (r_z - r_a)t = V_0 + V_z(t) - V_a(t) . \quad (3.36)$$

Das Gesamtvolumen ergibt sich aus dem Anfangsvolumen, dem Zuflussvolumen und dem Abflussvolumen. Je nach dem, ob die Differenz zwischen Zuflussrate und Abflussrate positiv oder negativ ist, nimmt das Gesamtvolumen zu oder ab solange, bis entweder die Badewanne überläuft oder leer ist. Entsprechend unserer Konvention nach Gl. (3.16) lässt sich dies durch einen Kasten darstellen, der gewichtete Eingänge hat, siehe Abb. 3.22. Der Faktor eins wird in der Zeichnung nicht notiert.

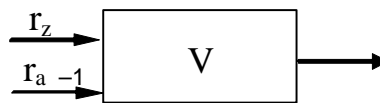


Abb. 3.19 Das integrative Element im Systemdiagramm

Im Folgenden wollen wir unter dem Gesichtspunkt der Anzahl der Dimensionen oder Zustandsvariablen einige elementare Systeme genauer ansehen. Beginnen wir mit dem einfachsten Fall.

3.3.1 Zustandsnormierung

Bei der Aufstellung von Wirkungsgraphen und Modellgleichungen kann es leicht passieren, dass man unsinnige Beziehungen modelliert, wenn man die inhärenten Dimensionen (Einheiten) der verwendeten Variablen und Konstanten nicht beachtet.

Beispiel *kinetische Energie*

Das Wortmodell für die Wirkungen zwischen kinetische Energie einer Masse und der Geschwindigkeit könnten lauten:

- Je größer die Masse m , umso größer die kinetische Energie E
- Je größer die Geschwindigkeit v , umso größer die kinetische Energie E

Die Modellierung könnte dann sein

$$E \sim m \cdot v \text{ bzw. } E = a \cdot m \cdot v$$

Die Prüfung der Dimensionen zeigt: E hat die Einheit Joule = $\text{kg m}^2/\text{s}^2$, während m in kg und v in m/s gemessen wird. Für eine dimensionslose Proportionalitätskonstante sollte aber noch ein Faktor m/s vorhanden sein. Daraus können wir schließen, dass wir besser eine Modellierung mit v^2 vornehmen sollten.

In diesem Fall zeigt also die Dimensionsprüfung nicht nur, ob wir bei den Variablen m und v die richtigen Einheiten verwendet haben (nämlich Meter und Sekunde und nicht km und Minuten), sondern gibt auch einen Hinweis auf Inkonsistenzen in der Modellierung.

Aus diesem Grund sollte man bei der Aufstellung der Modellgleichungen darauf achten, bei den Variablen und Konstanten die Dimensionen richtig zu wählen. Eine elegante Methode dafür besteht darin, alle Zustände grundsätzlich dimensionslos zu machen, indem man die Variablen durch einen Wert von ihr zu einem bestimmten Zeitpunkt, etwa t_0 , oder an einem Gleichgewichtspunkt, teilt.

Beispiel *Volumenänderung*

Das Volumen in Gleichung (3.36) habe die Einheit Liter oder Kubikmeter. Das Verhalten lässt sich einfach normieren, indem wir die neue dimensionslose Variable $U(t) = V(t)/V(0)$ einführen

$$U(t) = V(t)/V(0) = 1 + U_z(t) - U_a(t) \tag{3.37}$$

Dies bedeutet nicht, dass man die Variablen ohne Beachtung der Konstanten ansetzen soll oder nicht auf die Konsistenz der Einheiten achten müsste, in denen die Konstanten gemessen werden. Die dimensionslosen Differentialgleichungen der dynamischen Modellierung gestatten eine Formulierung, die unabhängig vom Problem eine bessere Wiedererkennung der Formulierung der eigentlichen Systemdynamik in vielen Anwendungen erlaubt.

Allgemein ersetzen wir in der Zustandsgleichung (3.32) und der Ausgabegleichung (3.33) jede Komponente $z_i(t)$ durch $k_i \cdot u_i(t)$ bzw. $z_i'(t)$ durch $k_i \cdot u_i'(t)$, wobei k_i jeweils eine typische, Einheiten-behaftete Konstante ist.

Auch die Zeit lässt sich so normieren: Skalieren wir die Zeit neu mit einem Faktor T , etwa der typischen Zeit zwischen zwei Ereignissen, etwa zwei Schwingungsmaxima, so erhalten wir eine normierte Zeit

$$\tau = t/T \quad \text{oder} \quad t = T\tau \quad \text{und} \quad dt = T d\tau \tag{3.38}$$

Dies setzen wir ebenfalls für die Zeit t ein, soweit sie explizit in den Gleichungen auftaucht. Für die Ableitungen gilt dann statt

$$\frac{dz_i}{dt} = f_i(\mathbf{z}(t), \mathbf{x}(t), t) \quad i\text{-te Zustandsgleichung} \tag{3.39}$$

die Gleichung

$$\frac{du_i}{d\tau} = \frac{T}{k_i} f_i(k_1 u_1(t), \dots, k_n u_n, \mathbf{x}(t), T\tau) \quad \text{normierte Zustandsgleichung} \tag{3.40}$$

Damit ist eine dimensionslose, normierte Lösung möglich. Eine diskrete Aussage für die konkrete Problemstellung erhält man aber erst, wenn man alle Ergebnisse der Lösung bzw. Simulation wieder in den dimensions- und Einheiten-behafteten Raum zurück transformiert hat.

3.3.2 Systeme ohne Zustandsvariable

Im einfachsten Fall haben wir überhaupt keine Zustandsvariable. Dies bedeutet, dass wir auch keine Veränderungsrate und dazu gehörenden Differentialgleichungen haben. Alle Ausgaben lassen sich direkt durch die Eingaben erschließen

$$\mathbf{y}(t) = \mathbf{g}(\mathbf{x}(t), t) \tag{3.41}$$

Ein Beispiel dafür ist die Funktion $y = a \sin^2(kt)$, deren Systemdiagramm in gezeigt ist.

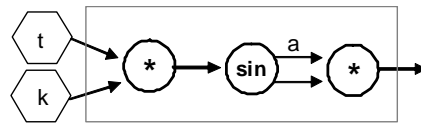


Abb. 3.20 Systemdiagramm eines Systems ohne Zustandsvariable

3.3.3 Systeme mit einer Zustandsvariablen

Eine der wichtigsten Elementarsysteme sind Systeme, die exponentiell wachsen oder abfallen.

Exponentielles Wachstum

Typisches Beispiel für ein Wachstum ist ein System, dessen Wachstumsrate vom Zustand selbst abhängig ist

$$z'(t) = w \cdot z(t) \quad \text{exponentielles Wachstum} \quad (3.42)$$

Die Lösung der Differentialgleichung und ihr Verhalten haben wir bereits in Gl. (3.11) und in Abb. 3.5 kennen gelernt. Eine interessante Variante des exponentiellen Abfalls ist gegeben, wenn zusätzlich ein additiver Term existiert, etwa in der Form

$$z'(t) = a(t) - w \cdot z(t) \quad \text{exponentielle Verzögerung} \quad (3.43)$$

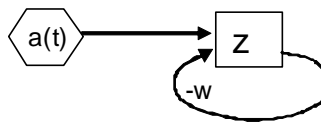


Abb. 3.21 Systemzustandsänderung durch verzögerte Rückkopplung

Betrachten wir dazu als konkretes, systemtypisches Beispiel das Speicherbecken aus Abb. 3.18. Nehmen wir dazu an, dass der Abfluss des Wassers aus dem Speicherbecken (d.h. Abfall dV_a/dt) mit sinkender Wasserhöhe h durch den Druckabfall $p \sim h$ auch proportional geringer wird. Dann ist mit $V(t) = b \cdot g \cdot h$ die entsprechende Differentialgleichung (3.35) mit der Proportionalitätskonstanten T für die nun zeitabhängige Abflussrate $r_a(t)$

$$\frac{dV_a(t)}{dt} \sim p \sim h = \frac{V(t)}{bg} \quad \text{oder} \quad \frac{dV_a(t)}{dt} = \frac{1}{T} V(t) =: r_a(t) \quad (3.44)$$

Die ursprüngliche Gleichung (3.35) wird damit zu

$$\frac{dV}{dt} = r_z - r_a(t) = r_z - \frac{1}{T} V(t) \quad (3.45)$$

entsprechend der Aufgabenstellung in Gl. (3.43).

Zur Lösung dieser Differenzialgleichung definieren wir uns die Hilfsvariable z

$$z(t) := \frac{dV}{dt} = r_z - \frac{1}{T} V(t) \quad \Rightarrow \quad \frac{dz(t)}{dt} = -\frac{1}{T} \frac{dV}{dt} \quad (3.46)$$

und erhalten durch Einsetzen der Definition für z(t)

$$-\frac{dz(t)}{dt} = \frac{1}{T} \frac{dV}{dt} = \frac{1}{T} z(t) \quad \text{oder} \quad \frac{\dot{z}}{z} = -\frac{1}{T}$$

Die Integration beider Seiten ergibt mit $z_0 := z(t_0) = r_z - V(t_0)/T$

$$\int_{t_0}^t \frac{\dot{z}}{z} dt' = \ln z - \ln z_0 = -\int_{t_0}^t \frac{1}{T} dt' = -\frac{t-t_0}{T},$$

und beide Seiten exponiert

$$z(t) = z_0 e^{-(t-t_0)/T}$$

Setzen wir die Definition (3.46) für z(t) ein, so erhalten wir

$$r_z - \frac{1}{T} V(t) = z(t) = z_0 e^{-(t-t_0)/T} \quad \Rightarrow \quad \frac{1}{T} V(t) = r_z - z_0 e^{-(t-t_0)/T}$$

$$V(t) = r_z T - z_0 T e^{-(t-t_0)/T}$$

und mit der Definition der Konstanten

$$A_0 := r_z T, \quad -z_0 T = -r_z T + V_0 = V_0 - A_0 := A$$

die Lösung

$$V(t) = A_0 + (V_0 - A_0) e^{-(t-t_0)/T} = A_0 + A e^{-(t-t_0)/T} \quad (3.47)$$

Dies entspricht einer "Abklingkurve", beispielsweise der Entladung eines Kondensators, mit der Zeitkonstante T von einem initialen Wert V_0 bis auf einen festen Wert A_0 . In der folgenden Abb. 3.22 ist dies dargestellt.

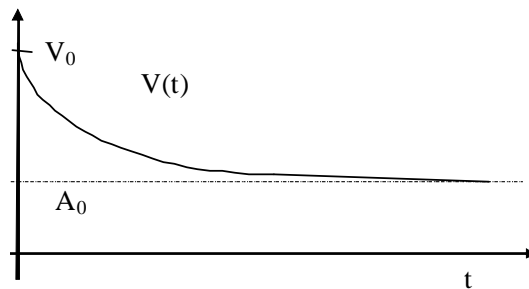


Abb. 3.22 Visualisierung der Differenzialgleichung durch Zufluss- und Abflussverhalten

Ist $V_0 < A_0$, so ist A negativ und die Kurve nähert sich der Geraden durch A_0 nicht von oben, sondern von unten an. Das Gesamtsystem wirkt also wie eine (exponentielle) Verzögerung, die das System beim Zustandsübergang vom Zustand $V = V_0$ zum Zustand $V = A_0$ erfährt.

Das Zuflussvolumen $V_z(t) = r_z \Delta t = r_z \cdot (t-t_0)$ ist linear, während das Abflussvolumen $V_a(t)$ vom Gesamtzustand $V(t)$ abhängig ist. Wie groß ist nun $V_a(t)$? Mit der Gesamtbilanz

$$V(t) = V_0 + V_z(t) - V_a(t)$$

ergibt sich durch Gl. (3.47)

$$\begin{aligned} V_a(t) &= V_0 + V_z(t) - V(t) = V_0 + r_z(t-t_0) - r_z T - (V_0 - r_z T) e^{-(t-t_0)/T} \\ &= r_z(t-t_0) - r_z T + r_z T e^{-(t-t_0)/T} + V_0 - V_0 e^{-(t-t_0)/T} \\ &= r_z(t-t_0) + (V_0 - r_z T)(1 - e^{-(t-t_0)/T}) \end{aligned} \quad (3.48)$$

Logistisches Wachstum

Ein weiteres Beispiel für Systeme mit einer Zustandsvariablen haben wir bereits beim Subsystem 3 (Konsumententwicklung) kennen gelernt. Hierbei ist die Zustandsänderung nicht nur proportional zu dem Zustand selbst, sondern hängt auch von den noch zur Verfügung stehenden Ressourcen nach Abzug der bestehenden Ressourcen ab. Dies lässt sich formulieren in der Form

$$z'(t) = az \cdot (1-bz) \quad \text{logistisches Wachstum} \quad (3.49)$$

mit dem maximalen Wert $z^* = 1/b$ für z , so dass die noch zur Verfügung stehenden Ressourcen als Differenz $(1-bz)$ auftreten und bei wachsendem z das Wachstum z' gegen null gehen lassen, bis $bz = 1$ gilt. Ist $z < 1/b = z^*$, so ist mit $z' > 0$ ein

Wachstum für z gegeben; ist $z > z^*$, so wird das Wachstum negativ und z nimmt ab. Es gibt also einen stabilen Punkt bei dieser Differentialgleichung, zu dem $z(t)$ hinstrebt; er liegt bei $z = z^*$. In Abb. 3.23 ist die dynamische Entwicklung von $z(t)$ gezeigt. Setzen wir einen Wert für $z(t_0)$ zum Zeitpunkt t_0 an, so kann man die Entwicklung von $z(t > t_0)$ in der Abbildung ab dem Zeitpunkt t_0 ablesen.

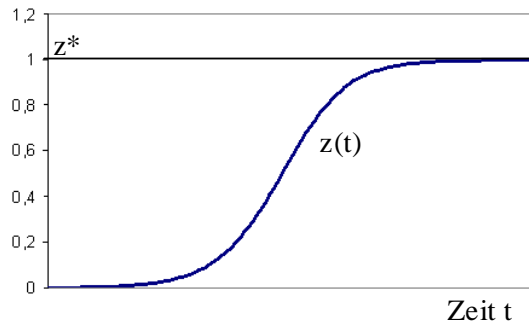


Abb. 3.23 Logistische Wachstumskurve für $z^*=1$

Aufgabe 3. 2

Betrachten Sie als Zustandsvariable nicht das Volumen, sondern die Höhe h . Wie lautet $h(t)$?

Aufgabe 3. 3

Nehmen Sie an, dass bei dem Speicherbecken in Abb. 3.18 der Zufluss nicht konstant ist, sondern durch ein Rohr aus einem anderen daneben gelegenen, immer auf eine Höhe h_{\max} gefüllten Speicherbecken erfolgt. Damit ist die Zuflussrate linear abhängig von der Differenz ($h_{\max}-h$) zwischen der Wasserstandshöhe h und dem maximalen Wert h_{\max} : je höher der Wasserstand, desto geringer der Zufluss.

- (a) Wie lautet nun die Differentialgleichung und ihre Lösung, wenn die Abflussrate konstant ist?
- (b) Wie verändern sich diese, wenn die Abflussrate proportional zur Wasserstandshöhe ist wie in Gl. (3.44) ?

Die bisher besprochene Aufgabenstellung benutzte eine Zustandsvariable. Betrachten wir nun typische Systeme mit zwei Zustandsvariablen.

3.3.4 Systeme mit zwei Zustandsvariablen

Bei zwei Zustandsvariablen $u(t)$ und $v(t)$ können getrennte Systeme vorliegen; die Änderungsraten u' und v' sind dann jeweils nur von der eigenen Variablen u bzw. v abhängig. In diesem Fall haben wir zwei einzelne Systeme, die jeweils für sich wie ein System mit nur einer Zustandsvariablen betrachtet werden können. Dieser Fall ist aber langweilig. Viel interessanter ist es, wenn die Variablen sich gegenseitig beeinflussen, also

$$\begin{aligned} u' &= f(u,v) \\ v' &= g(u,v) \end{aligned} \quad (3.50)$$

und beispielsweise linear

$$\begin{aligned} u' &= v \\ v' &= -a \cdot u \end{aligned} \quad (3.51)$$

Welche Lösung hat dieses Gleichungssystem? Dazu bilden wir die zweite Ableitung und setzen die bekannten Ausdrücke ein

$$\begin{aligned} u'' &= v' = -a \cdot u \\ v'' &= -a \cdot u' = -a \cdot v \end{aligned} \quad (3.52)$$

Wir erhalten für beide Zustandsvariablen gleiche Beziehungen vom Typ $x'' = -ax$. Eine Lösung für diese Differentialgleichung können wir erraten zu

$$x(t) = A \sin(\varphi + \omega t) \quad (3.53)$$

Wir prüfen nach, ob dies tatsächlich eine Lösung ist. Es gilt mit Gl.(3.53)

$$x' = A\omega \cos(\varphi + \omega t), \quad x'' = -A\omega^2 \sin(\varphi + \omega t) \Rightarrow x'' = -ax$$

mit $a = \omega^2$. Die Lösung ist also korrekt. Damit können wir folgern, dass die Zustände periodisch oszillieren; folgt $u(t)$ einer Sinusfunktion, so ist $v(t)$ eine negative Kosinusfunktion, also eine um den Phasenwinkel $\varphi = 90^\circ$ verzögertes Signal $u(t)$.

Wir können auch eine weitere Lösung der Differentialgleichung erraten, die Kosinusfunktion:

$$x(t) = A \cos(\varphi + \omega t) \Rightarrow x' = -A\omega \sin(\varphi + \omega t), \quad x'' = -A\omega^2 \cos(\varphi + \omega t)$$

Also ist auch die Überlagerung beider Lösungen eine Lösung von Gl.(3.52)

$$x(t) = A_1 \sin(\varphi_1 + \omega t) + A_2 \cos(\varphi_2 + \omega t) \quad (3.54)$$

$$\Rightarrow x' = A_1 \omega \cos(\varphi_1 + \omega t) - A_2 \omega \sin(\varphi_2 + \omega t)$$

$$x'' = -A_1 \omega^2 \sin(\varphi_1 + \omega t) - A_2 \omega^2 \cos(\varphi_2 + \omega t) = -\omega^2 x$$

Wir können übrigens die Oszillationen dämpfen, indem wir einen exponentiellen Abfall bei $v(t)$ einbauen, etwa entsprechend Gl.(3.45). Wir erhalten dann

$$\begin{aligned} u' &= v \\ v' &= -a \cdot u - b \cdot v \end{aligned}$$

Stabilität Linearer Systeme

Allgemein können wir ein solches lineares Differentialsystem für zwei Variablen schreiben als

$$\begin{aligned} u' &= a \cdot u + b \cdot v \\ v' &= c \cdot u + d \cdot v \end{aligned} \tag{3.55}$$

In Abb. 3.24 ist der Systemgraph eines linearen Systems aus zwei Zustandsvariablen zu sehen.

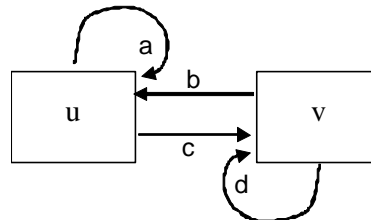


Abb. 3.24 Lineares System für zwei Zustandsvariable

Mit $\mathbf{z} = (u, v)^T$ und den in der 2×2 Matrix \mathbf{A} zusammengefassten Koeffizienten a, b, c, d kann man dies allgemein formulieren als

$$\mathbf{z}' = \mathbf{A} \mathbf{z} \tag{3.56}$$

Lineares System

Wann ist dieses lineare System stabil? Wann schwingt es?

Die obige Gleichung gilt für alle Vektoren \mathbf{z} , auch für die Eigenvektoren von \mathbf{A} . In diesem Fall haben wir

$$\mathbf{z}' = \mathbf{A} \mathbf{z} = \lambda \mathbf{z} \tag{3.57}$$

Die Lösung des Systems ist hierbei analog zu Gl. (3.11) und (3.13)

$$\mathbf{z}(t) = \mathbf{z}_0 e^{\lambda t} \tag{3.58}$$

mit komplexem $\lambda = r + i\omega$.

Ist der Eigenwert λ reell und damit $\omega = 0$, so haben wir nur reelles r und der Zustandsvektor erfährt bei $r > 0$ exponentielles Wachstum, bei $r < 0$ exponentiellen Abfall und bei $r = 0$ konstante Werte, siehe Abb. 3.5.

Ist allerdings nur einer der Eigenwerte komplex, so fängt das System an, periodisch zu schwingen: Für $\omega \neq 0$ können wir den Faktor $e^{\lambda t}$ in der Eulerschen Form formulieren

$$e^{\lambda t} = e^{(r+i\omega)t} = e^{rt} e^{i\omega t} = e^{rt} (\cos \omega t + i \sin \omega t) \quad (3.59)$$

Zusätzlich zu dem reellen Faktor, der exponentielles Wachstum oder Abfall bedeutet, haben wir hier noch einen periodisch schwingenden Anteil. Ist der Imaginärteil ω von λ aber null, so schwingt das System nicht.

Entscheidend für das Verhalten des linearen Systems sind also seine Eigenwerte. Für das konkrete System aus Gl. (3.55) errechnen wir sie als Lösung der Gleichung (3.57)

$$\mathbf{A} \mathbf{z} = \lambda \mathbf{z} \quad \text{bzw.} \quad (\mathbf{A} - \lambda \mathbf{I}) \mathbf{z} = \mathbf{0} \quad \text{oder} \quad |\mathbf{A} - \lambda \mathbf{I}| = 0$$

und damit für das System aus zwei Variablen

$$\begin{vmatrix} a-\lambda & b \\ c & d-\lambda \end{vmatrix} = (a-\lambda)(d-\lambda) - cb = \lambda^2 - \lambda(a+d) + (ad - cb) = 0 \quad (3.60)$$

Die Nullstellen des charakteristischen Polynoms sind die Eigenwerte. Mit den Abkürzungen $p := (a+d)$ und $q := (ad-cb)$ erhalten wir die Lösungen der quadratischen Gleichung

$$\lambda_{1,2} = \left(\frac{p}{2}\right) \pm \sqrt{\left(\frac{p}{2}\right)^2 - q} \quad (3.61)$$

Dabei können wir folgende Fälle unterscheiden:

- (1) Ist $p^2/4 - q < 0$, so ist der Eigenwert komplex mit nicht verschwindenden imaginärem Anteil; es ergeben sich zwei konjugiert komplexe Werte

$$\lambda_{1,2} = \left(\frac{p}{2}\right) \pm \sqrt{-1} \sqrt{q - \left(\frac{p}{2}\right)^2} = \left(\frac{p}{2}\right) \pm i \sqrt{q - \left(\frac{p}{2}\right)^2} = \alpha + i\beta \quad (3.62)$$

Damit sind bei $\beta > 0$ Oszillationen Lösung der Gleichung (3.57), was man als "Strudel" oder "Wirbel" im 2-dimensionalen Schaubild (u,v) deuten kann, siehe Abb. 3.25.

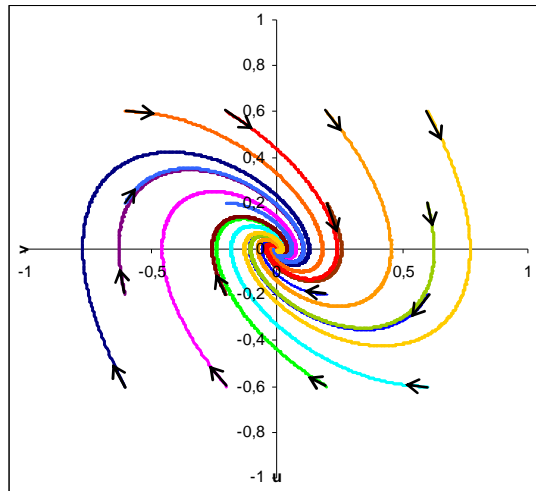
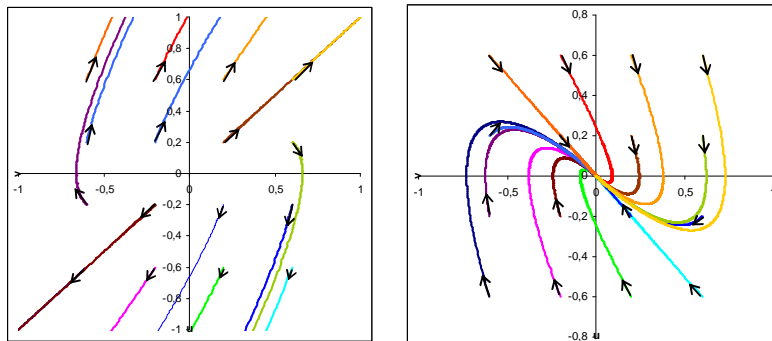


Abb. 3.25 2-dim Phasendiagramm eines linearen Systems mit $\alpha < 0$ und $\beta > 0$ ($a = 0, b = 1, c = -1, d = -1$)

- (2) Ist $p > 0$, so ist auch $\alpha = \text{Re}(\lambda_i) > 0$. Dies bedeutet eine positive Rückkopplung; das System wächst exponentiell an und ist damit instabil; unabhängig davon, ob Oszillationen vorliegen oder nicht, siehe Abb. 3.26(a).



(a) $\alpha > 0, \beta = 0$ ($c = -1, d = +2$) (b) $\alpha < 0, \beta = 0$ ($c = -1, d = -2$)

Abb. 3.26 2-dim Phasendiagramme eines linearen Systems mit $a = 0, b = 1$.

Erst bei $\alpha \leq 0$ erhalten wir ein langfristig stabiles System, siehe Abb. 3.26(b).

Nichtlineare Kopplungen

Bereits bei zwei Variablen können auch nichtlineare Wechselwirkungen existieren. Bekannte Beispiel dafür kommen aus der Populationsdynamik. Hierbei gibt es zwei Populationen unterschiedlicher Spezies, die jeweils ihre eigene Geburtenrate α und Sterberate β aufweisen. In diesem Fall erhalten wir für die Anzahl der Individuen N_1 bzw. N_2 analog zu Gl. (3.21) die Differentialgleichungen

$$\begin{aligned}\frac{dN_1}{dt} &= (\alpha_1 - \beta_1) N_1 \\ \frac{dN_2}{dt} &= (\alpha_2 - \beta_2) N_2\end{aligned}\quad (3.63)$$

Derartige Differentialgleichungen gelten nicht nur für Populationen, sondern auch für andere Systeme, etwa für Zu- und Abfluss von Wasser oder Warenmengen. In der obigen Form verhält sich jede Population analog zu Gl.(3.11): bei Überwiegen der Geburtenrate nimmt sie exponentiell zu, ansonsten schrumpft sie exponentiell zusammen. Interessant wird dieses System erst, wenn wir die Existenz beider Populationen von einander abhängig machen. Dazu betrachten wir zwei bekannte Fälle.

Beispiel Futterkonkurrenz

Angenommen, zu der natürlichen Sterblichkeit kommt noch Tod durch Verhungern hinzu. In diesem Fall enthält die Mortalitätsrate noch eine Funktion aller Individuen, etwa proportional zu der Summe beider Populationen

$$\beta_i = \gamma_i + c_i (N_1 + N_2) \quad i = 1, 2 \quad (3.64)$$

Da die Populationen mit der Zeit sich ändern, ändern sich auch die Sterberaten. Die vollständigen Differentialgleichungen sind dann

$$\begin{aligned}\frac{dN_1}{dt} &= (\alpha_1 - \gamma_1 - c_1 (N_1 + N_2)) N_1 = a_1 N_1 - c_1 N_1 (N_1 + N_2) \\ \frac{dN_2}{dt} &= (\alpha_2 - \gamma_2 - c_2 (N_1 + N_2)) N_2 = a_2 N_2 - c_2 N_2 (N_1 + N_2)\end{aligned}\quad (3.65)$$

mit $a_1 \equiv \alpha_1 - \gamma_1$, $a_2 \equiv \alpha_2 - \gamma_2$

Die zeitliche Entwicklung ist in Abb. 3.27 zu sehen.

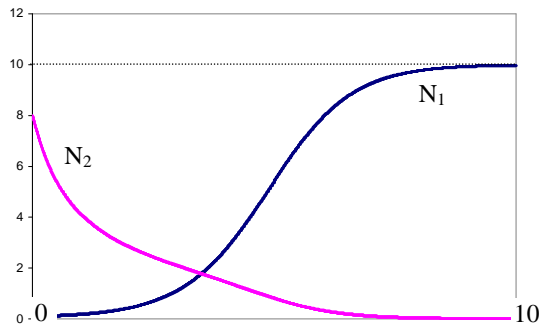


Abb. 3.27 Populationsentwicklung bei Futterkonkurrenz. Parameterwerte sind $N_1(0)=0,1, N_2(0)=8, a_1=1, c_1=0,1, a_2=0,1, c_2=0,1$

Wann ist das System stabil? Es lässt sich zeigen, dass (unabhängig von der Anfangszahl) Population 1 die Population vollständig bis zum Aussterben ($N_2 \rightarrow 0$) verdrängt, wenn $a_1/c_1 > a_2/c_2$ ist; im umgekehrten Fall wird Population 1 verdrängt.

Beispiel Räuber-Beute-System

Ein anderer wichtiger Fall tritt ein, wenn eine Population von der anderen lebt, etwa bei Füchsen und Hasen. Bei einem solchen „Räuber-Beute“-System, wie es von Volterra 1928 eingeführt wurde (und damit die Populationsdynamik begründete), ist die Sterberate von Spezies 1 geringer durch ein Nahrungsangebot, der Spezies 2, und die Sterberate von Spezies 2 höher durch das Gefressen werden durch Spezies 1:

$$\begin{aligned} \beta_1 &= \gamma_1 - c_1 N_2 && \text{Fressen von Spezies 2} \\ \beta_2 &= \gamma_2 + c_2 N_1 && \text{Gefressen werden von Spezies 1} \end{aligned} \tag{3.66}$$

Die vollständigen Differentialgleichungen sind dann

$$\begin{aligned} \frac{dN_1}{dt} &= (\alpha_1 - \gamma_1 + c_1 N_2) N_1 = a_1 N_1 + c_1 N_1 N_2 && \text{Räuber} \\ \frac{dN_2}{dt} &= (\alpha_2 - \gamma_2 - c_2 N_1) N_2 = a_2 N_2 - c_2 N_1 N_2 && \text{Beute} \end{aligned} \tag{3.67}$$

Dabei ist klar, dass $\alpha_1 - \gamma_1 < 0$ gelten muss, da ohne Beutetiere N_2 die Räuber aussterben, also eine negative Wachstumsrate haben müssen. Umgekehrt ist auch $\alpha_2 - \gamma_2 > 0$, da ohne Räuber N_1 die Beute sich ungehemmt vermehren können sollte.

Die zeitliche Entwicklung des Systems in Abb. 3.28 zeigt ein periodisches Verhalten: Wird die Zahl der Beutetiere durch die Räuber dezimiert, so nimmt mit einer Verzögerung auch die Zahl der Räuber ab. Wächst darauf hin die Anzahl der Beutetiere wieder an, so folgt dem mit Verzögerung auch die Räuberanzahl.

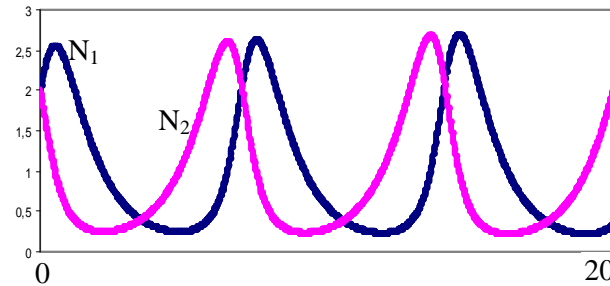


Abb. 3.28 Populationsentwicklung bei Räuber-Beute Abhängigkeit. Parameterwerte sind $N_1(0) = 2$, $N_2(0) = 2$, $a_1 = -1$, $c_1 = 1$, $a_2 = 1$, $c_2 = 1$

Im Unterschied zu dem Beispiel der Futterkonkurrenz kann die Räuberpopulation die Beutepopulation nicht verdrängen; sie hängt direkt von ihr ab.

Beispiel Bistabile Systeme

Bei der Modellierung können selbst kleine Ungenauigkeiten große Auswirkungen haben. Betrachten wir zum Beispiel das Wirkungssystem aus den Differentialgleichungen

$$\begin{aligned} u' &= b \cdot v & b > 0 \\ v' &= c \cdot u(1 - u^2) + d \cdot v & c > 0, d < 0 \end{aligned} \quad (3.68)$$

Die Differentialgleichung für v enthält fast ein logistisches Wachstum für u – bis auf den quadratischen Term, der stark nicht-linear eingeht und im Endeffekt einen kubischen Term in u bewirkt. Für $u^2 = 1$ wird der Kopplungseinfluss null, so dass sowohl bei $u = +1$ als auch bei $u = -1$ ein Fixpunkt vorliegt. In Abb. 3.29 (a) ist das Phasendiagramm des Zustandsraums gezeigt. Man erkennt zwei stabile Fixpunkte (Senken) und einen instabilen Sattelpunkt.

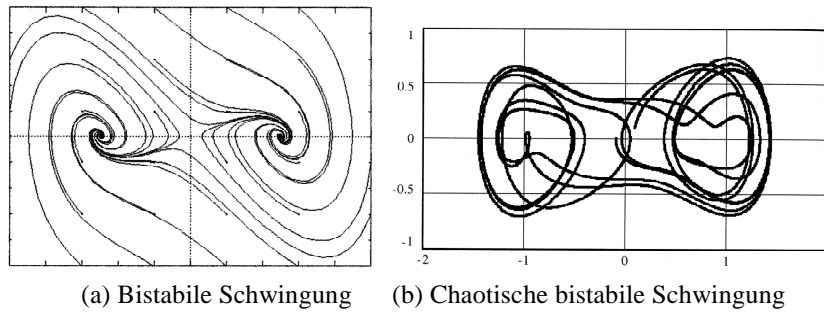


Abb. 3.29 Phasendiagramme bistabiler Schwinger

Geben wir noch zusätzlich einen regelmäßige Schwankung $f(t) = \cos(\omega t)$ auf die Kopplung beider Zustände

$$\begin{aligned}
 u' &= & b \cdot v & & b > 0 \\
 v' &= c \cdot u(1-u^2) + d \cdot v + r \cdot f(t) & c, r > 0, d < 0 & & (3.69)
 \end{aligned}$$

so erhalten wir ein System, das zwischen beiden Senken wechseln kann, siehe Abb. 3.29 (b). Dabei geschieht dies interessanterweise nicht periodisch, sondern sehr unregelmäßig und hängt von der numerischen Approximation ab. Schon kleine Änderungen verändern die Dynamik erheblich. Ein solches Verhalten entspricht einem (deterministischen) Chaos.

Aktivitäts-Hemmungs-Mechanismen: Muschelpigmentierung

Bei vielen Muscheln ist eine Pigmentierung der Schale zu beobachten. Dabei sind öfters Streifen oder Linien zu beobachten, die senkrecht oder waagrecht zur Wachstumsrichtung der Schale verlaufen. In Abb. 3.30 ist eine Muschel zu sehen, bei der die Streifen senkrecht zu dem spiralenförmigen Wachstum zu sehen sind.



Abb. 3.30 Pigmentierung bei der Muschel *Lyria planicostata taiwanica*

Wir können uns dies als eine Pigmentierungsinformation entlang einer Linie aus Wachstumszellen senkrecht zur Wachstumsrichtung vorstellen. Wie kann man sich den Pigmentierungsprozess vorstellen? Wie entwickelt sich ein derartiges Muster entlang einer Linie? Dazu erarbeitete Hans Meinhard einige interessante Modelle, die in dem lesenswerten Buch (Meinhard 2003) beschrieben sind.

Grundmodell seiner Betrachtungen ist ein System aus einem Aktivator u und einem Gegenspieler, dem Inhibitor v . Beide sind durch Differentialgleichungen miteinander verbunden, wobei der Aktivator für sich eine reine autokatalytische (sich selbst verstärkende) Funktion hat, etwa die Konzentration eines Moleküls, das verstärkt entsteht, wenn es vorhanden ist. Dies kann man modellieren durch eine Funktion der Form

$$u' = u(u - 1). \quad (3.70)$$

Im Unterschied zu dem logistischen Wachstum bedeutet hier jede Abweichung vom Punkt $u = 1$ eine Änderung, die bei $u > 1$ zu positivem Wachstum $u(t) \rightarrow \infty$ und bei $u < 1$ zu negativem Wachstum (Abnahme) und damit zu $u(t) \rightarrow 0$ führt: Der Punkt bei $u = 1$ ist nicht stabil.

Für eine Stabilität benötigt das System eine zweite Einflussgröße: den Inhibitor. Dieser kann beispielsweise vermehrt auftreten, wenn der Aktivator anwächst, und wird ansonsten abgebaut.

$$v' = u^2 - v. \quad (3.71)$$

Der Inhibitor ist mit $v' = 0$ stabil, wenn $u^2 = v$ ist.

Die Kopplung beider Systeme nehmen wir vor, indem wir die Inhibition umgekehrt proportional zur Aktivität einsetzen. Gl. (3.70) wird damit zu

$$u' = u(u/v - 1) = u^2/v - u \quad (3.72)$$

Diese Art der Beeinflussung hat einen großen Vorteil: Nehmen wir an, dass der Inhibitor ein Gleichgewicht erreicht nach einer Änderung im Aktivator, so ist mit $u^2 = v$ auch der Aktivator stabilisiert:

$$u' = u^2/u^2 - u = 1 - u$$

Ist $u = 1$, so verschwindet u' und u ist stabil. Ist aber $u > 1$, so folgt $u' < 0$ und somit eine Abnahme, bis $u = 1$ resultiert. Umgekehrt bei $u < 1$ folgt $u' > 0$ und damit ein Anwachsen von u , bis es wieder auf 1 stabil wird.

Wir erweitern nun die lokale Abhängigkeit zwischen Aktivator und Inhibitor zu einer Wechselwirkung mit weiter entfernt arbeitenden Zentren (Zellen). Da wir bei der Muschel eine Wachstumslinie beobachten, sind die wechselwirkenden Zellen entlang dieser Linie aufgereiht. Für eine Wechselwirkung modellieren wir deshalb einen Stoffaustausch der Inhibitoren bzw. Aktivatoren entlang einer eindimensionalen Richtung x zusätzlich zu der zeitlichen Modellierung. Dabei müssen wir bedenken, dass eine einfache, konstante Änderung der Konzentration zwar eine Materialdiffusion bewirkt, aber dabei genauso viele Moleküle in eine Zelle hinein gehen wie hinausgehen. In diesem Fall ändert sich die Konzentration der fraglichen Substanz nicht. Für eine echte Wechselwirkung über Zellgrenzen hinweg benötigen wir also eine nicht-konstante Änderung, beispielsweise eine räumliche Ableitung zweiten Grades $\partial^2 u/\partial x^2$. Mit diesen Überlegungen ergänzen wir unser System (3.70) und (3.72) zu

$$u' = b_u + s \cdot u^2/v - r_u u + c_u \frac{\partial^2 u}{\partial x^2} \quad \text{Aktivator-Inhibitor-} \quad (3.73)$$

$$v' = b_v + s \cdot u^2 - r_v v + c_v \frac{\partial^2 v}{\partial x^2} \quad \text{Mechanismus} \quad (3.74)$$

wobei wir die verschiedenen Einflüsse mit Faktoren s, r_u, r_v, c_u, c_v versehen haben. Der Wechselwirkungsmechanismus funktioniert dabei wie folgt: Ausgehend von kleinen, initialen Aktivitäten b_u und b_v und konstanten Abbauraten r_u und r_v nehmen wir an, dass eine einzige Zelle eine leicht erhöhte Konzentrationen von u und v hat, beispielsweise durch eine stochastische Fluktuation des Stoffwechsels. Dann bleibt die Konzentration u lokal dort länger erhalten, während sich der Inhibitor v durch eine deutlich erhöhte Diffusion (mind. 7 mal schneller) verbreitet. Dadurch kommt es zu einer erhöhten Autokatalyse von u in der Zelle, und die Aktivierung steigt. Da aber die Diffusion von u ebenfalls vorangeht, stabilisiert sich anschließend der neue Konzentrationswert in der Zelle. Dabei können die Nachbarzellen nicht das Gleiche tun, da der Inhibitor zuerst ankommt und dort die erhöhte Aktivierung verhindert. Es kommt also zu zufälligen Ungleichmäßigkeiten entlang der Wachstumslinie, die sich stabilisieren. In Abb. 3.31 ist dies durch eine Grafik verdeutlicht.

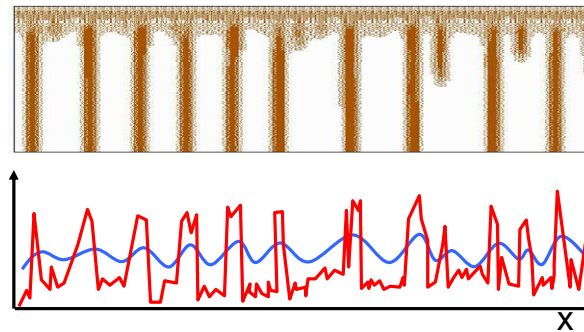


Abb. 3.31 Ausprägung der Aktivität entlang einer Wachstumslinie x .

In ihr sind sowohl die Aktivitätskonzentrationen als auch die Inhibitorkonzentration eingezeichnet. Typisch sind die lokalen, starken Konzentrationsgefälle des Aktivators, während der Inhibitor stark diffundiert weniger lokale Unterschiede aufweist.

Je nach dem, ob die Pigmentierung von u oder v bewirkt wird, bilden sich Streifen oder ihr farbliches Komplement (helle Streifen auf dunklem Grund) aus.

Alternative Modellierung1

Bei diesem Modell stellt sich die Frage: Ist das das einzige Modell, um eine autokatalytische Reaktion zu modellieren? Welche anderen Wechselwirkungen zwischen zwei Zuständen sind noch denkbar und sinnvoll?

Bei unserem Beispiel der Muschelpigmentierung können wir noch viele weitere Arten der Pigmentierung beobachten, beispielsweise breite, regelmäßig angeordnete Streifen im Unterschied zu den obigen unregelmäßigen Streifen. Dies lässt sich durch einen Sättigungsmechanismus modellieren, der die Inhibition ergänzt. Dazu wandeln wir den Inhibitionsterm $s \cdot u^2/v$ in Gl. (3.73) so ab, dass er bei größeren Stärken von u neutralisiert wird:

$$u^2 \text{ wird zu } \frac{u^2}{1+s_u u^2} \quad (3.75)$$

Bei kleinen Aktivator-Konzentrationen $1 \ll s_u u^2$ reduziert sich der Bruch zum ursprünglichen autokatalytischen Term; bei größeren Aktivitäten wird der Bruch konstant. Mit diesen Überlegungen wandeln wir die Gleichungen (3.73) und (3.74) um in folgende Gleichungen

$$u' = b_u + s \cdot u^{*2} v - r_u u + c_u \frac{\partial^2 u}{\partial x^2} \quad \text{Aktivator-Substrat-} \quad (3.76)$$

$$v' = b_v - s \cdot u^{*2} v - r_v v + c_v \frac{\partial^2 v}{\partial x^2} \quad \text{Mechanismus} \quad (3.77)$$

wobei der autokatalytische Term $s \cdot u^{*2} v$ nicht nur durch die Menge des Substrats v , sondern auch durch die Sättigung im Term u^{*2} bestimmt wird:

$$u^{*2} = \frac{u^2}{1 + s_u u^2} \quad (3.78)$$

Der autokatalytische Term produziert die Aktivatorkonzentration auf Kosten des Substrats, das sich verringert. Aus diesem Grund ist dieser Term in Gl. (3.77) subtrahiert. Die Formulierung als Aktivator-Substrat-Mechanismus erzeugt breite Maxima, die in regelmäßigen Abständen erzeugt werden und sich in Regionen höherer Substratkonzentration verschieben können.

Alternative Modellierung2

In der Formulierung von (Meinhard 2003) ist die geringe Initialproduktion b_u anstatt in Gl. (3.76) im Term von Gl. (3.78) additiv enthalten. Dadurch wird bei lokalem Ausdünnen des Substrats v die Gesamtaktivität geringer als im ersten Fall, bei dem die Initialproduktion b_u unabhängig von v ist. Wird die Diffusion zu einem frühen Entwicklungszeitpunkt gestoppt, so bleiben alle Zellen mit einer Aktivatorkonzentration aktiv, die oberhalb eines Schwellenwerts liegt; alle anderen Zellen werden vollständig inaktiv.

Ein weiterer Mechanismus besteht darin, anstelle einer Inhibition den hemmenden Einfluss über die Erhöhung der Aktivator-Abbaurrate r_u auszuüben. In diesem Fall erhalten wir die Differentialgleichungen

$$u' = b_u + s \cdot u^2 - r_u u v + c_u \frac{\partial^2 u}{\partial x^2} \quad \text{Aktivator-Abbau-} \quad (3.79)$$

$$v' = b_v + s \cdot u^2 - r_v v + c_v \frac{\partial^2 v}{\partial x^2} \quad \text{Mechanismus} \quad (3.80)$$

In diesem Fall ist der Abbau proportional zur Konzentration des Inhibitors, der wiederum zusammen mit dem Aktivator erzeugt wird. Hier fällt die Breite der Aktivatormaxima lokal schnell ab, da der Inhibitor dort auch stark erzeugt wird. Die geringere Lebensdauer des Aktivators kann zu Oszillationen führen, die dann periodische Muster zur Folge haben können. Biologisch gesehen ist der molekulare Aufbau-Abbaumechanismus nicht besonders energiesparend und damit unwahrscheinlicher als andere Mechanismen.

3.3.5 Systeme mit drei Zustandsvariablen

Je mehr Zustandsvariable wir einführen, um so komplexer werden die möglichen Systemreaktionen. Den Fall linearer Systeme haben wir bereits beim Fall zweier Zustandsvariablen behandelt; hier kommen keine neue Erkenntnisse hinzu.

Aus den vielfältigen Systemen mit drei Variablen möchte ich einen Spezialfall herausnehmen: Im Fall der Wechselwirkung von zwei Zustandsvariablen kann eine dritte Variable als „Vermittler“ wirken. Beispielsweise können wir in unserem Beispiel der Muschelpigmentierung eine dritte Substanz w einführen, die den Aktivator u ebenfalls hemmt und deren Produktion von dem schnell diffundierenden Inhibitor v kontrolliert wird:

$$u' = b_u + \frac{s}{s_u + w^2} - r_u u + c_u \frac{\partial^2 u}{\partial x^2} \quad \text{Aktivator-Inhibitor-} \quad (3.81)$$

$$v' = r_v u - r_v v + c_v \frac{\partial^2 v}{\partial x^2} \quad \text{Inhibitions-Mechanismus} \quad (3.82)$$

$$w' = \frac{s}{s_w + u^2 / v^2} - r_w w + c_w \frac{\partial^2 w}{\partial x^2} \quad (3.83)$$

Ist der Aktivator u maximal, so wird die Produktion des Inhibitors w stark gehemmt und umgekehrt. Zusammen bilden sie ein bistabiles System. Die Produktion von w wird wiederum durch v gefördert, so dass v ebenfalls als Inhibitor für u agiert. Ist die Pigmentierung anstatt an u an w gebunden, so entsprechen den schmalen Pigmentstreifen und breiten Zwischenräumen bei u den breiten Streifen und schmalen Zwischenräumen bei w : Die Streifenbildung wirkt komplementär.

3.4 Gleichgewichtssysteme

In Systemen können wir Vorgänge beobachten, die für sich allein schnell ablaufen und einem stabilen Gleichgewicht zustreben. Betrachten wir beispielsweise die Regelung einer Raumtemperatur durch ein Thermostat, so wird bei Absinken der Temperatur unter die voreingestellten Temperaturvorgabe das Thermostat sich einschalten und die Heizung betätigen. Dies wird solange durchgeführt, bis die ansteigende Raumtemperatur das Thermostat wieder ausschaltet. Wir empfinden dabei die Temperatur im 24 Stunden-Zeitmaßstab als „konstant“ und ignorieren dabei die Tatsache, dass in Wirklichkeit die Temperatur um unsere gewünschte Raumtemperatur im 15-Minuten-Zeitmaßstab schwankt. Wir reduzieren also die Dynamik des Systems auf einen konstanten Wert, die Vorgabetemperatur, und kapseln damit das rückgekoppelte System in einem Subsystem ab, von dem wir nur das Ergebnis im großen Zeitmaßstab betrachten. Dies modelliert zwar die Wirklichkeit nicht exakt, reicht aber für unsere Annahme über die Temperatur und die Weiterverwendung in einem komplexeren Modell (z.B. Energiebilanz des Gesamtgebäudes) vollkommen aus.

Diese Vorgehensweise können wir verallgemeinern: Haben wir sehr komplexe Systeme, so können wir versuchen, die Komplexität dadurch zu verringern, dass wir rückgekoppelte Prozesse, die sich sehr schnell stabilisieren und zu einem Gleichgewicht konvergieren, in Subsystemen zusammenfassen und die Subsysteme auf den Gleichgewichtszustand reduzieren.

Beispiel Badewanne

In unserem Beispiel von der Badewanne, die gleichzeitig Wasserzufluss und Wasserabfluss hat, gingen wir von einem Zufluss- und einem Abflussmechanismus aus, die sich überlagern und zusammen nach einer gewissen Zeit, charakterisiert durch eine Zeitkonstante T , ein Gleichgewicht erreichen. Die Überlagerung ist nach Gl. (3.45)

$$\frac{dV}{dt} = r_z - r_a(t) \quad \text{mit} \quad r_a(t) = V(t)/T$$

Das Volumen stabilisiert sich beim Wert $V(\infty) = A_0 = r_z T$ und damit bei einer Abflussrate von

$$r_a(\infty) = V(\infty)/T = r_z \quad (3.84)$$

Gleichgewicht herrscht also, wenn Zufluss und Abfluss genau gleichen Wert haben. Dies ist bei der Kreuzung beider von einander unbeeinflusster Kurven in Abb. 3.32 der Fall.

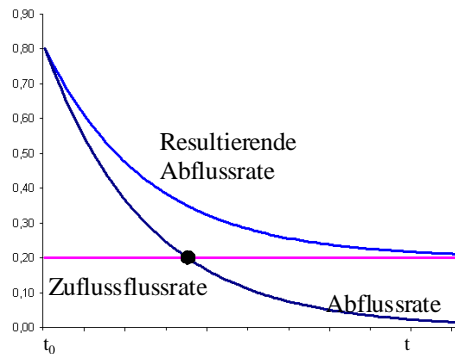


Abb. 3.32 Die Zufluss, Abfluss und Gleichgewicht beim Speicher

Da aber hier die Abflussrate vom Gesamtzustand abhängig ist, geht sie mit einer gewissen Zeitverzögerung zu dem Gleichgewichtszustand über.

Wir sahen in diesem Beispiel, dass sich das langsame Verhalten der Abflussrate durch Angabe des Gleichgewichtspunktes abkürzen lässt. Dieser ergibt sich aus dem Schnittpunkt der gegenläufigen Funktionen.

Der Vorteil einer solchen Vorgehensweise ist die Reduzierung der Dynamik auf den Zustand im Gleichgewicht und damit eine drastische Vereinfachung des Systems. Der Nachteil besteht darin, dass die Vorhersage ungenau wird, wenn wir einen Zeitpunkt innerhalb des durch die Zeitkonstante T stark beeinflussten Bereichs betrachten. Hier gilt eben die Gleichgewichtsbedingung noch nicht; das Volumen ändert sich noch laufend.

Die Reduzierung auf ein Gleichgewicht als vereinfachende Betrachtungsweise ist in verschiedenen Bereichen üblich, beispielweise in der Makroökonomie.

Beispiel Warenmengen-Preisdynamik

Betrachten wir die Abhängigkeit der Produktion einer Ware als Funktion des Preises, so ergibt sich qualitativ eine ansteigende Kurve $y_p(x)$ wie in Abb. 3.33. Es werden mehr Waren einer Sorte produziert, wenn der Preis, der erzielt werden kann, höher ist. Andererseits kaufen weniger Konsumenten eine Ware, wenn sie teuer ist, siehe Kurve $y_k(x)$. Wie viel sollte von der Ware produziert werden? Nun, nach einer gewissen Zeit, in der die Käufernachfrage die Regale leergefegt hat oder in der die Ladenhüter aussortiert wurden, stellt sich ein Gleichgewicht zwischen Angebot und Nachfrage ein. Dieses Gleichgewicht ist am Schnittpunkt beider Kurven zu beobachten. Die gesamte Makroökonomie beruht auf der Annahme, dass man nur die „großen“ Zeiträume modellieren muss, nicht die kurzzeitige Dynamik. Es reicht also, nur den Gleichgewichtspunkt zu betrachten und als Konstante in der Gesamtdynamik anzunehmen.

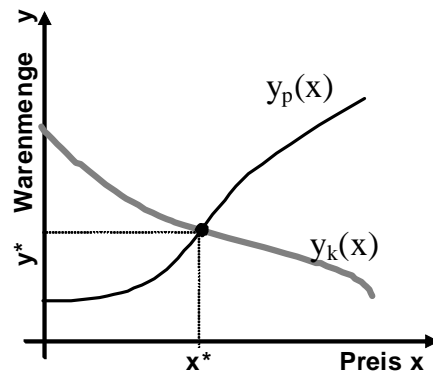


Abb. 3.33 Der Gleichgewichtszustand als Folge gegenläufiger Mechanismen

In Wirklichkeit müsste man ein solches System als ein rückgekoppeltes System aus zwei Variablen modellieren. Dazu nehmen wir der Einfachheit halber lineare Zusammenhänge an und modellieren die Warenproduktion y_p und den Warenkonsum y_k als Geraden

$$\begin{aligned} y_p(x) &= y_p(0) + ax \\ y_k(x) &= y_k(0) - bx \end{aligned} \quad (3.85)$$

Der Gleichgewichtspunkt ist durch die Bedingung

$$y_p(x^*) = y_k(x^*) \quad (3.86)$$

gegeben und ergibt

$$\begin{aligned} y_p(0) + ax^* &= y_k(0) - bx^* \Leftrightarrow (a+b)x^* = y_k(0) - y_p(0) \\ \Leftrightarrow x^* &= (y_k(0) - y_p(0))/(a+b) \\ \Leftrightarrow y^* &= y_p(0) + ax^* \end{aligned} \quad (3.87)$$

Dies ist die statische Gleichgewichtslösung. Die Dynamik modellieren wir aber anders. Die Kurven in Abb. 3.33 legen nahe, dass bei einer Preiserhöhung die Nachfrage sinkt und das Angebot steigt. Dies begründet sich in dem "Kaufpotential", der Differenz $y_k - y_p$, das trotz Preiserhöhung Δx noch vorhanden sein wird und Anreiz zur Produktionssteigerung bleibt. Also ist die Produktionssteigerung Δy_p pro Preisänderung proportional zum Kaufpotential

$$\Delta y_p / \Delta x \sim (y_k - y_p) \Leftrightarrow \Delta y_p / \Delta x = c_1 (y_k - y_p) \quad c_1 > 0 \quad (3.88)$$

während das absolute Konsumpotential y_k dadurch schwindet

$$\Delta y_k / \Delta x = -c_2 (y_k - y_p) \quad c_2 > 0 \quad (3.89)$$

Im Grenzübergang $\Delta x \rightarrow 0$ erhalten wir die Differentialgleichungen

$$\begin{aligned} y_p' &= c_1 (y_k - y_p) = -c_1 y_p + c_1 y_k \\ y_k' &= -c_2 (y_k - y_p) = +c_2 y_p - c_2 y_k \end{aligned} \quad (3.90)$$

Wie ist das Systemverhalten dieses gekoppelten Systems? Die Gleichungen (3.90) entsprechen einem linearen Differentialgleichungssystem (3.56) mit der Lösung (3.60). Für unser System ist somit

$$\lambda^2 + \lambda(c_1 + c_2) + (c_1 c_2 - c_1 c_2) = \lambda^2 + \lambda(c_1 + c_2) = 0$$

oder für $\lambda \neq 0$

$$\lambda = -(c_1 + c_2) < 0 \quad c_1, c_2 > 0 \quad (3.91)$$

Damit ist klar: Das System oszilliert nicht und hat einen exponentiellen Abfall des Wachstums. Die Startwerte beschränken sich auf positive Koeffizienten c_1, c_2 . In Abb. 3.34 (a) ist die Warenentwicklung bei Preissteigerung zu sehen.

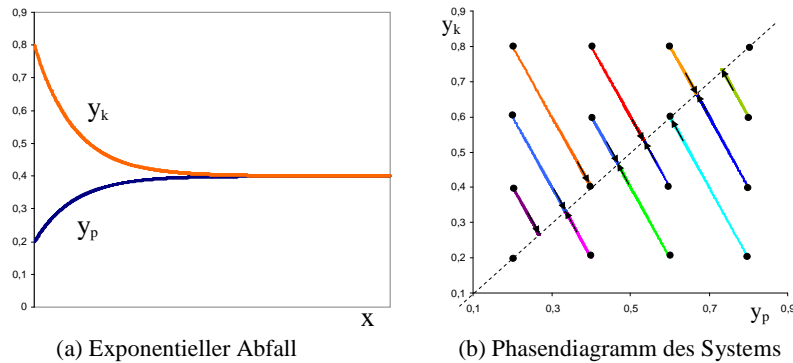


Abb. 3.34 Verhalten des Systems bei Preisanstieg bei $c_1 = 1, c_2 = 2$

Das Gesamtverhalten bei verschiedenen Startpunkten ist in Abb. 3.34 (b) gezeigt. Ausgehend von einem der 16 dargestellten Startwerte konvergieren die Zustände erwartungsgemäß zu einer Geraden, der Winkelhalbierenden. Dies entspricht dem Gleichgewicht mit $y_k = y_p = y^*$. Die Trajektorien sind Geraden, da mit Gl. (3.90) die Beziehung

$$\frac{y_k'}{c_2} = -\frac{y_p'}{c_1} \quad \text{und damit} \quad \frac{\Delta y_k}{\Delta x c_2} = -\frac{\Delta y_p}{\Delta x c_1} \quad \text{bzw.} \quad \frac{\Delta y_k}{\Delta y_p} = -\frac{c_2}{c_1} \quad (3.92)$$

resultiert. Die Gerade hat die negative Steigung c_2/c_1 , in Abb. 3.34 (b) die Steigung -2 .

Durch direktes Einsetzen der Lösungen der Gleichgewichtsbedingung in das System ersparen wir uns die Modellierung der Dynamik und kürzen den Simulationsprozess ab. Bei größeren Systemen ist dies durchaus von Nutzen und in Wirtschaftssimulationen durchaus üblich.

Beispiel Gleichgewichtsmodell für Steuerpolitik

In Abb. 3.35 ist als Beispiel die Grundstruktur der Modellierung eines einfachen Gleichgewichtsmodells zur Analyse von Steuerpolitiken gezeigt.

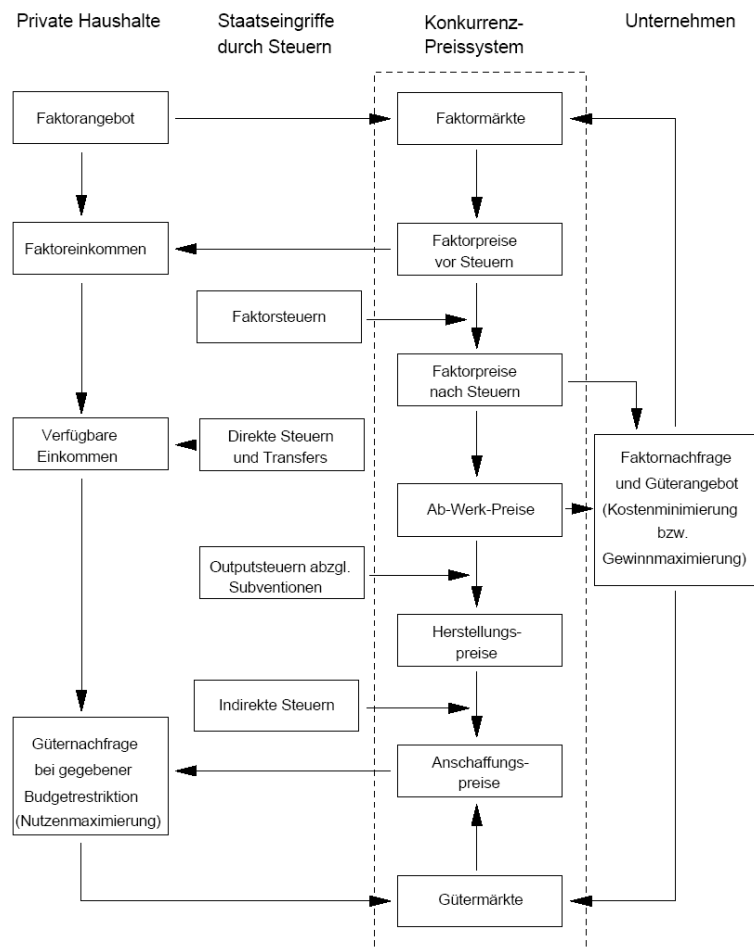


Abb. 3.35 Beispiel eines Gleichgewichtsmodells (aus [IAB253])

Zur Vereinfachung wird auf Produktionsverflechtungen und Außenhandelsbeziehungen verzichtet. Man nimmt dabei an, dass alle Handelnden rationale, wirtschaftlich optimale Entscheidungen treffen und der dabei entstehende Preismechanismus für Angebot und Nachfrage entscheidend ist.

In der Abbildung ist im Wesentlichen die Wirkstruktur von Steueränderungen gezeigt; die dahinter liegende Dynamik von Angebot und Nachfrage ist als statisches Gleichgewicht in den jeweiligen Kästen (Gütermärkte, Anschaffungspreise usw.) vorausgesetzt. Schreitet der Staat mit Steuereingriffen ein, so stellt sich jeweils ein neues Gleichgewicht darin ein, das man im Vergleich mit dem vorherigen Gleichgewicht quantitativ und qualitativ beurteilen kann.

Voraussetzung einer quantitativ korrekten Analyse ist nicht nur die korrekte Modellierung, sondern auch die Anpassung der Parameter des Modells an die Wirklichkeit. Dazu gibt es verschiedene Verfahren.

3.5 Schätzung der Parameter

Ähnlich wie bei den linearen Systemen müssen wir auch die Parameter θ der aufgestellten und verwendeten Modelle aus den Beobachtungen erschließen. Eine einfache Methode dafür besteht darin, die Parameter $\theta = (\theta_1, \dots, \theta_n)$ iterativ so lange zu optimieren, bis der Fehler $\varepsilon = y(t) - \hat{y}(t|\theta)$ im quadratischen Mittel

$$R_N(\theta) = \frac{1}{N} \sum_{t=1}^N \varepsilon_{t\theta}^2 \quad (3.93)$$

minimal wird

$$\theta_N^* = \arg \min_{\theta} R_N(\theta) . \quad (3.94)$$

Damit haben wir uns folgende Strategie gewählt: Suche diejenige Werte der Parameter θ , die den kleinste Abweichung der Ausgabe von der beobachteten Ausgabe bewirken. Allgemein bedeutet dies die Optimierung einer Zielfunktion

$$R_N(\theta) = \frac{1}{N} \sum_{t=1}^N f(\varepsilon_{t\theta}) = \langle f(\varepsilon_{t\theta}) \rangle_t \quad (3.95)$$

bei der die Mittelung (Erwartungswert) einer Funktion f errechnet wird, die vom Fehler $\varepsilon_{t\theta}$ abhängt.

Wählen wir zum Beispiel als Funktion den negativen Logarithmus der Wahrscheinlichkeitsdichte

$$f(\varepsilon_{t\theta}) = - \ln p(\varepsilon_{t\theta}) \quad \text{maximum likelihood-Schätzung} \quad (3.96)$$

so ist $f(\cdot)$ genau dann minimal, wenn die *likelihood*-Schätzung $p(\cdot)$, und damit auch $\ln p(\cdot)$, maximal ist. Bei minimalem $f(\cdot)$ wird das Maximum der Fehlerdichte $p(\cdot)$ gesucht. Dies bedeutet, als Zielfunktion die Entropie des Fehlers

$$R_N(\boldsymbol{\theta}) = \langle f(\boldsymbol{\varepsilon}_{t\theta}) \rangle_t = \langle -\ln p(\boldsymbol{\varepsilon}_{t\theta}) \rangle_t = H(\boldsymbol{\varepsilon}). \quad (3.97)$$

zu minimieren. Für den speziellen Fall, dass der Fehler normalverteilt ist mit der Gauß-Funktion

$$p(\boldsymbol{\varepsilon}_{t\theta}) = A e^{-\frac{\boldsymbol{\varepsilon}_{t\theta}^2}{2\sigma^2}} \quad (3.98)$$

wird die Entropie zu

$$H(\boldsymbol{\varepsilon}) = \langle -\ln p(\boldsymbol{\varepsilon}_{t\theta}) \rangle_t = \langle -\ln A e^{-\frac{\boldsymbol{\varepsilon}_{t\theta}^2}{2\sigma^2}} \rangle_t = \langle \boldsymbol{\varepsilon}_{t\theta}^2 / 2\sigma^2 \rangle_t - \ln A \quad (3.99)$$

Die Maximierung der *likelihood*-Wahrscheinlichkeit bzw. Minimierung der Entropie bedeutet in diesem Fall eine Minimierung des erwarteten quadratischen Fehlers.

Wir haben nun eine Zielfunktion gefunden und die Anpassung der Parameter als Minimierung der Zielfunktion konzipiert. Wie gelangen wir nun konkret zu unserem gewünschten Parametersatz? Eine bewährte Strategie besteht darin, nach jeder Beobachtung die Parameter zu verbessern und damit die Ausgabe des Modells den Beobachtungen der Realität anzupassen. Dazu können wir zwar den Gradientenalgorithmus aus Abschnitt 2.6 verwenden, aber wir wollen stattdessen eine andere Lernmethode einführen: die Newton-Raphson-Methode.

Bei dieser Methode wird die Nullstelle einer Funktion durch die Ableitung und die Funktion selbst bestimmt. Angenommen, wir kennen den Funktionswert $f(x_1)$ an der Stelle x_1 und bilden dort die Ableitung $f'(x_1)$, also die Tangente. Diese schneidet die x -Achse an der Stelle x_2 , siehe Abb. 3.36.

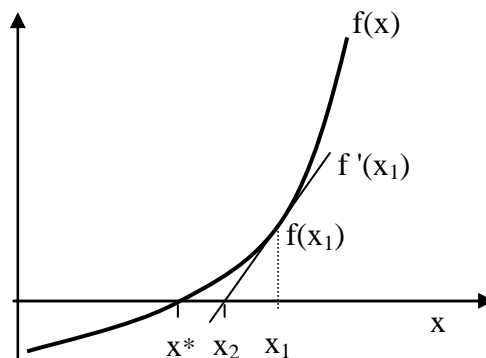


Abb. 3.36 Die Newton-Raphson Approximation

Also ist die Ableitung f' an der Stelle x_1 identisch mit der Steigung der Tangente

$$f'(x_1) = \frac{f(x_1) - 0}{x_1 - x_2} \Leftrightarrow x_1 - x_2 = f'(x_1)^{-1} f(x_1)$$

$$x_2 = x_1 - f'(x_1)^{-1} f(x_1)$$

Die Schrittweite wird nun mit Hilfe eines Parameters γ so angepasst, dass der Iterationsschritt vom t -ten zum $t+1$ -ten Schritt nicht zu groß und nicht zu klein wird

$$x_{t+1} = x_t - \gamma f'(x_t)^{-1} f(x_t) \quad (\text{Newton-Raphson-Algorithmus}) \quad (3.100)$$

Wie können wir dies für unsere Suche nach dem Minimum der Funktion $R(\theta)$ nutzen? Wir wissen, dass an der Stelle θ^* nicht nur $R(\theta)$ minimal ist, sondern auch die Ableitung $f(\theta) = R'(\theta)$ null wird. Dies können wir nutzen und für die Funktion $f(x)$ in Gl.(3.100) die Ableitung der Zielfunktion $R(\theta)$ einsetzen. Wir erhalten so den Iterationsalgorithmus für die Parameter θ

$$\theta_{t+1} = \theta_t - \gamma R''(\theta_t)^{-1} R'(\theta_t) \quad (3.101)$$

Bei einem n -dimensionalen Vektor θ ist $R'(\theta_t) = \text{grad } R(\theta)$ ein n -dimensionaler Vektor und $R''(\theta_t) = (\partial^2 R / \partial \theta_i \partial \theta_j)$ eine $n \times n$ Matrix, die Hesse-Matrix.

Verfügen wir nicht über die analytische Form der Zielfunktion, so können wir trotzdem das Minimum bestimmen. Dazu approximieren wir die Ableitung durch

$$R'(\theta_t) \approx \frac{R(\theta_{t-1}) - R(\theta_t)}{\theta_{t-1} - \theta_t} \quad (3.102)$$

und die zweite Ableitung durch

$$R''(\theta_t) \approx \frac{R'(\theta_{t-1}) - R'(\theta_t)}{\theta_{t-1} - \theta_t} \quad (3.103)$$

und setzen dies in Gl.(3.101) ein. Wir erhalten

$$\theta_{t+1} = \theta_t - \gamma \frac{R(\theta_{t-1}) - R(\theta_t)}{R'(\theta_{t-1}) - R'(\theta_t)} \quad (3.104)$$

Wir benötigen also in diesem Fall nicht nur einen Messwert der Zielfunktion, sondern sogar drei, da $R'(\theta_{t-1})$ auch von $R(\theta_{t-1})$ und $R(\theta_{t-2})$ bestimmt wird.

Betrachten wir nun ein Beispiel für die Optimierungsmethode zur Parameterschätzung.

Beispiel

Eine Papierherstellungsmaschine wird durch ein Modell mit den Zustandsgleichungen für zwei Zustände z_1 und z_2 , den Eingaben x_1 für den Papierbreifluss und x_2 für den Luftstrom, sowie dem Parameter c

$$\begin{aligned}\frac{dz_1}{dt} &= x_1 - c g_1(z_1, z_2) \\ \frac{dz_2}{dt} &= x_2 - g_2(z_1, z_2)\end{aligned}\quad (3.105)$$

sowie der Ausgabe

$$\hat{y}(t|c) = c g_1(z_1, z_2) \quad (3.106)$$

mit den bekannten Funktionen $g_1(z_1, z_2)$ und $g_2(z_1, z_2)$ beschrieben (Ljung, Glad 1994). Die Zielfunktion ist

$$R(c) = \frac{1}{100} \sum_{k=1}^{100} (\hat{y}(t|c) - y(k\Delta t))^2 \quad (3.107)$$

Sei die Eingabe $x_1(k\Delta t)$ zufällig gewählt und die dazu gehörende Ausgabe $y(k\Delta t)$ zu den diskreten Zeitpunkten $k\Delta t$ mit den obigen Gleichungen erzeugt, siehe Abb. 3.37. Verwendet wurde dabei ein Parameterwert von $c = 0,8$.

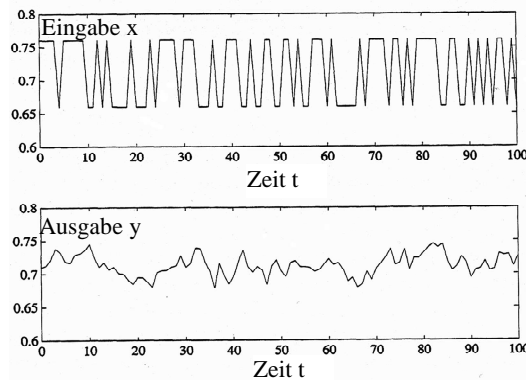


Abb. 3.37 Eingabe und Ausgabe der Papierherstellung (nach Ljung 1994)

Angenommen, wir verfügen nur über die obigen Testdaten. Können wir mit Hilfe der Zielfunktion den ursprünglichen Parameter c korrekt schätzen? Dazu zeichnen wir uns den Verlauf der Zielfunktion (Abb. 3.38) in Abhängigkeit vom Parameter c .

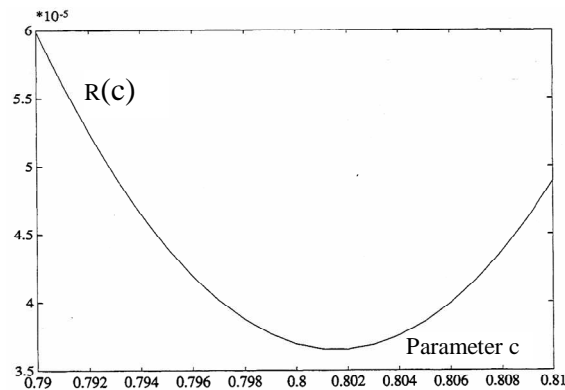


Abb. 3.38 Verlauf der Zielfunktion in Abhängigkeit des Parameters c

Man sieht deutlich, dass das Minimum bei ca. $c^* = 0,802$ angenommen wird und in guter Übereinstimmung mit dem tatsächlich eingesetzten Wert von $c = 0,8$ ist.

In dem vorliegenden Beispiel ließ sich der freie Parameter einfach bestimmen. Anders ist die Situation, wenn nicht nur ein Parameter, sondern eine Vielzahl davon zu bestimmen ist. Hier sind die Erfolge nur begrenzt, siehe etwa (Francois 2001). Eine wichtige Vorgehensweise besteht deshalb darin, das Gesamtmodell in Untermodellen zu spezifizieren und dann die Parameter für jedes Untermodell separat zu schätzen.

Bei der angewandten Gleichgewichtsmodellierung hat sich dabei ein anderes Vorgehen durchgesetzt. Dazu werden die Parameter für einen Referenzzeitraum, etwa ein Jahr, genau so angepasst, dass Modellaussagen und gemessene wirtschaftliche Daten wie Input-Output-Tabellen, Einkommens- und Verbrauchsstichproben, Steuerstatistiken usw. übereinstimmen; das Modell wird „kalibriert“. Man unterstellt dabei, dass die Wirtschaft während dieses Zeitraums im Gleichgewicht war. Konkret werden dabei die Gleichgewichtsbedingungen nach den Parametern aufgelöst und die gemessenen Werte eingesetzt.

Dieses Vorgehen bringt neben Vorteilen auch Probleme mit sich: Da der dynamische Anteil des Modells ausgeblendet wird, müssen wichtige dynamische Anteile wie „Flexibilitäten“ oder „Elastizitäten“ bei Preisen und Mengen durch zusätzliche Parameter bzw. Einflussgrößen abgeschätzt werden. Da ein statisches Gleichgewicht auch die Technologie- und Präferenzstruktur nicht explizit erfasst, werden systematische Sensitivitätsanalysen angewendet, um die Wirkung alternativer Annahmen (*Einflussshypothesen*) oder Änderungen (*Elastizität*) der wesentlichen Parameter zu erfassen und auf wirtschaftliche Konsistenz zu überprüfen. Die damit erzielte Robustheit des Modells ist eine wichtige Voraussetzung, um aus den Ergebnissen der Modellierung schließlich politische Handlungsempfehlungen

abzuleiten. In Abb. 3.39 ist ein Überblick über die fünf prinzipiellen Arbeitsschritte einer solchen Gleichgewichtsmodellierung und –analyse gezeigt.

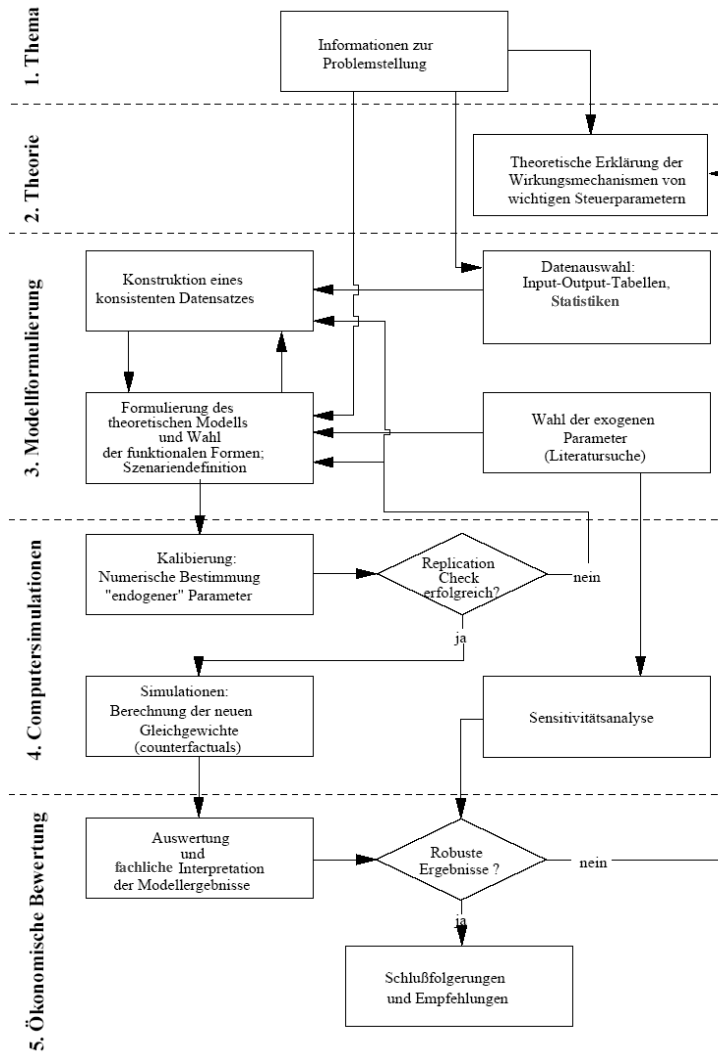


Abb. 3.39 Arbeitsschritte einer Gleichgewichtsanalyse (aus [IAB253])

Die Schritte bestehen aus folgenden Aktionen:

1. Beschaffen aller Informationen
2. Bildung einer Theorie, wie die Wirkungsmechanismen ablaufen könnten

3. Bildung einer einschlägigen Datenbank mit allen benötigten Tabellen, Statistiken und Eingabewerten, um die Submechanismen zu modellieren. Dazu gehört auch eine Modellierung der Mechanismen, die von den Daten gestützt wird.
4. Steht die eigentliche Simulation, so müssen ihre Parameterwerte an die Realität angepasst werden. Dies ist die Aufgabe des „*Replication check*“. Ist auch dieser erfolgreich, so kann man die Robustheit der Parameterkonstellation erforschen, um einen Eindruck von den zu erwartenden Ergebnissen zu erhalten.

Nun folgt die eigentliche Simulation, die die Fragen „was wäre, wenn...“ beantworten soll und verschiedene Szenarien durchspielt.

5. Abschließend müssen wir die Ergebnisse der Simulation bewerten: Sind sie eindeutig, oder müssen weitere Simulationsläufe erfolgen? Waren die Robustheitstests erfolgreich, oder können wir den Ergebnissen nicht trauen und müssen sie als Ausreißer interpretieren?

Je nach Bewertung können wir uns mit den Ergebnissen zufrieden geben und sie dem breiten Publikum bekannt geben, oder aber die Datenbasis bzw. die Modellierung verändern.

4 Hierarchische Systeme

Große, komplexe Systeme lassen sich sowohl von der Modellierung als auch von der Analyse der Ergebnisse schwer handhaben. Es hat sich deshalb bewährt, je nach Bedarf und Fragestellung sie in einzelne Subsysteme (Modell-Module) zu unterteilen, die jeweils einzeln verbessert, aktualisiert oder ausgetauscht werden können. Auch eine Zusammenarbeit mit anderen Institutionen, die sich auf Teilaspekte (Module) konzentrieren, wird damit vereinfacht.

Jedes der einzelnen Subsysteme muss in einem derartigen Gesamtsystem extra modelliert werden, wobei verschiedene Modelle zur Auswahl stehen: Input-Output-Tabellen, Lineare Modelle, Nicht-Lineare Modelle, Dynamische Systeme oder Gleichgewichtssysteme. Es ist durchaus möglich, wissensbasiert Regeln und Gleichungen für die Wechselwirkung zwischen Subsystemen anzugeben, ohne dabei die inneren Mechanismen für das beobachtete Verhalten angeben zu können.

In diesem Kapitel werden wir anhand einiger Beispiele verschiedene Arten kennen lernen, Modelle zu konstruieren. Betrachten wir zunächst den typischen Fall von historisch gewachsenen Modellhierarchien.

4.1 Hierarchiebildung durch Aggregation

Eine typische Art, hierarchische Modelle zu bauen, ist die Zusammenfassung von bereits fertigen Teilmodellen zu einem Gesamtmodell. In Kapitel 3.2 entwickelten wir ein hierarchisches System, das als Weltmodell in verschiedene, inhaltliche Teilmodelle gegliedert war. Das Gesamtmodell funktionierte also nur mit allen Teilen. Im Unterschied dazu sind historisch gewachsene Modelle darauf angewiesen, im Detaillierungsgrad mit der Zeit zu wachsen. Dies bedeutet die Notwendigkeit, die globalen Aussagen des Globalmodells mit den lokalen Aussagen der Submodelle abzugleichen.

Ein typisches Beispiel dafür sind Wirtschaftsmodelle. Im Bereich der Makroökonomie gewann in den letzten Jahrzehnten eine Initiative Raum, die nicht nur die Wirtschaft lokal modellieren möchte, sondern auch die globalen Einflüsse abbildet. Ausgehend von einer ursprünglich lokalen Initiative in USA bildete sich eine Initiativgruppe INFORUM, die ein globales Modellsystem aufbaut. Das Weltmodell GLODYM bildet die bilateralen Handelsbeziehungen (Wirtschaftsverflech-

tung, Geldbilanz, Energiebilanz) von 53 Ländern und Regionen ab und entstand auf Basis einer europäisch-asiatischen Zusammenarbeit. Oberste Einheit sind Wirtschaftsregionen, die wiederum in sektorale Gebiete (Länder usw.) unterteilt sein können. Das hierarchische Gesamtmodell ist in Abb. 4.1 abgebildet. Dabei sind 18 Länder besonders detailliert modelliert.

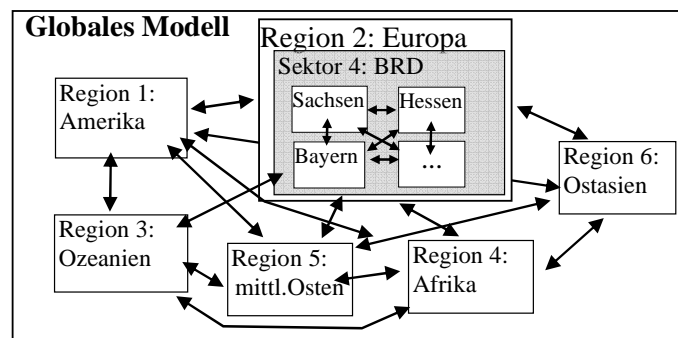


Abb. 4.1 Der hierarchische Aufbau eines globalen Wirtschaftssimulationsmodells

Die deutsche Version der Verflechtungsmodellierung im INFORUM-Verbund ist INFORGE (*Interindustry Forecasting Germany*, s. Meyer 1999), unterteilt sich in einzelne Bundesländer (Modell LÄNDER) und beschreibt die regionalen Wirkungen. Alternativ dazu gibt es eine Modellierung für Länderübergreifende Regionen und Städte (REGIO). Eine besondere Version davon wurde vom Forschungsinstitut IAB der Agentur für Arbeit erstellt und modelliert nur die in INFORGE wenig genau ausgeprägten Arten von Arbeit, etwa Wochenendarbeit, Saisonarbeit, Arbeitslosigkeit, Minijobs usw.

Obwohl das INFORGE-Modell sehr stark in sich geschlossen ist, also nur wenige externe (exogene) Vorgaben und viele interne (endogene) Variablen hat, lassen sich erst durch eine internationale Zusammenarbeit die wirtschaftlichen Effekte der exportabhängigen deutschen Wirtschaft erfolgreich modellieren.

In Abb. 4.2 ist die INFORGE-Modellstruktur gezeigt. Bei einer solchen Hierarchisierung werden die verschiedenartigen Einflüsse baumartig in Untergruppen zusammengefasst.

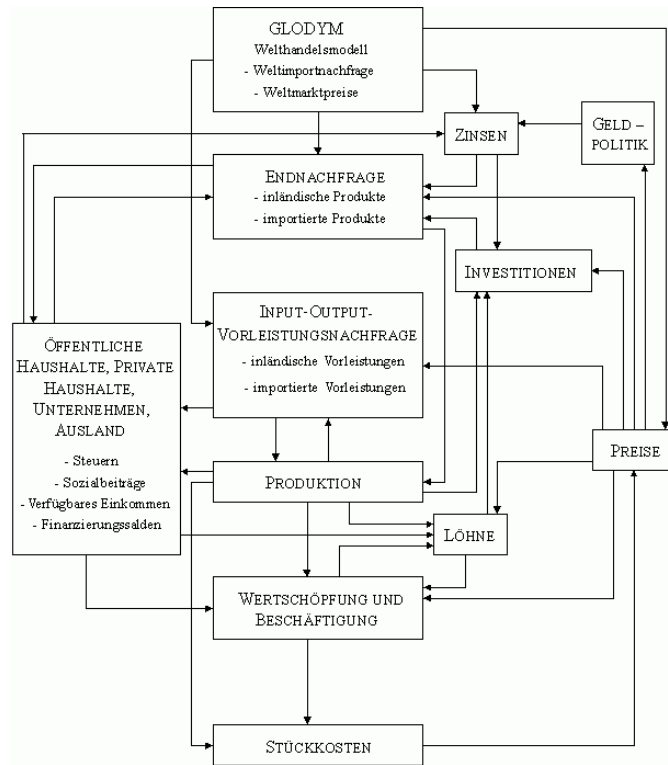


Abb. 4.2 Das INFORGE-Modell (aus Schnur 2002)

In Abb. 4.3 ist dies am Beispiel des INFORGE-Modells gezeigt. Dabei ist jede Untergruppe wiederum von durchschnittlich 250 Variablen bestimmt. Auch das Konsumverhalten wird modularisiert: jede der 43 Konsumverwendungszwecke kann wiederum in die 59 Produktbereiche unterteilt werden, für die er sich auswirkt.

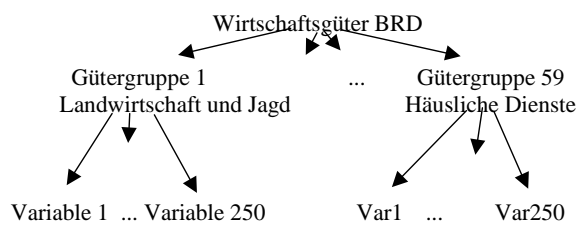


Abb. 4.3 Aufteilung der Produktion in Untergruppen

Die Gütergruppen können auch in 11 Wirtschaftsbereiche (Verarbeitendes Gewerbe, Kredit/Versicherungsgewerbe, Verkehr, Gastgewerbe, Baugewerbe, Energie- und Wasserversorgung etc.) mit Hilfe einer Input-Output-Tabelle (Kapitel 2.1) zusammengefasst werden und so mit der Entwicklung der Wirtschaftsbereiche die wirtschaftliche Entwicklung (Strukturwandel) der Bundesrepublik beschrieben werden.

Natürlich wirken sich in diesem Modell auch Politik und Gesetzgebung der 16 Bundesländer aus. Es gibt deshalb das Submodell LÄNDER, das die Wechselwirkungen zwischen den Bundesländern anhand der Produktionsbereiche zusammenfasst. Für jedes Bundesland ist die Produktion in jeweils 11 Wirtschaftsbereichen zusammengefasst. Dies gestattet sowohl eine Aussage über die Wirtschaftsentwicklung und Strukturänderungen einzelner Branchen (Steinkohleabbau!) als auch der einzelner Bundesländer. In Abb. 4.4 ist dies übersichtswise gezeigt. Der obere grau schraffierte Block zeigt die Komponenten der Endnachfrage für das Bruttoinlandsprodukt. Der untere graue Kasten modelliert die Auswirkungen der Endnachfrage und der Wertschöpfung auf die 11 Wirtschaftszweige im INFORGE-Modell, wobei hier nur die landesbezogenen Beträge betrachtet werden.

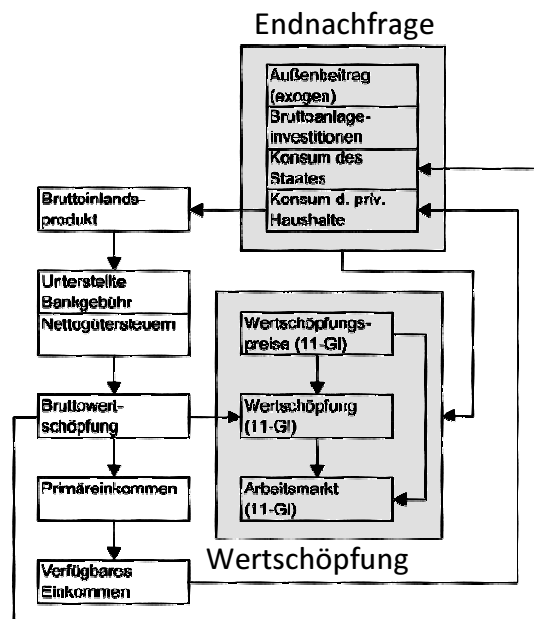


Abb. 4.4 Die Wirkungsstruktur der Ländermodelle (nach Distelkamp 2003)

Jedes dieser 16 Ländermodelle enthält eigene Konstanten und damit eine eigene Entwicklung seiner Variablen. Die Abstimmung aller Ländermodelle miteinander („Skalierung“) und dem INFORGE-Modell von Deutschland, das nur die gesamten Wirtschaftszweige betrachtet und nicht die Länder, ist ein wichtiger Prozess. Dabei wird eine Input-Output-Analyse vorgenommen, wobei sich die Bruttowertschöpfung (z.B. Produktion, Investitionen, Konsum) der 11 verschiedenen Wirtschaftsbereiche auf die 16 Bundesländer verteilt. Es ergibt sich jeweils eine 11×16-Matrix, siehe Abb. 4.5.

Länderdaten der Wirtschaftsbereiche

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
r_i																	1
																	2
																	3
																	4
																	5
																	6
																	7
																	8
																	9
																	10
																	11
																	INFORGE-Daten der Wirtschaftsbereiche
																	s_j

Abb. 4.5 Die Input-Output-Tabellenstruktur für die Güternachfrage/Produktion

Die Abstimmung der Länderdaten mit den INFORGE-Daten wird mit Hilfe eines einfachen Verfahrens vorgenommen. Dazu wird zunächst eine provisorische, geschätzte Matrix $A(0)$ initial erstellt; für jedes Bundesland werden unkoordiniert die Zahlen zusammengetragen und in die entsprechende Spalte eingesetzt. Eine Spaltensumme s_j der Matrix ist also die Summe der Wertschöpfung aller Wirtschaftsbereiche in diesem Bundesland j ; eine Zeilensumme r_i der Matrix ist die Summe der Wertschöpfungen aller Bundesländer für einen Wirtschaftsbereich i . Nun besteht die Aufgabe darin, diese Matrix aus unkoordinierten Daten mit den Daten des INFORGE-Modells abzugleichen. Dazu sind nur die pauschalen Zeilensummen aus INFORGE bekannt; die Aufteilung auf die einzelnen Länder muss ermittelt werden. Sei $r_i(0)$ die Zeilensumme der Initialmatrix und r_i^* die Zeilensumme von INFORGE, die zu erreichen ist. Für diesen Prozess werden alle Tabelleneinträge A_{ij} der Zeile mit dem selben Koeffizienten multipliziert

$$A_{ij}(1) = A_{ij}(0) \cdot r_i^*/r_i(0) \text{ mit } j = 1 \dots n \tag{4.1}$$

Dies führen wir für alle Zeilen durch. Damit haben alle Zeilen die richtigen Zeilensummen, weil

$$r_i(1) = \sum_{j=1}^n A_{ij}(1) = \sum_{j=1}^n A_{ij}(0) \cdot \frac{r_i^*}{r_i(0)} = \frac{r_i^*}{r_i(0)} \sum_{j=1}^n A_{ij}(0) = \frac{r_i^*}{r_i(0)} r_i(0) = r_i^* \quad (4.2)$$

Leider haben wir die Tabellenwerte dadurch so verändert, dass dann die Spalten falsche Spaltensummen aufweisen. Dies korrigieren wir, indem wir danach auch alle Spalten mit dem passenden Koeffizienten multiplizieren

$$A_{ij}(2) = A_{ij}(1) s_j^*/s_j(0) \quad \text{mit } i = 1 \dots m \quad (4.3)$$

Leider haben wir aber durch die Korrektur wieder falsche Zeilensummen erhalten, so dass wir den Algorithmus erneut durchführen müssen. Der Algorithmus konvergiert und nach einer gewissen Anzahl von Iterationen können wir anhalten.

Formal können wir beide Operationen, die Korrektur der Zeilensumme und die der Spaltensumme, als eine Operationsfolge hinschreiben. Dazu bilden wir eine Diagonalmatrix Matrix $R(k)$, die als n Hauptdiagonalelemente die Koeffizienten $r_i^*/r_i(k)$ enthält, und eine Diagonalmatrix $S(k)$, die als Hauptdiagonalelemente die m Koeffizienten $s_j^*/s_j(k)$ enthält. Eine Multiplikation mit S von rechts an A multipliziert die j -te Spalte von A mit dem Koeffizienten $s_j^*/s_j(k)$, so dass $A \cdot S$ die Spaltenskalierung durchführt. Entsprechendes gilt auch für die Zeilen, so dass die Iteration vom k -ten Schritt auf den $k+1$ -ten Schritt auch als

$$\begin{aligned} \hat{A}(k+1) &= R(k) A(k) \\ A(k+1) &= \hat{A}(k+1) S(k+1) \end{aligned}$$

schreiben lässt und damit

$$A(k+1) = R(k) A(k) S(k+1) \quad \text{RAS-Skalierung} \quad (4.4)$$

wird. Das Verfahren wird „RAS“-Verfahren genannt. Da es mit den (positiven) Diagonalmatrizen zweifach proportional skaliert, wird es auch als „biproportionale Skalierung“ bezeichnet. Allgemein wird das Verfahren für die Anpassung von existierenden Input-Output-Tabellen an neue Nebenbedingungen, neue Zeilen- und Spaltensummen, benutzt.

Das Ergebnis konvergiert zwar in der Praxis nach einigen Iterationen, aber eine eindeutige Lösung ist nicht garantiert, da die $n \times m$ Matrixkoeffizienten durch die $n+m$ Nebenbedingungen nicht eindeutig festgelegt werden können. Trotzdem liefert es bei einer guten Ausgangsmatrix eine brauchbare Lösung.

4.2 Hierarchiebildung durch lineare Multi-Level-Modellierung

Die Zusammenfassung (*aggregation*) bestehender Module zu einem Gesamtsystem ist nur ein Ansatz. Ein anderer besteht darin, existierenden Kontext schrittweise in das Modell hineinzunehmen, um es Schritt für Schritt zu verfeinern und komplex aufzubauen.

In den Sozialwissenschaften hat sich in letzter Zeit die Erkenntnis verbreitet, dass man die Eigenschaften der Individuen nicht nur auf der Ebene von Einzelpersonen, Gruppen oder Gesellschaften beschreiben sollte, sondern alle Ebenen miteinander verbinden muss, um ein korrektes Gesamtverhalten zu bekommen. Die Eigenschaften der Individuen kondensieren sich in Eigenschaften der Gruppe und die wiederum in Eigenschaften der Gesellschaft. In Abb. 4.6 sind die Unterteilungen visualisiert: Aus einzelnen, isolierten Einheiten wird ein Ganzes (das dunkle, große Oval) und das in den gesellschaftlichen Rahmen (schraffierter Kasten) eingebettet.

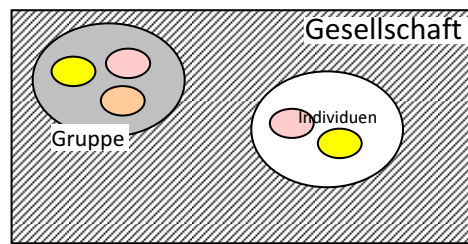


Abb. 4.6 Die vielschichtige Zusammensetzung gesellschaftlichen Verhaltens

Wir wählen uns also keine Abstraktionsebene aus, auf der modelliert wird, sondern beziehen alle Ebenen in die Modellierung ein, wobei die Parameter der einen Ebene durch Mechanismen der darunter liegenden Ebene modelliert werden. Konkret ist dieser Gedanke allerdings meist nur in linearen Systemen aus Kapitel 2.2 verwirklicht. Betrachten wir dazu als Beispiel eine Aussage für eine Schulklasse i über die mittleren Leistungspunkte y_i in Abhängigkeit vom Arbeitseinsatz x_i der Klasse.

$$y_i = a_0 + a_1 x_i + \varepsilon_i$$

Dies modellieren wir mit dem Anfangswert a_0 , der Steigung a_1 und einem Rauschterm ε für jede der Gruppen. Möchten wir eine Aussage über einen einzelnen Schüler machen, so muss die obige Gleichung noch mit dem Index j des Schülers innerhalb der Gruppe (Klasse) versehen werden. Dies sei die Ebene 1 der Modellierung.

$$\text{Ebene 1: } y_{ij} = a_{0j} + a_{1j} x_{ij} + \varepsilon_{ij} \quad (4.5)$$

Nehmen wir an, dass die Parameter a_{0j} und a_{1j} selbst nur Mittelwerte der Individuen sind und von anderen Variablen, etwa der individuellen Fernsehzeit w_j , beeinflusst werden, so lässt sich dies als Funktion einer anderen Variablen w_j für jedes Mitglied j schreiben:

$$\begin{aligned} \text{Ebene 2: } a_{0j} &= \alpha_{00} + \alpha_{01}w_j + u_j \\ a_{1j} &= \alpha_{10} + \alpha_{11}w_j + v_j \end{aligned} \quad (4.6)$$

Auch hier gibt es Parameter α_{ij} und die Rauschterme u_j und v_j der Messungen. Je nach dem, welche Abhängigkeiten man annimmt oder weglässt, resultieren unterschiedliche Multilevel-Modelle. Setzen wir beispielsweise $x_{ij} = w_j \equiv 0$, so resultiert das ANOVA-Modell, das nur Zufallseinflüsse kennt; bei $w_j = v_j \equiv 0$ (ANCOVA-Modell) ist der lineare Einfluss erhalten.

Man kann diesen Gedanken auch auf die nächste Ebene übertragen und so fort; üblich sind aber nur 2 oder 3 Ebenen. Natürlich kann man die Gleichung aus Gl. (4.6) auch direkt in Gl. (4.5) einsetzen, aber dies erzeugt nur eine komplizierte Gleichung mit „gemischten Effekten“ (*mixed model*). Die Herkunft der Einzeleffekte lässt sich leichter in einem Multilevel-Modell erfassen.

Allgemein gilt der Ansatz auf Ebene 1 wie auch auf Ebene 2 auch für mehrere Variablen entsprechend unserer multiplen linearen Regression, s. Kapitel 2. Die Wahl der Variablenzahl pro Ebene ist Sache der inhaltlichen Modellierung und nicht immer einfach zu bestimmen: Soll man einen Einfluss auf der gleichen Ebene durch eine weitere Variable einbringen, oder als Kontext betrachten und der nächsten Ebene zufügen? Hier helfen Tests weiter, ähnlich wie im Beispiel des nächsten Abschnitts.

Beispiel Modellierung des Einflusses der Tabakindustrie

Angenommen, wir betrachten als Individuen die Senatoren des US-Senats und möchten herausfinden, durch was ihr Abstimmungsverhalten bestimmt wird, etwa bei Abstimmungen über Gesetze, die die Tabakindustrie betreffen. Das Abstimmungsverhalten jedes Kongressmitglieds ist in öffentlich zugängigen Datenbanken erfasst.

Im Unterschied zu Deutschland müssen die US-Senatoren außerdem offen legen, von wem sie in welcher Höhe Spenden erhalten haben. Auch ist aus Statistiken bekannt, dass Republikaner häufiger im Sinne der Tabakindustrie gestimmt haben als Demokraten. Wenn überhaupt ein Einfluss auf die Abstimmungen vorliegt, dann kann also wahrscheinlich der Gruppeneinfluss p , ob Republikaner ($p = 1$) oder Demokrat ($p = 0$) und der Spendenfluss m [Tausend \$] dafür verantwortlich sein. Wir modellieren also zunächst das Abstimmungsverhalten y , die relative Anzahl der Zustimmungen des Abgeordneten zu den von der Tabakindustrie befürworteten Vorlagen, mit unserem einfachen Level-1 Ansatz

$$y_{ij} = a_{0j} + a_{1j}p_{ij} + a_{2j}m_{ij} + \epsilon_{ij} \quad (4.7)$$

und versuchen, die Parameter mit möglichst geringstem mittleren quadratischen Fehler anzupassen. Für diesen Zweck gibt es spezielle Programmpakete, die mit dem *maximum-likelihood*-Ansatz bei gegebenem Datensatz die Parameter iterativ anpasst. Wenden wir dies auf unser vorgegebenes Problem an, so erhalten wir die Koeffizienten.

Warum sollten wir die Daten mit einem Multilevel-Ansatz modellieren, wenn auch ein einfacher linearer Level-1-Ansatz geht? Dafür gibt es mehrere Argumente.

- Zum einen können wir grafisch visuell die Abhängigkeiten der Entscheidungen vom Spendenaufkommen zeichnen und untersuchen. Dabei sehen wir, dass es ziemlich unterschiedliche Abhängigkeiten gibt: In manchen Staaten ist die Abhängigkeit positiv, aber in manchen Staaten hat die Regressionsgerade eine negative Steigung (z.B. ID oder SC, s. Abb. 4.7). Dies legt eine Abhängigkeit der Level-1 Modellierung von den Staaten nahe.

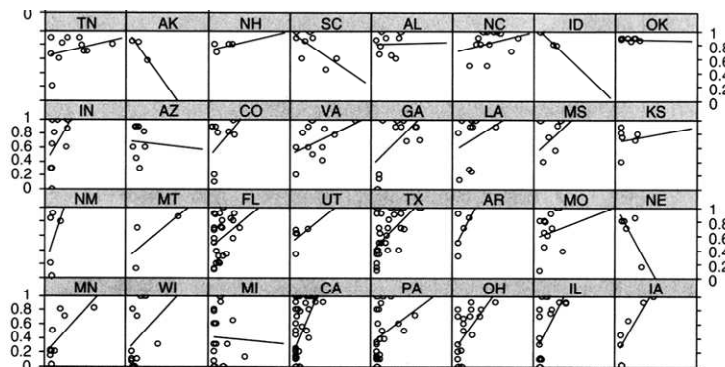


Abb. 4.7 Der mittlere Spendeneinfluss je nach US-Staat (nach Luke 2004)

- Weiterhin können wir testen, ob die Voraussetzung für den linearen Ansatz mit der Fehlerminimierung gegeben ist, nämlich die Unabhängigkeit der Beobachtungen. Dazu testen wir die Datensamples aller Senatoren und bilden den Interklassen-Korrelationskoeffizienten (ICC) ρ

$$\rho = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_\varepsilon^2} \quad \text{inter class correlation coefficient} \quad (4.8)$$

Die Varianzen der Rauschanteile ε von Ebene 1 und u von Ebene 2 erhalten wir, indem wir die Parameter eines Null-Modells an die Daten anpassen. Ein Null-Modell ist das Modell aus Gl. (4.5), das wir erhalten, wenn wir in Gl. (4.5) $x_{ij} = 0$ und in Gl. (4.6) $w_j = 0$ setzen, also das ANOVA-Modell verwenden. Hier ist a_0 die mittlere Schätzung über alle Kongressmitglieder, u_j modelliert die Variabilität zwischen den Staaten i und ε_{ij} die Variabilität zwischen den Mitgliedern j eines Staates i .

Die Anpassung ergibt in unserem Beispiel $\sigma_u = 0,035$ und $\sigma_\varepsilon = 0,093$, also $\rho = 0,27$ (Luke 2004). Dies ist relativ hoch: Die Zugehörigkeit zu einem Staat i erklärt also 27% der Varianz. Der hohe ICC-Wert deutet darauf hin, dass die

Beobachtungen nicht unabhängig von einander sind und damit die Voraussetzung der einfachen Modellierung nicht erfüllt ist. Für eine bessere Modellierung sollten wir also eine weitere Ebene einziehen, auf der die Staatenzugehörigkeit als Einflussfaktor modelliert wird. Der Grund dafür scheint zu sein, dass die Senatoren eines Staates ähnlicher zueinander sind als zu anderen Staaten.

- Ein wichtiges Argument für die Multilevel-Modellierung ist grundsätzlicher Natur: Wenn es einen vernünftigen Grund dafür gibt, eine Multilevel-Struktur anzunehmen, etwa weil die Mentalität eines Staates seine Repräsentanten beeinflusst, dann muss man ein solches Modell ausprobieren, um falsche Aussagen zu vermeiden.

Wir stellen deshalb zwei Multilevelmodelle auf. Basis der Modellierung soll der Staateneinfluss sein. Ein plausibler Indikator dafür ist die Frage, wie stark der Staat mit der Tabakindustrie verknüpft ist. Dies können wir durch die mittlere Tabakernte (gemessen in geernteter Anbaufläche) pro Staat messen. Für die Multilevel-Modellierung müssen zuerst die Parameter der Level-1 Modellierung angepasst werden, und danach ihre Modellierung durch eine Level-2 Modellierung näher untersucht werden.

Das einfachste Modell (**Modell 0**) ist das ANOVA-Modell. Hier erhalten wir bei $p = 0$ und $m = 0$ den Mittelwert des Abstimmungsverhaltens: $a_{0j} = 0,52$. Dies bedeutet, dass im Mittelwert 52% aller Senatoren des Kongresses für ein pro-Tabak-Gesetz stimmen.

Für **Modell 1** erhalten wir mit der Modellierung

$$\begin{aligned} a_{0j} &= \alpha_0 + u_{0j} & \alpha_0 &= 0,22 \\ a_{1j} &= \beta_0 + u_{1j} & \beta_0 &= 0,48 \\ a_{2j} &= \gamma_0 + u_{2j} & \gamma_0 &= 0,0046 \end{aligned} \quad (4.9)$$

Bei der Anpassung der Parameter ergeben sich obigen Werte (Luke 2004). Wir interpretieren dabei α_0 als den Mittelwert über alle Zustimmungsentscheidungen, wenn $p = m = 0$ ist. Der Effekt „Republikaner“ bewirkt für $p = 1$ einen Anstieg von y um β_0 , also um 48%. Der Effekt „Spendengeld“ bewirkt ($m = 1$) eine Zunahme der Pro-Tabak-Entscheidungen um 0,46% pro 1000\$.

Betrachten wir nun **Modell 2**. Die Modellierung lautet für Level 2

$$\begin{aligned} a_{0j} &= \alpha_0 + \alpha_1 e_j + u_{0j} & \alpha_0 &= 0,22, \alpha_1 = 0,0005 \\ a_{1j} &= \beta_0 + u_{1j} & \beta_0 &= 0,48 \\ a_{2j} &= \gamma_0 + u_{2j} & \gamma_0 &= 0,0045 \end{aligned} \quad (4.10)$$

und für **Modell 3**

$$\begin{aligned} a_{0j} &= \alpha_0 + \alpha_1 e_j + u_{0j} & \alpha_0 &= 0,18, \alpha_1 = 0,0027 \\ a_{1j} &= \beta_0 + \beta_1 e_j + u_{1j} & \beta_0 &= 0,5, \beta_2 = -0,0016 \end{aligned} \quad (4.11)$$

$$a_{2j} = \gamma_0 + \gamma_1 e_j + u_{2j}$$

$$\gamma_0 = 0,005, \gamma_1 = -0,00002$$

Die Anpassung der Parameter erfolgte wieder durch *maximum likelihood*-Approximation.

Vergleichen wir die drei Modelle mit einander, so finden wir folgendes:

- Die Variablen p (party) und m (money) sind hoch-signifikante Prädiktoren für alle drei Modelle.
- Die Variable e ist in Modell 2 hoch signifikant, hat aber eine geringe Wirkung: pro 1000 *acres* gibt es nur $\alpha_1 = 0,05\%$ mehr an Tabak-positiven Stimmverhalten.
- In Modell 3 beobachten wir eine hohe Wechselwirkung zwischen den Ebenen: Die mittlere Verschiebung durch den Staateneinfluss e erhöht sich von $\alpha_1 = 0,0005$ auf $\alpha_1 = 0,0027$ um das 5-fache; dafür dämpft der Effekt e der Tabakernte deutlich den Anstieg von y bei den Effekten der Einflussvariablen p („Partei“) und m („Spendengeld“). In Abb. 4.8 sind die Abhängigkeiten des Abstimmungsverhaltens vom Spendengeld bei verschiedenen Erntewerten visualisiert.

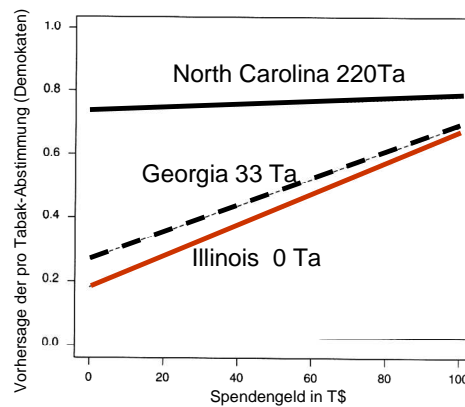


Abb. 4.8 Der mittlere Spendeneinfluss je nach Tabakernte (in Tausend acres) (nach Luke 2004)

Man kann deutlich erkennen, dass bei $p = 1$ (bei den Demokraten) der Geldeinfluss bei geringer Tabakernte (z.B. Illinois) wesentlich größer ist als bei hohem Erntevolumen (North Carolina), wohl weil die Lobby der Tabakbauern in diesen Staaten schon bei den Wahlen stärker sind. Das Multi-Level-Modell vermag also nicht nur wesentlich genauer, sondern auch differenzierter die Wirkungseinflüsse zu beschreiben als die einfachen Modelle.

Eine solche Modellbildung stellt eine *bottom-up*-Methode dar, ein hierarchisches Modell zu erstellen. Wir betrachten nun eine Methode, ein vorhandenes, komple-

xes Modell durch einen *top-down*-Ansatz in „natürliche“ Untermodelle zu untergliedern.

4.3 Hierarchiebildung und Strukturanalyse

Angenommen, wir haben bereits ein Modell gebildet. Mit der Zeit wird es immer komplexer und wir verstehen kaum noch die eigentlichen Abhängigkeiten und Funktionen innerhalb des Modells. In diesem Fall hat es sich bewährt, das Modell in Submodelle zu untergliedern und es damit auf die „wesentlichen“ Strukturen zu reduzieren. Eine solche Zerlegung hat dabei folgende Vorteile:

- Ein klares Bild der kausalen Struktur kann bei der Konstruktion helfen, Modellierungsfehler zu vermeiden.
- Mit Hilfe der klaren Strukturen kann der Anwender leichter das Modell verstehen, d.h. seine Ziele und Annahmen zu erfassen und seine Hauptarchitektur zu verfolgen.
- Mit der Kenntnis der wichtigsten Ziele und Annahmen kann man auch leichter die theoretischen Eigenschaften und damit das Anwendungsgebiet des Modells abschätzen. Auch die Ergebnisse lassen sich so leichter interpretieren.
- Durch die Zerlegung des Modells in Mengen von zusammenhängenden Variablen lassen sich leichter die Punkte im Modell lokalisieren, die für eine bestimmte Analyse ausschlaggebend sind, unabhängig vom wesentlich größeren Modell, in das sie eingebettet sind.

Als Richtschnur soll uns dabei der Gedanke dienen, die „logisch-kausale“ Struktur des Modells zu erschließen.

Als wichtiges Hilfsmittel dabei dient uns die graphische Form des Modells. Betrachten wir dazu die allgemeinen differentiellen Systemgleichungen, wie sie in Abschnitt 3.3 eingeführt wurden

$$\mathbf{z}(t+1) = \mathbf{F}(\mathbf{z}(t), \mathbf{x}(t), t) \quad (4.12)$$

Im linearen, diskreten Fall ist dies

$$\mathbf{z}(t+1) = (\mathbf{I} + \mathbf{A}) \mathbf{z}(t) + \mathbf{B} \mathbf{x}(t) \quad (4.13)$$

Haben wir keine linearen Wirkungen, so können wir die Funktion $\mathbf{F}(\cdot)$ in einer Taylorreihe entwickeln und erhalten mit den Jacobi-Matrizen

$$\mathbf{A} = (A_{ij}) = \left(\frac{\partial F_i}{\partial z_j} \right), \quad \mathbf{B} = (B_{ij}) = \left(\frac{\partial F_i}{\partial x_j} \right) \quad (4.14)$$

lineare Approximationen erster Ordnung. Im Übrigen gelten die folgenden Betrachtungen ganz allgemein für alle Arten von Wirkungen zwischen Variablen z_i , unabhängig davon, ob sie linear oder nichtlinear sind.

Kombinieren wir die Eingabe (den Kontext) \mathbf{x} mit dem Zustand \mathbf{z} zu einem gemeinsamen Vektor $\mathbf{w} = (\mathbf{z}, \mathbf{x})^\top$, so erhalten wir

$$\mathbf{z}(t+1) = \begin{bmatrix} \mathbf{I} + \mathbf{A} & \mathbf{B} \end{bmatrix} \begin{pmatrix} \mathbf{z}(t) \\ \mathbf{x}(t) \end{pmatrix} = \mathbf{D} \mathbf{w}(t) \quad (4.15)$$

$$\mathbf{w}(t+1) = (\mathbf{z}(t+1), \mathbf{x}(t+1))^\top$$

Bei einem linearen System lässt sich der neue Zustand als Multiplikation mit einer Systemmatrix beschreiben. Wie wir in Abschnitt 3.1 sahen, entspricht eine solche Wirkungsmatrix einem Wirkgraphen. In ihm ist für jede Variable ein Knoten vorgesehen; existiert eine Wirkung von Variable i auf Variable j , so ist im korrespondierenden Graphen eine gerichtete Kante (Pfeil) von Knoten i auf Knoten j vorhanden. Die Matrix \mathbf{D} , in dessen Element D_{ij} die Wirkung von Knoten i auf Knoten j beschrieben wird, wird *Adjazenzmatrix* des Graphen genannt. Sie enthält für eine Verbindung eine eins; ist keine Verbindung enthalten, so enthält sie eine null. Die Stärke c_{ij} der Verbindung (die Größe des Einflusses) kann durch die Zuordnung $D_{ij} = c_{ij}$ ausgedrückt werden.

In den folgenden Abschnitten wollen wir einige Methoden kennen lernen, um die Komplexität in Modellen zu reduzieren und durch die vereinfachte Struktur des Modells Kausalitätsbeziehungen besser zu verstehen. Dabei folgen wir im Wesentlichen den Arbeiten von (Gilli 2004).

4.3.1 Graphenbasierte Analyse

Wir haben also einen Wirkungsgraphen und seine Adjazenzmatrix und wollen innere, logische Wirkungssubgraphen identifizieren um die Struktur besser zu verstehen. Dabei interessieren wir uns für besonders stark zusammenhängende Teile der Wirkstruktur und die innewohnenden Kausalketten. Was heißt "stark zusammenhängend"? Dazu müssen wir einige Begriffe der Graphentheorie einführen.

DEF. Sei V die Menge der Knoten $V = \{V_i\}$ und E die Menge der Kanten $E = \{E_{ij}\}$. Dann ist ein *Graph* $G = (V, E)$ das Tupel aus den Knoten und den dazwischen verlaufenden Kanten.

Für unseren Zweck identifizieren wir die Variablen z_i bzw. w_i mit den Knoten V_i und zeichnen genau dann eine Kante E_{ij} von V_i nach V_j , wenn z_i in der Funktion \mathbf{F} von Gl. (4.12) bzw. w_i in Gl.(4.15) auftritt. Die Struktur des Graphen lässt sich mit Hilfe der Adjazenzmatrix beschreiben:

DEF. Die *Adjazenzmatrix* $\mathbf{A} = (A_{ij})$ lässt sich elementweise definieren durch

$$A_{ij} = \begin{cases} 1 & E_{ij} \in E \\ 0 & \text{sonst} \end{cases}$$

- Ist $E_{ij} = E_{ji}$, so spricht man von einem *ungerichteten* Graphen. Wir betrachten aber zunächst nur gerichtete Graphen.
- Sei eine Knotenfolge V_1, V_2, \dots, V_n aus n Knoten gegeben, wobei die Kanten $E_{i,i+1} = (V_i, V_{i+1})$, $i = 1, \dots, n-1$ existieren. Eine solche Folge von Kanten wird auch als *Weg* bezeichnet.
- Ein *Zyklus* liegt vor, wenn der *Startknoten* V_1 und der *Endknoten* V_n des Weges identisch sind.
- Ein *Pfad* ist ein Weg, bei dem keine gleichen Knoten existieren, also $V_i \neq V_j$ bei $i \neq j$ und i, j aus $1 \dots n$.
- Sind Startknoten und Endknoten eines Pfades identisch, so spricht man von einem *Kreis*.

Nun können wir mit den eingeführten Begriffen den Begriff „streng zusammenhängend“ definieren:

DEF. Ein *streng zusammenhängender* Subgraph $G_s = (V_s, E_s) \subset G$ eines Graphen G ist durch folgende Eigenschaft gekennzeichnet: Seien $V_i, V_j \in V_s$. Wenn ein Pfad von Knoten V_i zu Knoten V_j existiert, so gibt es auch einen Pfad von V_j zu V_i .

Wie ermitteln wir die streng zusammenhängenden Komponenten (*Blöcke*) eines Graphen? Es existiert ein dazu ein einfacher Algorithmus (unter vielen):

- 1) Berechne die Menge $R(V_0)$, die Menge aller Knoten $V_i \in V$, die von V_0 aus erreichbar sind. Dies sind die Pfade von V_0 zu V_i . Da nur die Menge gefragt ist, reicht es, die Pfade zu betrachten, die keine Knoten gemeinsam haben (untereinander *knotendisjunkt* sind).
- 2) Berechne die Menge $Q(V_0)$, die Menge aller Knoten $V_i \in V$, von denen ein Pfad zu V_0 existiert. Auch hier sind die Pfade untereinander knotendisjunkt.
- 3) Berechne die Schnittmenge $R(V_0) \cap Q(V_0)$. Diese Menge definiert unsere gesuchte Menge von streng zusammenhängenden Komponenten.

Betrachten wir dazu folgendes Beispiel: Seien die Abhängigkeiten zwischen acht Variablen gegeben mit

$$\begin{aligned} z_1 &= F_1(z_2, z_3, z_4, z_7) \\ z_2 &= F_2(z_7) \\ z_3 &= F_3(z_6, z_9) \end{aligned} \tag{4.16}$$

$$z_4 = F_4(z_7)$$

$$z_6 = F_6(z_1, z_3, z_4)$$

$$z_7 = F_7(z_2, z_5)$$

$$z_8 = F_8(z_5, z_7, z_9)$$

In Abb. 4.9 ist der korrespondierende Graph und seine Adjazenzmatrix gezeigt. Dabei bezeichnen wir den einzelnen Knoten direkt durch seinen Index.

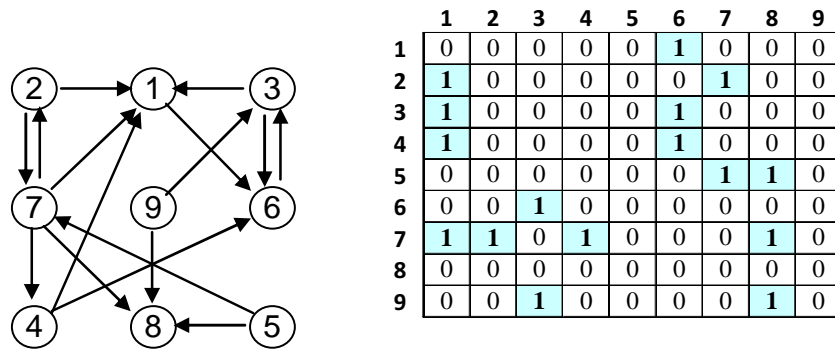


Abb. 4.9 (a) Graph und (b) korrespondierende Adjazenzmatrix

Wie wir sehen, reicht es für die logische Wirkungsstruktur aus, eine einfache Adjazenzmatrix mit 1/0-Gewichten aufzustellen.

Nun ermitteln wir die streng zusammenhängenden Komponenten des Graphen durch eine Breitensuche. Beispielsweise für den Knoten V_2 ist dies mit $R(2)$, $Q(2)$ gegeben:

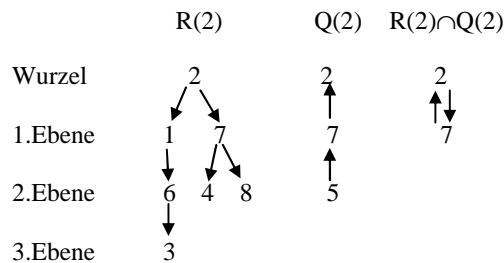


Abb. 4.10 Die streng zusammenhängende Komponente mit V_2

Um eine weitere streng zusammenhängende Komponente (Block) zu ermitteln, müssen vorher alle Knoten und Kanten der bereits ermittelten Komponenten im Graphen gestrichen werden, s. Abb. 4.11 (a). Danach kann man erneut den obigen

Algorithmus darauf anwenden, etwa auf den Knoten V_1 , s. Abb. 4.11 (b). In unserem Beispiel sind alle restlichen Knoten nicht streng zusammenhängend, jeder Knoten bildet einen Block. Man beachte, dass wir bei beiden Blöcken auch mit einem anderen Knoten hätten beginnen können, etwa beim ersten Block mit 7 statt mit 2, und trotzdem denselben Block erhalten hätten. Wir erhalten so die Mengen $\{1,3,6\}, \{2,7\}, \{4\}, \{5\}, \{8\}, \{9\}$.

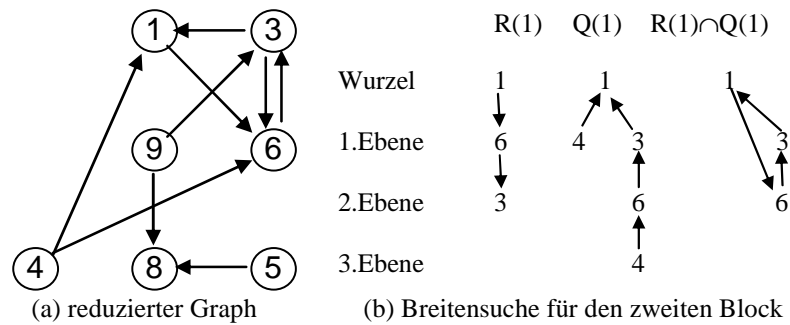


Abb. 4.11 Ermittlung der nächsten streng zusammenhängenden Komponente

Nun fassen wir im alten Graphen alle Knoten eines Blocks zu einem neuen Knoten zusammen und erhalten so einen neuen Graphen, den *reduzierten* Graphen. Für unser Beispiel ist er in Abb. 4.12 abgebildet, zusammen mit seiner Adjazenzmatrix.

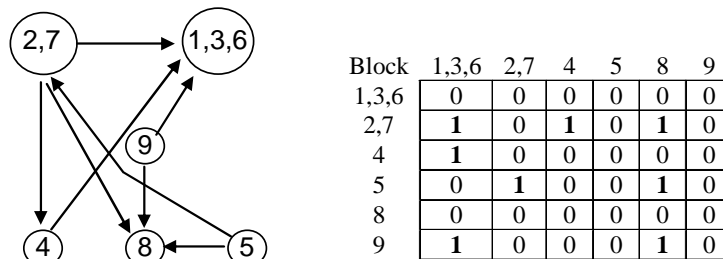


Abb. 4.12 Reduzierter Graph der Blöcke und seine Adjazenzmatrix

Nun wäre es sinnvoll, auch in dem reduzierten Graphen Blöcke zu bilden. Ist dies möglich, gibt es streng zusammenhängende Komponenten in diesem Graphen? Die Antwort lautet: Nein! Der Grund dafür liegt in dem Konstruktionsverfahren der Blöcke: Würde innerhalb des reduzierten Graphen ein Kreis existieren, so wäre er bereits bei der Bildung einer der Blöcke erkannt worden: Alle seine Kno-

ten wären sowohl bei R als auch bei Q aufgetreten und somit in den Block eingegliedert worden. Da wiederum jeder der eingegliederten Knoten streng zusammenhängend mit den anderen Knoten seines Blockes war, wären alle betroffenen Blöcke zu einem großen Block vereinigt worden, was nicht der Fall war. Also gibt es keine zwei Blöcke im reduzierten Graphen, die streng zusammenhängend sind.

Es hat also keinen Sinn, den obigen Algorithmus für die Suche streng zusammenhängender Komponenten auf den reduzierten Graphen anzuwenden. Wie können wir dann den Graphen weiter zusammenfassen und komprimieren, ohne dabei logisch-kausale Zusammenhänge zu verlieren?

Modellseparierung nicht-zyklischer Kausalketten

Für diesen Zweck ordnen wir den Graphen in Ebenen an. Dies geht, weil ja keine Kreise existieren können. Ebene 0 enthält immer die Eingaben x_i (exogenen Variablen), jede Ebene darunter die Wirkungen der Komponenten darüber, und die letzte Ebene die Variablen, auf die nur Wirkungen gehen, Ausgabevariablen y_i . Für unser Beispiel ist dies in Abb. 4.13(a) gezeigt. Dabei können wir auch die Zeilen und Spalten der dazu gehörenden Adjazenzmatrix entsprechend anordnen, so dass zuerst die Blöcke (Knoten) der Ebene 0, dann der Ebene 1 usw. eingetragen werden. Die Spalten aller Komponenten der Ebene 0 können wir weglassen, da sie ja keine Einflüsse (gerichtete Kanten) empfangen, siehe Abb. 4.13(b). Die resultierende Adjazenzmatrix hat eine Form, die als „obere Dreiecksmatrix“ bekannt ist: Alle von null verschiedenen Einträge finden sich oberhalb der Hauptdiagonale, da eine Wirkung $A_{ji} \neq 0$ mit $j > i$ auf eine Rückwirkung innerhalb einer Ebene oder eine Ebene hinauf bedeuten würde. Dies ist aber von der Konstruktion her ausgeschlossen.

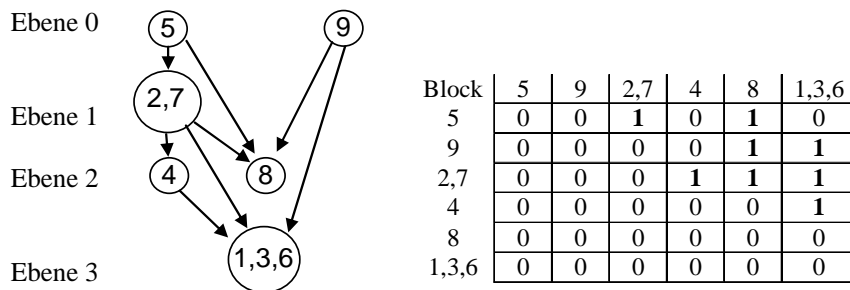


Abb. 4.13 (a) angeordneter Graph (b) Adjazenzmatrix

Die Adjazenzmatrix zeigt nur eine direkte Wirkung von einer Variablen auf die andere an. Uns geht es aber um Kausalitätsstrukturen, also ganze Ketten von Wirkungen, bei denen eine Variable über andere hinweg eine dritte Variable beeinflussen kann. Wollen wir auch diesen Gedanken einfließen lassen, so erweitern

wir die Matrix der direkten Wirkungen um alle indirekten Wirkungen. Die so erreichte Matrix heißt Erreichbarkeitsmatrix C .

DEF Die *Erreichbarkeitsmatrix* $C = (C_{ij})$ lässt sich elementweise definieren durch

$$C_{ij} = \begin{cases} 1 & \text{ex. Pfad von } V_i \text{ nach } V_j \text{ der Länge } 1 \\ "i" & \text{ex. Pfad von } V_i \text{ nach } V_j \text{ der Länge } > 1 \\ 0 & \text{sonst} \end{cases}$$

Die Adjazenzmatrix wird also mit einigen Einträgen angereichert, die anstelle einer Null ein „i“ für „indirekte Wirkung“ anzeigen.

Nun wenden wir uns wieder der Frage zu: Wie kann ein Modell in Untermodelle unterteilt werden? Konkret bedeutet das für einen Block i , dass wir eine Menge von Knoten bzw. Blöcke suchen, die Block i enthält und möglichst klein ist, aber von Block i verschieden. Die größte Menge ist leicht zu finden: Sie ist das Gesamtmodell selbst. Wie aber finden wir die kleinste Menge?

Reduzierte Strukturen

Möchten wir ein Submodell bezüglich eines direkten Einflusses i finden, das das gleiche Verhalten wie das Gesamtmodell vollbringt, so müssen wir nur alle Variablen finden, auf die der Einfluss i Auswirkungen hat und die wiederum auf die Ausgabe j Auswirkungen haben. Für die Lösung benutzen wir einen ähnlichen Gedanken wie bei dem Finden der streng zusammenhängenden Komponenten selbst:

- 1) Berechne die Menge $R(V_i)$, die Menge aller Blöcke $V_k \in V$, die von V_i beeinflusst sind. Dies sind alle Blöcke, die von Block V_i erreicht werden können, also alle Blöcke in der i -ten Zeile der Erreichbarkeitsmatrix C , bei denen $C_{ik} \neq 0$ ist, zuzüglich V_i . Beispiel: Block 9 beeinflusst $R(V_9) = \{9, 8, (1,3,6)\}$.
- 2) Berechne die Menge $Q(V_j)$, die Menge aller Knoten $V_k \in V$, von denen ein Einfluss auf V_j existiert. Dies sind alle Blöcke, die Block V_j erreichen kann, also alle Blöcke in der j -ten Spalte der Erreichbarkeitsmatrix C , bei denen $C_{kj} \neq 0$ ist, zuzüglich V_j . Beispiel: Block 1 wird beeinflusst von $Q(V_1) = \{(1,3,6), 9, 5, (2,7), 4\}$.
- 3) Berechne die Schnittmenge $R(V_i) \cap Q(V_j)$. Diese Menge definiert unsere gesuchte Menge von Komponenten, die Wirkungen von Block i auf Block j ausdrücken.

Beispiel: Das Submodell für den Einfluss von Block 9 auf Block 1 ist $R(V_9) \cap Q(V_1) = \{(1,3,6),9\}$; Blöcke 4,5,(2,7),8 brauchen wir nicht.

Damit haben wir alle Knoten zusammengefasst, die für die direkten Auswirkungen von Knoten i auf Knoten j zuständig sind. Alle anderen Knoten können wir wegfassen lassen, ohne das Systemverhalten zu verändern.

Minimale Strukturen

Unsere Überlegungen für eine Vereinfachung des Modells können wir mit weiteren Ansätzen vorantreiben. Da wir hier nur die prinzipiellen, kausalen Einflüsse und nicht die quantitativen Beeinflussungen untersuchen, können wir den Erreichbarkeitsgraphen weiter vereinfachen. Dazu können wir alle Verbindungen weglassen, die nicht unbedingt erforderlich sind, um die kausale Struktur wiederzugeben. Dies sind insbesondere alle Einflüsse, die sowohl direkt als auch indirekt von Block i nach Block j gehen. Markieren wir zuerst die Wirkungspfade, die nicht wegfassen dürfen, ohne die ein Einfluss nicht möglich wäre und ohne die somit die kausale Struktur zerstört sein würde („Basisverbindung“), so können wir alle übrigen Direktverbindungen weglassen, da sie nicht notwendig sind. An ihrer Stelle reichen die indirekten Verbindungen und die Basisverbindungen aus, um die Kausalstruktur zu erhalten. Der resultierende Graph wird „Minimalgraph“ genannt.

In Abb. 4.14(a) ist dies an unserem Beispiel gezeigt. Die Basisverbindungen sind durchgehend, die entbehrlichen Verbindungen gestrichelt eingezeichnet. Die Erreichbarkeitsmatrix in Abb. 4.14(b) enthält an den entsprechenden Einträgen ein „b“, „i“ oder „d“ für Basisverbindung, Indirektverbindung oder Direktverbindung.

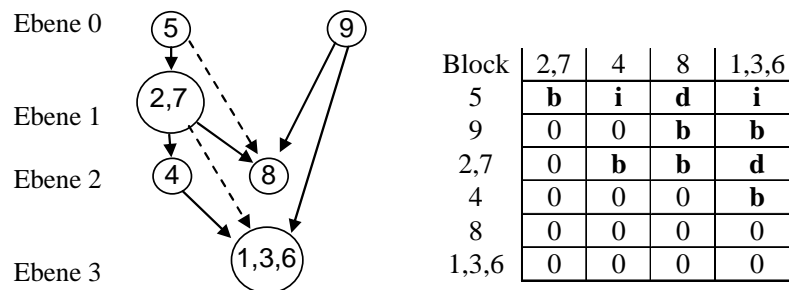


Abb. 4.14 (a) Minimalgraph (b) Erreichbarkeitsmatrix

Würden wir neue kausale Wirkungen in das Modell einbauen, so können sie nur bei Einträgen mit „0“ erfolgen; bei einem Eintrag mit „i“ erhalten wir keine neue Wirkstruktur.

Abhängigkeitsstrukturen

Innerhalb der Blöcke sind die Variablen gegenseitig abhängig; sie bilden einen Kreis. In ökonomischen Modellen gibt es hunderte bis tausende solcher Kreise. Um auch die Struktur solcher komplexer Strukturen besser zu verstehen und eine Strukturierung innerhalb streng zusammenhängender Subgraphen (Manchmal fast das gesamte Modell!) zu ermöglichen, ist es sinnvoll, nach Elementen zu suchen, die für den Kreislauf verantwortlich sind. Dies ist

- die Knotenmenge, von deren Elemente (Variable) mindestens eines in jedem der Kreisläufe vorhanden ist („essentielle Rückkopplungsknotenmenge“), sowie
- die Kantenmenge, aus der in jedem Kreislauf mindestens eine Kante vorhanden ist („essentielle Rückkopplungskantenmenge“).

Solche Mengen lassen sich durch Algorithmen ermitteln, s. z.B. (Gilli 1992). Beispielsweise können wir Graphen betrachten, in denen die Kreisläufe den enthaltenen Knoten bzw. Kanten gegenübergestellt werden. Am Beispiel des Blocks $\{1,3,6\}$ ist dies in Abb. 4.15 gezeigt. Wir sehen, dass die essentiellen Rückkopplungsknotenmengen aus $\{3\}$ und $\{6\}$ bestehen: jede der Mengen deckt alle Kreise ab. Die essentielle Rückkopplungskantenmenge ist $\{d\}$.

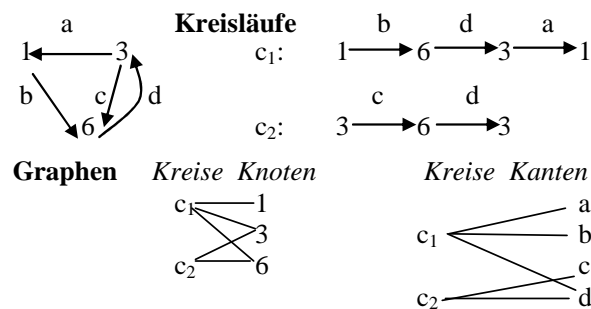


Abb. 4.15 Essentielle Mengen und bipartite Graphen

Die Graphen haben die Eigenschaft, dass sie sich in zwei Mengen („Kreise“ sowie „Knoten“ bzw. „Kanten“) derart aufteilen lassen, dass alle Verbindungen zwischen Elementen der Gesamtmenge aus der einen Menge („Kreise“) in die andere („Knoten“) führen. Verbindungen innerhalb einer Menge gibt es nicht. Solche Graphen heißen *bipartite* Graphen.

4.3.2 Zeitverschobene Einflüsse

Betrachten wir den Fall, dass Variablen auftauchen, die einen Wert von einem früheren Zeitpunkt enthalten. Beispielsweise wird Gl. (4.16) zu

$$\begin{aligned} z_1 &= F_1(z_2, z_3, z_4, z_7) \\ z_2 &= F_2(z_7) \\ z_3 &= F_3(z_6, z_9) \\ z_4 &= F_4(z_7) \\ z_6 &= F_6(z_1, z_3, z_4, z_8(t-1)) \\ z_7 &= F_7(z_2, z_5, z_3(t-1)) \\ z_8 &= F_8(z_5, z_7, z_9) \end{aligned}$$

Die so definierten Gleichungen enthalten die Variablen z_6 und z_7 , die abhängig von dem Wert für z_8 bzw. z_3 an einem früheren Zeitpunkt sind. Wie können wir solche Fälle behandeln? Wie erreichen wir eine kausale Dekomposition des Modells über Zeitgrenzen hinweg?

Die Lösungsidee besteht darin, Kopien des Modells zu den Zeitpunkten t und $t+1$ zu erstellen und den zeitlichen Einfluss in einen räumlichen Einfluss umzuwandeln, den wir mit unseren bisher entwickelten Methoden behandeln können. Formal können wir mit Gl. (4.12) die Systemgleichungen schreiben als

$$\mathbf{z}(t) = \mathbf{F}(\mathbf{z}(t), \mathbf{z}(t-1), \mathbf{x}(t))$$

Angenommen, externe Beeinflussungen erstrecken sich über drei Zeitpunkte. Dann gilt für drei Perioden (Periode 0 hat keinen Vorgänger)

$$\begin{aligned} \mathbf{z}(t+0) &= \mathbf{F}(\mathbf{z}(t+0), \mathbf{x}(t)) \\ \mathbf{z}(t+1) &= \mathbf{F}(\mathbf{z}(t+0), \mathbf{z}(t+1), \mathbf{x}(t+1)) \\ \mathbf{z}(t+2) &= \mathbf{F}(\mathbf{z}(t+1), \mathbf{z}(t+2), \mathbf{x}(t+2)) \end{aligned} \quad (4.17)$$

Fassen wir die drei Spaltenvektoren von \mathbf{z} zu einem großen Spaltenvektor $\mathbf{z}(t) = (\mathbf{z}(t+0), \mathbf{z}(t+1), \mathbf{z}(t+2))^T$ und die drei Spaltenvektoren von \mathbf{x} zu einem großen Spaltenvektor $\mathbf{x}(t) = (\mathbf{x}(t+0), \mathbf{x}(t+1), \mathbf{x}(t+2))^T$ zusammen, so lässt sich Gl.(4.17) schreiben als

$$\mathbf{z}(t) = \mathbf{F}(\mathbf{z}(t), \mathbf{x}(t))$$

und damit wie vorher behandeln. Die Erreichbarkeitsmatrix \mathbf{C} des resultierenden Graphen besteht aus der ursprünglichen Matrix \mathbf{C}_0 und neuen Erreichbarkeitsmatrizen \mathbf{C}_1 und \mathbf{C}_2 , die die Erreichbarkeit eines Knotens aus dem Graphen zum Zeitpunkt $t+1$ bzw. $t+2$ von einem Knoten aus dem Graphen zum Zeitpunkt t beschreiben

$$\mathbf{C} = \begin{pmatrix} \mathbf{C}_0 & \mathbf{C}_1 & \mathbf{C}_2 \\ \mathbf{0} & \mathbf{C}_0 & \mathbf{C}_1 \\ \mathbf{0} & \mathbf{0} & \mathbf{C}_0 \end{pmatrix} \quad (4.18)$$

In unserem Beispiel besteht dann der Gesamtgraph aus Kopien des Graphen zum Zeitpunkt null, die durch zeitliche Einflüsse verbunden sind. In (4.16) ist der Gesamtgraph zu sehen, wobei die Einflüsse von einem Zeitpunkt zum nächsten mit größerer Strichdicke gezeichnet sind.

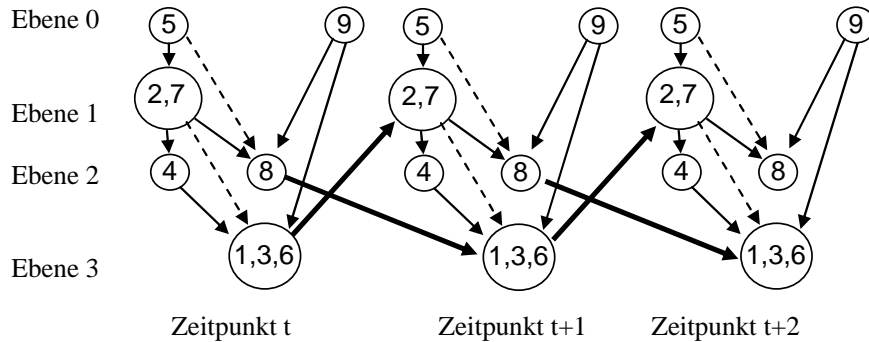


Abb. 4.16 Der Gesamtgraph aus Kopien des Graphen zu verschiedenen Zeitpunkten.

Die einzelnen Erreichbarkeitsmatrizen in (4.17) zeigen deutlich, wie die kausalen Wirkungen mit der Zeit immer weiter streuen. Es reicht also, für das Verhalten nur die ersten drei Zeitpunkte aufzuschreiben; danach ist die Streuung zu groß.

Block	C_0				C_1				C_2			
	2,7	4	8	1,3,6	2,7	4	8	1,3,6	2,7	4	8	1,3,6
5	b	i	d	i	i	i	i	i	i	i	i	i
9	0	0	b	b	i	i	i	i	i	i	i	i
2,7	0	b	b	d	i	i	i	i	i	i	i	i
4	0	0	0	b	i	i	i	i	i	i	i	i
8	0	0	0	0	0	0	0	b	i	i	i	i
1,3,6	0	0	0	0	b	i	i	i	i	i	i	i

Abb. 4.17 Erreichbarkeitsmatrizen C_0 , C_1 , C_2

4.3.3 Reduktion der Komplexität von Modellen

Man kann vor der Analyse eine elementare Vorverarbeitung der Abhängigkeiten durchführen, um die Anzahl der Variablen zu verringern.

Variablensubstitution

Folgende zwei Fälle bieten sich an:

(a) Eine unabhängige Variable i beeinflusst nur **eine** andere Variable k .

$$\begin{aligned} z_i &= f(z_1, z_2, \dots) \\ z_k &= g(z_1, \dots, z_i, \dots) \quad \Rightarrow \quad z_k = g(z_1, \dots, f(z_1, z_2, \dots), \dots) \end{aligned} \quad (4.19)$$

Dann können beide Variablen zusammengefasst werden, siehe Abb. 4.18.

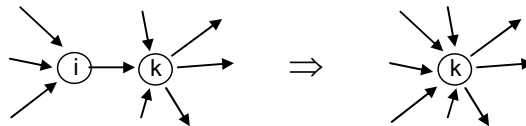


Abb. 4.18 Variablensubstitution (a)

(b) Eine Variable k wird nur von **einer** anderen Variable i beeinflusst.

$$\begin{aligned} z_i &= f(z_1, z_2, \dots) \\ z_k &= g(z_i) \quad \Rightarrow \quad z_k = g(f(z_1, z_2, \dots)) \end{aligned} \quad (4.20)$$

Dann können beide Variablen zusammengefasst werden, siehe Abb. 4.19.

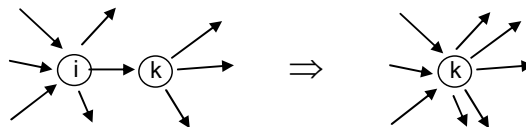


Abb. 4.19 Variablensubstitution (b)

Nebenbedingung dieser Operationen ist zum einen, dass keine Schleifen existieren dürfen (z_k als Eingang von z_i), zum anderen, dass z_i kein Eingang (ohne Vorgängerknoten) und z_k kein Ausgang (ohne Nachfolgerknoten) sein dürfen.

Die Methode wurde bei verschiedenen Modellen angewendet und erzielte dabei 30-60% Reduktion der Modellgröße!

Verschmelzung in Submodelle

Angenommen, es existiert ein Variablenpaar mit direkten kausalen Auswirkungen in beiden Richtungen. Dann kann man direkt beide Knoten in einem Block vereinigen und Mehrfachkanten nur einmal notieren, s. Abb. 4.20.

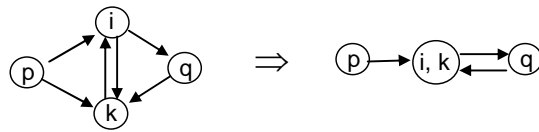


Abb. 4.20 Zusammenfassung zweier rückgekoppelter Variabler i und k

Reduktion der graphischen Repräsentation

Die Graphische Repräsentation kann auf einen Minimalgraphen ausgedünnt werden, auch wenn Rückkopplungen existieren. Dazu geht man folgendermaßen vor:

- 1) Berechne die essentiellen Rückkopplungskantenmengen
- 2) Streiche alle Kanten aus dieser Menge
- 3) Berechne den minimalen Graphen der resultierenden Ebenenstruktur
- 4) Addiere nun aus der Menge der essentiellen Rückkopplungskanten die Kanten, die Basiskanten sind, wieder zum Graphen.

Eine weitere Möglichkeit besteht darin, bei der Verbindung von Untermodellen jede Verbindung (Einfluss) nur einmal zu zeichnen und damit die Kausalität anzudeuten, siehe Abb. 4.21.

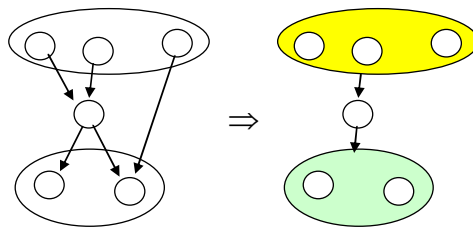


Abb. 4.21 Vereinfachung der Einflüsse bei Submodellen

5 Simulation

Der Begriff „Simulation“ ist sehr vielschichtig. Im Unterschied zur öffentlichen Wahrnehmung von Simulation als "bunte Bildchen, die nett anzusehen sind", ist die eigentliche Simulation eine wichtige Methode, Probleme zu lösen, die zu komplex für eine theoretische Behandlung sind. In diesem Kapitel soll ein Überblick über die Möglichkeiten und Probleme der Simulation gegeben werden. Wir beginnen mit der deterministischen Simulation, um den Standpunkt klarer zu machen.

5.1 Deterministische Simulation

Eine deterministischen Simulation erinnert stark an herkömmliche Programmierung: Die Simulation bedient sich normaler Programmlogik, um die Logik der Aufgabenstellung nachzuvollziehen. Dabei unterscheiden wir zwischen der diskreten Simulation, die diskrete Ereignisse in eine zeitliche oder kausale Reihenfolge bringt und die Wechselwirkungen dieser Ereignisse simuliert, und der kontinuierlichen Simulation, die vorwiegend bei dynamischen Systemen eingesetzt wird. Beginnen wir mit dem letzteren.

5.1.1 Analytische Modellierung vs. Simulation diskreter Systeme

Die Anwendungsbereiche der diskreten Simulation sind Fertigungsprozesse in Industrieunternehmen, Telekommunikationsnetzen, der Gesundheitssektor und alle Arten von Warteschlangenproblemen und Arbeits- (Last-)verteilung. Bei diesen Anwendungen haben wir komplexe Situationen mit vielen Ausnahmen, so dass eine analytische Modellierung zur Problemlösung wegen der Komplexität meist ausscheidet oder nur sehr schwer möglich ist. An einem Beispiel soll dies verdeutlicht werden.

Beispiel Lagerhaltung

In einem Betrieb wird täglich eine zunächst als konstant angenommene Menge R eines Vorproduktes benötigt und einem Lager entnommen. Ist das Lager aufgebraucht, so sollte inzwischen eine frühere Bestellung eingetroffen sein. Die neue Bestellung von B Einheiten wird aufgrund der Verbrauchszeit T dimensioniert.

Einerseits kostet jede Bestellung einen festen Bearbeitungsaufwand F , so dass eine tägliche Bestellung zu teuer wird, andererseits bindet eine Lagerung auch Kapital, so dass auch wenige große Bestellungen ungünstig sind. Es stellt sich nun die Frage: Wie viel sollte die Firma bestellen, um eine möglichst preisgünstige Lagerhaltung zu realisieren?

Bei der Lösung dieses Problems mit einem analytischen Ansatz geht man folgendermaßen vor:

- Die Gesamtkosten K teilen sich auf in Bestellkosten und Lagerkosten pro Tag, also

$$K = K_B + K_L \quad (5.1)$$

- Die Tagesmenge R hat einen Bestellkostenanteil von R/B an den Fixkosten F der Bestellung von B Einheiten, so dass pro Tagesmenge die Bestellkosten K_B

$$K_B = \frac{RF}{B}$$

anfallen.

- Die Kosten der Lagerung sollen nur von den Kapitalkosten H pro Stück abhängen. Während der Zeit T werden die Vorprodukte gleichmäßig aufgebraucht, so dass mit der Anzahl $L(t)$ der gelagerten Produkte

$$L(t) = B - Rt \quad t = 0, 1, \dots, T$$

für die Lagerkosten $K_L(t)$ pro Tag bei $L \geq 0$ folgt

$$K_L(t) = H L(t) = H (B - Rt) \quad t = 0, 1, \dots, T$$

und die mittleren Lagerkosten K_L pro Tag bis zur Neulieferung

$$\begin{aligned} K_L &= \frac{1}{T} \int_0^T H L(t) dt = \frac{1}{T} \int_0^T H (B - Rt) dt = \frac{1}{T} H (Bt - R \frac{1}{2} t^2) \Big|_0^T \\ &= \frac{1}{T} H (BT - \frac{1}{2} RT^2) = H (B - \frac{1}{2} RT) \end{aligned} \quad (5.2)$$

- Die Bestellmenge B soll bei Neulieferung aufgebraucht sein, um nicht unnötige Lagerkosten zu haben. Also ist

$$B = RT$$

Mit Gl.(5.2) folgt daraus

$$K_L = H (B - \frac{1}{2} B) = \frac{1}{2} HB$$

Die Gesamtkosten sind also pro Tag

$$K(B) = K_B + K_L = \frac{RF}{B} + \frac{1}{2}HB$$

Bei welcher Bestellmenge sind die Gesamtkosten minimal? Dazu bestimmen wir das Extremum von $K(B)$

$$\frac{dK(B^*)}{dB} = -\frac{RF}{B^{*2}} + \frac{1}{2}H = 0 \Leftrightarrow B^{*2}\frac{1}{2}H - RF = 0$$

oder

$$B^* = \sqrt{\frac{2RF}{H}} \tag{5.3}$$

mit den Kosten

$$\begin{aligned} K(B^*) &= \frac{RF}{B^*} + \frac{1}{2}HB^* = \frac{B^*RF}{B^{*2}} + \frac{1}{2}HB^* = \frac{1}{2}HB^* + \frac{1}{2}HB^* \\ &= HB^* = \sqrt{2RFH} \end{aligned} \tag{5.4}$$

Damit haben wir direkt die optimale Bestellmenge mit den minimalen Kosten bestimmt. Natürlich lässt sich diese Lösung auch erreichen, wenn wir für festes B ein Simulationssystem schreiben, bei dem der Verbrauch R pro Tag simuliert wird und dabei Bestellkosten und Kapitalbindung mit einbeziehen und die auflaufenden Kosten über mehrere Bestellungen ermitteln. Variieren wir dann die Vorgabe B , so können wir verschiedene Kosten $K(B)$ feststellen. und dann mit einer diskreten Iteration, etwa der Newton-Raphson-Iteration aus Abschnitt 3.5, die optimale Bestellmenge finden. Dazu wird der Parameter B^* gesucht, der minimale Kosten verursacht. Dies entspricht der Nullstelle der Ableitung von K , so dass wir für die Zielfunktion von Abschnitt 3.5 $R(\theta) = K(B)$ setzen können. Für diesen Fall lautet die Iteration dann

$$\begin{aligned} B_{t+1} &= B_t - \gamma \frac{K(B_{t-1}) - K(B_t)}{K'(B_{t-1}) - K'(B_t)} \\ \text{mit } K'(B_{t-1}) &= \frac{K(B_{t-2}) - K(B_{t-1})}{B_{t-2} - B_{t-1}}, K'(B_t) = \frac{K(B_{t-1}) - K(B_t)}{B_{t-1} - B_t} \end{aligned} \tag{5.5}$$

Nun fragen wir uns, wozu wir ein aufwendiges Simulationsprogramm schreiben sollen, wenn wir B^* und K^* direkt analytisch ermitteln können? Die Antwort kommt aus der Praxis: Die analytische Lösung gilt nur, wenn alle Annahmen über Verbrauch R , Finanzkosten H und Bearbeitungskosten F zutreffen. Insbesondere müssen sie konstant sein, um zur analytischen Lösung zu gelangen. In der Praxis sind aber meist weder T , noch R oder H konstant, sondern ändern sich häufig im Betrieb, je nach Konjunktur, Preisen und Produktionsorganisation. Die diskreten Kosten sind deshalb laufend anders. Aus diesem Grund bietet es sich an, nicht nur die Bestellmengen für die Vielzahl der benötigten Vorprodukte einzeln

zu optimieren, sondern dies durch ein Simulationsprogramm mit allen jeweils gültigen Parameterwerten durchführen zu lassen.

5.1.2 Zeitkontinuierliche vs. zeitdiskrete Modellierung und Simulation

In Kapitel 3 sahen wir, wie sich das zeitliche Verhalten komplexer Systeme modellieren lässt. In Abb. 5.1 ist der zeitliche Verlauf der Variablen eines rückgekoppelten Modells visualisiert, das durch Absenken der Schwelle in einen schwingenden Zustand gerät. Dabei wird angenommen, dass eine Eingabevariable x einen Zustand z_1 so ändert, dass damit eine Oszillation des Zustands z_2 auslöst, deren Wert die Ausgabe y des Systems bildet.

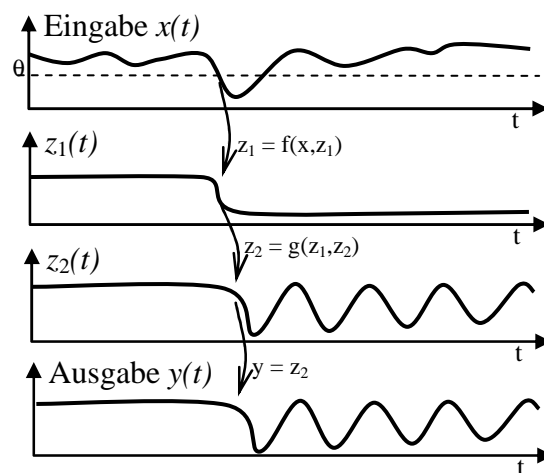


Abb. 5.1 Zeitlicher Ablauf von kausalen Eingaben

Die Darstellung geht davon aus, dass wir ein zeitkontinuierliches System haben, etwa ein chemisches, biologisches oder ökonomisches System. Auch für analoge Computersysteme gilt diese Annahme. Die übliche Simulation geht aber davon aus, dass die Variablen x durch Zahlen mit endlicher Präzision numerisch in einem digitalen Rechner repräsentiert werden (**Wert-Diskretisierung**). Jede Zahl wird dabei binär durch die n Koeffizienten $b_i \in \{0,1\}$ repräsentiert:

$$x = \sum_{i=0}^{n-1} b_i 2^i \quad (5.6)$$

Auch die zeitliche Entwicklung kann nur an diskreten Zeitpunkten beobachtet werden, etwa an äquidistanten Zeitpunkten $t_0 + k\Delta t$ mit der Abtastfrequenz (*Samplingfrequenz*) $f_0 = 1/\Delta t$ (**Zeit-Diskretisierung**) durch einen Analog-

Digitalkonverter ADC. Die Situation aus Abb. 5.1 ändert sich damit in die Situation in Abb. 5.2.

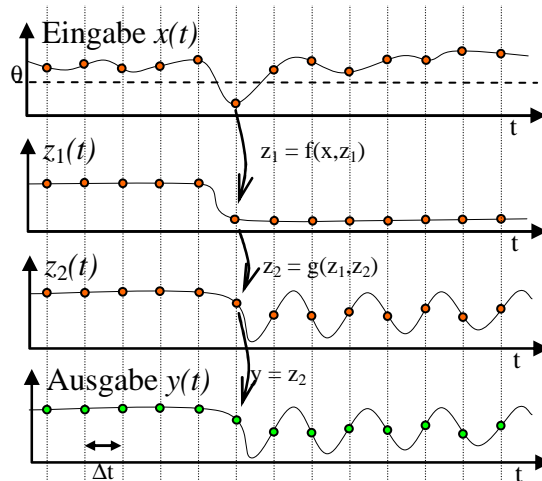


Abb. 5.2 Diskretisierung in der Zeit.

Ein erstes Problem bei dieser Abbildung kontinuierlicher Vorgänge auf diskrete Zeitpunkte tritt auf, wenn wir für alle Variablen einen gemeinsamen Zeitabstand Δt bestimmen wollen: Welchen sollen wir wählen? Dazu müssen wir die Nebeneffekte berücksichtigen, die eine zeitliche Diskretisierung mit sich bringt. Betrachten wir dazu $y(t)$ in Abb. 5.2. Wie wir sehen, wird durch das Abtasten die Oszillation nicht korrekt wiedergegeben; die Ursache liegt darin, dass die Oszillationsfrequenz höher ist als die Abtastfrequenz f_0 . Allgemein gilt nach dem Nyquist-Theorem, dass bei einer Abtastfrequenz f_0 nur die Anteile eines Ereignisses beobachtet werden können, die Frequenzen kleiner als $f_0/2$ haben. Für unseren Zeitabstand Δt heißt das, dass bei Variablen i , die sich in einem Zeitabschnitt Δt_i signifikant ändern können, ein Abtastabstand von maximal $\Delta t = \Delta t_i/2$ verwendet werden muss, um keine Information zu verlieren. Dies gilt für jede Variable i für sich, so dass man für jede Variable ein eigenes Abtastintervall (*Samplingintervall*) definieren kann.

Der **Vorteil** eines solchen Systems liegt in der minimierten Anzahl von Messwerten, die gespeichert und verarbeitet werden müssen. Dies bedeutet neben der geringeren Größe der Speicherplätze bzw. Datenbanken und der geringeren CPU-Belastung aber auch weniger Aufwand. Bei medizinischen Messwerten bedeutet das sowohl weniger Geld für Labortests bzw. Personal als auch weniger Belastung für die Patienten durch unnötige Messprozeduren. Diese Überlegung ist in der klinischen Praxis üblich: die Anzahl der Messvariablen und der zeitliche Abstand

der Messungen richtet sich nach dem Zustand der Patienten und ändert sich kontinuierlich.

Allerdings bringt ein solches System auch **Nachteile** mit sich. Besonders folgende Fragen müssen gelöst werden:

- Die Messzeitpunkte sind nicht identisch. Wenn wir mehrere Variable haben, welche sollen wir als „zeitgleich“ ansehen und verwenden? Wie sollen wir eine Diagnose von mehreren Werten durchführen? Dies ist ein Problem temporaler Datenbanken und ist in relationalen Datenbanken nur schwer lösbar. Spezielle temporale Erweiterungen der relationalen Technik helfen hier weiter.
- Wenn ein gemeinsamer Messzeitpunkt nicht gefunden wird und wir uns dazu entschließen, die vorhandenen Messwerte zu interpolieren, so müssen wir Messwerte neu konstruieren. Dies ist als „Resampling“ bekannt.
- „Missing values“: Manchmal fehlen auch einfach und schlicht Werte, weil aus obigen Gründen nicht gemessen wurde: Wir müssen Werte ergänzen.

In den beiden letzten Fällen können wir die Werte durch Interpolation erschließen. Betrachten wir die zeitliche Folge der Messwerte als eine funktionale Abhängigkeit, so können wir die Messwerte im einfachsten Fall durch lineare Interpolation, sonst durch Polynome f wie in Kapitel 2 approximieren. Die analytisch „glatte“ Funktion $f(\cdot)$ stellt eine Mittelung der zufällig davon abweichenden Werte dar. Beim *Resampling* können wir nun für jeden beliebigen Zeitpunkt einen Wert $f(t)$ ablesen und notieren. Die resultierenden, rekonstruierten Messwerte sind damit zeitsynchron zu den Samples der anderen Variablen.

Leider aber sind sie in dieser Form nicht brauchbar: Die neuen, „gemessenen“ Werte sind ja keine verrauschten Werte, sondern deren Erwartungswerte. Dies würde aber künstlich Information (über den Erwartungswert) in die Messwerte einführen. Auch die Benutzung des Median-Wertes (die Nutzung des $(n+1)/2$ -ten Messwerts aller n Abweichungen, aufsteigend geordnet) bringt ein solches Problem mit sich.

Stattdessen müssen wir die fehlenden Werte so ergänzen, dass die Zufallsproportionen der Messwerte gewahrt bleiben. Kennen wir die Verteilung der Abweichung nicht, so wird häufig angenommen, dass eine Normalverteilung um den Erwartungswert mit der gemessenen Standardabweichung σ vorliegt. Es reicht dann, einen Zufallswert aus einer Normalverteilung zu ziehen und auf den Erwartungswert $f(t)$ aufzuaddieren. Genauer ist es aber, die gemessene Statistik der Variablen zu verwenden, also einen zufälligen Wert aus allen gemessenen Abweichungen zu ziehen und an dieser Stelle einzusetzen.

Beispiel *Medizin. Diagnostik bei sept. Schock*

Bei der Mortalitätsprognose für septischen Schock auf der Intensivstation spielen systolischer und diastolischer Blutdruck sowie die Anzahl der Thrombozyten im Blut eine Rolle. Eine gute Diagnose lässt sich als eine Funktion aus die-

sen drei Werten formulieren. Leider aber werden beide Blutdruckwerte in der klinischen Praxis sehr oft am Tag gemessen, manchmal sogar kontinuierlich; die Thrombozytenzahl dagegen nur einmal pro Tag. Eine derartige Diagnose ist also retrospektiv im Normalfall nur einmal pro Tag möglich. Nun ist aber bekannt, dass sich zugrunde liegende Sepsisdynamik sehr schnell entwickeln kann, innerhalb von Stunden. Die gesamte Sepsisdynamik muss also mit nur einem einzigen Sample in 24 Stunden beurteilt werden! Es ist klar, dass dies nicht ausreichend ist: Der zeitliche Abstand zwischen den Samples sollte bei Verdacht auf septischen Schock für Thrombozyten verkleinert werden; es lassen dabei sich immer Blutdruckwerte so finden, dass alle Messwerte zeitsynchron sind.

Dieses Beispiel zeigt schon die Hauptproblematik beim Resampling und Messwert-synchronisation: Eine rein formale Messwertbehandlung muss immer kritisch vor dem Hintergrund der Annahmen betrachtet werden; meist sind die Annahmen das schwache Glied in der Argumentationskette. Für ein gutes Resampling müssen auch tragfähige Aussagen über die Dynamik und Varianz der Messvariablen vorliegen.

5.1.3 Ereignisorientierte Modellierung und Simulation

Bei der zeitdiskreten Modellierung ersetzen wir kontinuierliche Messwerte durch zeitdiskrete Messungen. Haben wir mehrere diskrete Messwerte an diskreten Zeitpunkten, so können wir die Ursachen und Wirkungen auf einer einzigen Zeitachse anordnen und das ganze System als einen Ablauf von diskreten Einzelereignissen modellieren. Anstelle von n Variablen auf n parallelen Zeitachsen erhalten wir n Variablen auf einer Zeitachse. In Abb. 5.3 ist ein solches System für drei Variable x_1 , x_2 , x_3 gezeigt.

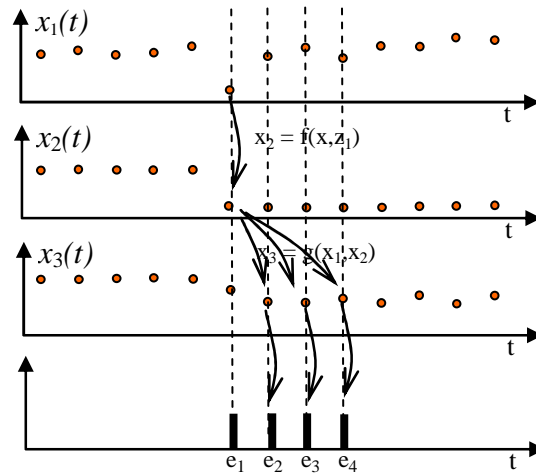


Abb. 5.3 Ereignisanordnung in der Zeit

Die Einflüsse lassen sich durch eine Kausalkette mit den Ereignissen e_1 bis e_4 modellieren. Dabei tritt in den Hintergrund, dass die ursprünglichen Messereignisse an diskreten Zeitpunkten stattgefunden haben; die Zeit zwischen den Ereignissen kann ignoriert werden und ist nur noch ein logisches Attribut. Die Zeit bedeutet hier nur noch eine Ordnungsrelation, ein Index der Einzelereignisse, und sagt uns, in welcher Reihenfolge die Ereignisse abgearbeitet werden müssen.

Beispiel *Simulation eines Supermarkts*

Wir möchten die Kundenströme in einem Supermarkt simulieren. Dazu erstellen wir uns zuerst ein Modell, das die wesentlichen Eigenschaften des Markts enthält. Angenommen, der Markt hat einen Eingang, einen großen Raum mit Warenregalen und vor dem Ausgang eine Kasse. In Abb. 5.4 ist dies gezeichnet.

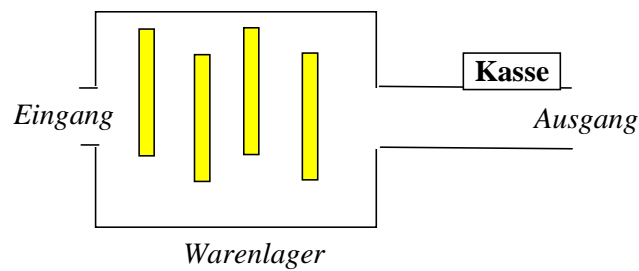


Abb. 5.4 Modell des Supermarkts

Dabei modellieren wir mit folgenden Variablen

- Eingabe* $x = \{\Phi, a, b, c, \dots\}$ die Kunden am Eingang, wobei $x = \Phi$ = „kein Kunde“ und $x = a$ = „Kunde a betritt den Laden“ bedeutet.
- Einkaufszeit* $T_{\alpha L}$ des Kunden $\alpha = a, b, \dots$. Dies entspricht der Aufenthaltszeit im Warenlager
- Kasse mit der Länge z_w der *Warteschlange*, wobei die Reihenfolge der Kunden vor der Kasse in FIFO-Ordnung (*First In - First Out*) erfolgt, und der Zeit $T_{\alpha w}$, die der Kunde α an der Kasse braucht, um abgerechnet zu werden.
- Ausgabe* $y = \{\Phi, a, b, c, \dots\}$ die Kunden $\alpha = a, b, c, \dots$ am Ausgang, wobei $x = \Phi$ bedeutet: "Kein Kunde verlässt den Laden".

Betrachten wir einen typischen Ablauf. Angenommen, Kunde a kommt in den Laden, bedient sich und wird danach an der Kasse abgefertigt. Dies ist in Abb. 5.5 visualisiert.

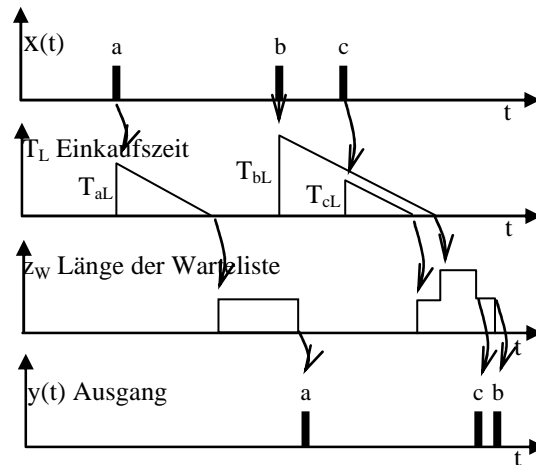


Abb. 5.5 Kausalkette der Kundenbedienung

Die Einkaufszeit ist als Funktion T_{aL} - t notiert, so dass sie durch einen Zeitähler modelliert ist, der bis auf null abläuft. Nach dem Ablauf wird Kunde a in die Kassenwarteschlange eingegliedert, bearbeitet und nach Ablauf der Kassenzeit am Ausgang registriert. Bei dem nachfolgenden Kunden b und c ändert sich die lineare Reihenfolge: b kauft länger ein als c und kommt damit in die Warteschlange an der Kasse hinter c . Die Länge der Schlange erhöht sich um eins, verringert sich nach Bedienung von c um eins und geht nach Bedienung von b wieder auf null. Nach dem Verlassen der Kasse erfolgt der Ausgang, zuerst von c , dann von b .

An dem Beispiel sehen wir nicht nur die Kausalkette der Ereignisse wie bei Kunde a , sondern auch, wie ein Ereignis (Kunde c) ein anderes Ereignis (Kunde b) überholen kann. Dieser nicht-monotone Ereignisablauf zeichnet eine Simulation aus, die auch asynchron parallele Handlungsabläufe sauber abbildet.

Manche Programmiersprachen (wie SIMULA etc.) unterstützen diesen Prozess, indem eine Zeitleiste definiert wird, in die einzelne Ereignisse eingehängt werden können. In der Sprache Java beispielsweise lässt sich dies durch eine Warteschlange modellieren (Datentyp „List“), in die bedarfsweise an beliebiger Stelle neue Elemente eingefügt werden können, wenn neue Ereignisse für die Eingabe bzw. Ausgabe generiert werden. Bei der Simulation wird dann nur noch sequentiell auf diese Ereigniskette zugegriffen und abgearbeitet.

Technisch gesehen muss die Kausalkette auf die Zeitachse abgebildet werden. Dazu werden Datenstrukturen aufgebaut, die nach der Ereignisgenerierung ein schnelles Auffinden der Listenelemente mit der richtigen Zeitkennung und das Einfügen des auszuführenden Elements ermöglichen. Eine Idee besteht darin, für jedes Ereignis, das in der Zukunft zum Zeitpunkt T_i stattfinden soll, einen Zähler

einzurichten und ihn initial auf die Anzahl der Zeitticks zu setzen, die bis T_i auftreten werden. Alle Ereignisse (Objekte) hängen in einer ungeordneten Warteschlange, siehe Abb. 5.6.

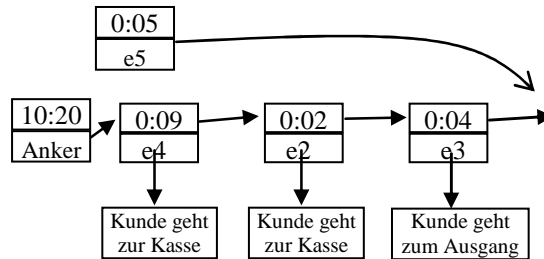


Abb. 5.6 Datenstruktur für die Ereignisanordnung in der Zeit

Bei jedem Zeittick geht der Prozessor durch alle Elemente durch und dekrementiert sie jeweils um eins. Jedes Element, bei dem der Zähler eine Null erhält, ist zur Ausführung vorgesehen und wird nach der Ausführung der angehängten Methode gelöscht. Vorher erzeugt es aber vielleicht weitere Ereignisse, die zur Erledigung in die Schlange gehängt werden. Eine solche Datenstruktur hat den Vorteil, dass sie sehr schnell aufgebaut werden kann und jedes neues Ereignis sofort zur Verfügung steht. Der Nachteil ist aber auch klar: bei jedem Zeittick wird nicht nur die Zeit im Anker erhöht, sondern man muss auch die Liste durchgehen und die Zeit von jedem Ereignis um eins verkleinern.

Ein schnelleres Verfahren ist in Abb. 5.7 möglich. Hier ist eine geordnete Liste dargestellt, bei der ein neues Ereignis e5 für 10.25 Uhr eingehängt wird.

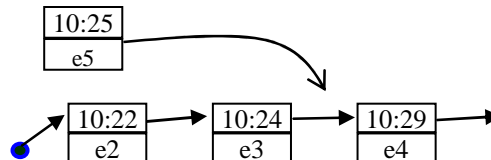


Abb. 5.7 Datenstruktur für die Ereignisanordnung in der Zeit

Alle Listeneinträge sind bereits geordnet und bleiben geordnet, so dass es ausreicht, das erste Element auf die Zeit zu überprüfen. Ist das Ereignis nicht fällig, so sind es auch alle anderen nicht in der Liste.

Die Struktur ist zwar schnell beim jedem Zeittick, aber sie hat einen Nachteil: bei der Einsortierung eines Ereignisses ist sie langsam. Eine solche Struktur lässt sich

aber dadurch effizienter machen, dass man sie in einen Baum einbaut, in dem die einzelnen Elemente in $O(\log n)$ -Zeit zugreifbar sind anstatt linear.

5.2 Stochastische Simulation

In unserem Beispiel der Wirtschaftssimulation in Abschnitt 5.1.1 gingen wir davon aus, dass die Zahl der Vorprodukte, die Lieferzeit usw. konstant sind. Dies ist aber nicht realistisch. Stattdessen sind die Mengen und Zeiten nur Erwartungswerte für zufallsbehaftete Werte, die stark schwanken können und deshalb in der diskreten Simulation auch als Zufallsvariablen modelliert sein müssen. Eine solche Modellierung muss die typischen Eigenschaften der Variablen widerspiegeln, insbesondere ihre Wahrscheinlichkeitsverteilung. Wir betrachten deshalb zunächst die typischen Eigenschaften von Zufallsvariablen und ihre Erzeugung.

5.2.1 Eigenschaften von Zufallszahlen

Als „Zufallszahlen“ bezeichnen wir Variable, die ihren Wert zufällig annehmen können. Dabei ist nicht der Wert zufällig, sondern das Ereignis, einen (evtl. vordefinierten, festen) Wert anzunehmen. Angenommen, wir haben eine Funktion $P(w)$ vorgegeben, welche die Wahrscheinlichkeit ausdrückt, mit der ein auftretender Wert x kleiner oder gleich einem Wert w ist:

$$P(w) = \text{Prob}(x \mid x \leq w) \quad (5.7)$$

Sie wird **Verteilungsfunktion** der Zufallsvariablen X (*cumulative distribution function* cdf) genannt. Sie charakterisiert vollkommen die Verteilung. Sie kann auch als Integral einer anderen Funktion $p(x)$ geschrieben werden

$$P(w) = \int_{-\infty}^w p(x) dx \quad \text{Verteilungsfunktion} \quad (5.8)$$

der **Wahrscheinlichkeitsdichte** $p(x)$ (*probability density function pdf*) $p(x) \geq 0$. Es ist klar, dass bei $w = \infty$ der Wert $P(\infty) = 1$ resultiert: jeder endliche Wert w ist kleiner als Unendlich. Dies bedeutet

$$P(\infty) = \int_{-\infty}^{+\infty} p(x) dx = 1 \quad (5.9)$$

Die einfachsten Arten von Zufallsvariablen sind Variablen mit einer konstanten, uniformen Verteilung. In Abb. 5.8 ist die gezeigt. Sie ist hier von $x = 0$ bis $x = 1$ konstant eins und geht ab $x = 1$ auf null zurück.

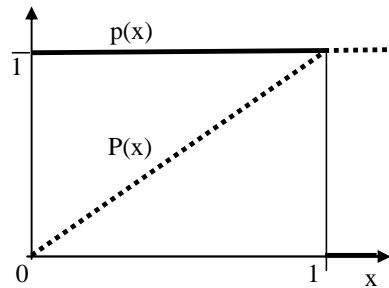


Abb. 5.8 Die Wahrscheinlichkeitsdichte einer uniformen Verteilung

Formal kann man sie definieren als

$$p(x) = \begin{cases} 1 & \text{für } 0 \leq x \leq 1 \\ 0 & \text{sonst} \end{cases} \quad (5.10)$$

wobei die Gesamtwahrscheinlichkeit $P(0 \leq x \leq 1)$ das Integral ist, die Fläche unter der Kurve $p(x)$. Sie wird auch als „Wahrscheinlichkeitsmasse“ bezeichnet.

$$P = \int_0^1 p(x) dx = x \Big|_0^1 = 1 \quad (5.11)$$

Alle Ereignisse geschehen wie gewünscht im Intervall von null bis eins, was für unsere uniforme Verteilung bedeutet

$$P(w) = \int_0^w 1 dx = x \Big|_0^w = w \quad (5.12)$$

eine lineare Funktion in w , die ab $x = 1$ konstant bleibt. Sie ist in Abb. 5.8 gestrichelt eingezeichnet. Für die uniforme Verteilung im Intervall $[x_0, x_1]$ kürzen wir sie mit $U(x_0, x_1)$ ab. Erwartungswert und Streuung ergeben sich dann direkt aus den Intervallgrenzen.

Wie errechnen wir allgemein Erwartungswert und Streuung einer Verteilung in einem Intervall?

Der **Erwartungswert** μ einer Zufallsvariablen im Intervall $[x_0, x_1]$ ist definiert

$$\mu = \frac{\int_{x_0}^{x_1} x \cdot p(x) dx}{\int_{x_0}^{x_1} p(x) dx} \quad (5.13)$$

Dabei dient der Nenner zur Normierung auf die Wahrscheinlichkeit des Intervalls. Für die uniforme Verteilung im Intervall $[x_0, x_1]$ gilt allgemein mit konstanter Dichte $p(x) = c$

$$\mu = \int_{x_0}^{x_1} x \cdot c \, dx \bigg/ \int_{x_0}^{x_1} c \, dx = c \frac{x^2}{2} \bigg|_{x_0}^{x_1} \bigg/ c x \bigg|_{x_0}^{x_1} = \frac{1}{2} \frac{(x_1^2 - x_0^2)}{x_1 - x_0} \quad (5.14)$$

Für das Intervall $[0,1]$ ist also $\mu = 1/2$.

Die **Varianz** σ^2 in einem Intervall $[x_0, x_1]$ ist definiert als

$$\sigma^2 = \int_{x_0}^{x_1} (x - \mu)^2 p(x) dx \bigg/ \int_{x_0}^{x_1} p(x) dx \quad (5.15)$$

und damit für unsere uniforme Verteilung

$$\begin{aligned} \sigma^2 &= \int_{x_0}^{x_1} (x - \mu)^2 c dx \bigg/ \int_{x_0}^{x_1} c dx = c \frac{(x - \mu)^3}{3} \bigg|_{x_0}^{x_1} \bigg/ c x \bigg|_{x_0}^{x_1} \\ &= \frac{(x_1 - \mu)^3 - (x_0 - \mu)^3}{3(x_1 - x_0)} \end{aligned} \quad (5.16)$$

Für das Intervall $[0,1]$ ergibt sich so zu

$$\sigma^2 = \frac{(1/2)^3 - (-1/2)^3}{3} = \frac{2/8}{3} = \frac{1}{12}$$

Aufgabe 5. 1

Angenommen, die Wahrscheinlichkeitsdichte p ist im Intervall $[0,d]$ konstant mit $p(x) = c$. Welchen Erwartungswert μ und Standardabweichung σ hat sie?

5.2.2 Messen von Zufallszahlen

Wie erhalten wir die Funktionen $P(\cdot)$ oder $p(\cdot)$ einer Zufallsvariablen? Die Dichtefunktion $p(x)$ ist der Grenzwert eines Histogramms, das dadurch entsteht, dass man den Wertebereich der Zufallsvariablen x in Intervalle Δx einteilt und für jedes Intervall zählt, wie oft die Zufallsvariable darin beobachtet wurde. In Abb. 5.9 ist ein Beispiel zu sehen. Für eine sehr große Anzahl beobachteter x und $\Delta x \rightarrow 0$ erhalten wir dann die Dichte $p(x)$ der Zufallsvariablen. In jedem Abschnitt i gilt dann

$$p_i = \frac{\text{Anzahl}_i}{\sum_k \text{Anzahl}_k} \quad (5.17)$$

Wie erstellen wir uns im Detail ein solches Histogramm?

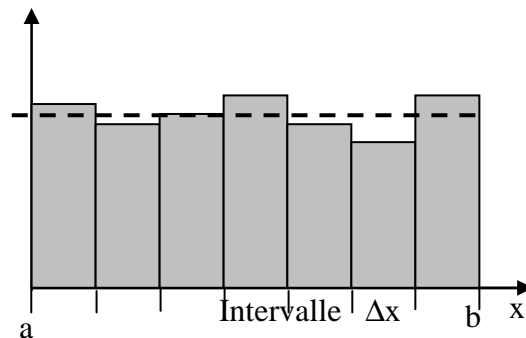


Abb. 5.9 Histogramm einer Beobachtung einer uniformen Verteilung

Dazu legen wir zuerst die Anzahl N der Intervalle fest und damit die Intervallbreite $\Delta x = (b-a)/N$. Dann deklarieren wir ein Feld $A[]$ aus N Elementen $A[0]$, $A[1]$, ..., $A[N-1]$. Nun lesen wir die Zufallszahlen ein. Bei jedem Aufruf einer Zufallszahl wird nun die Anzahl der Elemente im richtigen Intervall mit dem Index i um eins inkrementiert nach der Formel

$$i := \left\lfloor \frac{x-a}{\Delta x} \right\rfloor, \quad A[i] = A[i]+1 \quad (5.18)$$

wobei das Abschneiden der Nachkommastellen durch eine Integer-Division erreicht werden kann. Diese Operation wird n -mal durchgeführt, wobei aus praktischen Erfahrungen sinnvollerweise $n > 10N$ gewählt werden sollte. Anschließend wird die Anzahl der Ereignisse $A[i]$ für jedes Intervall mit der Gesamtzahl n auf die relative Häufigkeit p_i pro Intervall normiert:

$$p_i = \frac{A[i]}{n} \quad (5.19)$$

Man beachte, dass die Erstellung von mehrdimensionalen pdf von Verbund-Zufallszahlen (Zufallsvektoren (x_1, x_2, \dots, x_k)) problematisch wird, je mehr Dimensionen beteiligt sind. Setzt man z.B. 10 Intervalle pro Zufallsvariable an, so benötigt man nach obiger Heuristik mehr als $n = 100$ Beobachtungen, um sie zu füllen. Betrachtet man aber die multidimensionale pdf von 10 Variablen, so ist die Auflösung zwar immer noch 10 Intervalle pro Dimension, aber wir haben 10^{10} Intervalle zu füllen mit 10^{11} Beobachtungen – in der Praxis meist nicht möglich. Solche hochdimensionalen pdf, die zur numerischen Berechnung der Entropie u.d.gl. nötig sind, können wir nur mit anderen Methoden (oder gar nicht) erhalten.

5.2.3 Erzeugung und Simulation von Zufallszahlen

Zufallszahlen werden meist als Folge generiert, bei der jede Zahl aus ihrem Vorgänger erzeugt wird. Dazu werden an die Zufallszahlen einige Forderungen gestellt:

- *Unabhängigkeit*
Jede erzeugte Zufallszahl sollte unabhängig sein von der vorigen. Dies sollte durch Tests sicher gestellt werden.
- *konstante Dichte*
Für eine uniforme Verteilung müssen alle Zahlen aus dem Intervall $[0,1]$ eine konstante Dichte besitzen; eine Bevorzugung bestimmter Bereiche ist nicht zulässig.
- *hohe Besetzungsdichte*
Ein weiteres Kriterium ist die Anzahl der in einem Teilintervall vorhandenen Zahlen: die Besetzungsdichte. Idealerweise sollten keine Lücken zwischen den potentiellen Zahlen existieren, aber da wir nur mit endlicher Genauigkeit zählen (Werte-Diskretisierung), ist dies nicht möglich. Trotzdem sollten die Lücken gleichmäßig im Intervall verteilt sein.
- *Reproduzierbarkeit*
Obwohl jede Zahl von der anderen unabhängig sein soll, ist es sehr praktisch, wenn man eine Folge wiederholen kann, um verschiedene Algorithmen bei gleichen Bedingungen (z.B. gleicher Folge von zufälligen Ereignissen) miteinander vergleichen zu können. Muss man die ganze Folge nicht abspeichern, sondern kann sie bei Eingabe des Startwerts wiederholen lassen, so spart man Speicherplatz.

Früher wurden Hardwareeffekte verwendet, um Zufallszahlen zu erzeugen. Beispielsweise wurde das weiße, thermische Rauschen (zufällige Spannungsschwankungen an Bauelementen) verstärkt und einer Analog-Digital-Wandlung zugeführt. Die beobachteten Zahlen werden allen Forderungen gerecht, bis auf die letzte: Sie waren leider nicht reproduzierbar. Aus diesem Grund setzten sich die mit weniger Aufwand algorithmisch erzeugten Zufallszahlen durch, die reproduzierbar sind und keine spezielle Hardware benötigen. Allerdings ist hier zu fragen: können deterministische Algorithmen „echte“ Zufallszahlen, also zufällig erzeugte Zahlen, hervorbringen? Die Antwort lautet „Nein“: nach der Definition werden diese Zufallszahlen nicht zufällig erzeugt. Man spricht deshalb eher von „Pseudo-Zufallszahlen“, da diese Zahlen die Eigenschaften „echter“ Zufallszahlen nur approximieren.

Die ersten algorithmischen Verfahren (von Neumann, Metropolis 1940) versuchten, mit einem willkürlichen Prozess zufällige Zahlen zu erzeugen. Dazu wurde eine vierstellige Zahl, etwa 7.182, quadriert. Die so erzeugten Zahlen werden, falls nötig, auf 8 Stellen mit Nullen aufgefüllt. Von der so erzeugten Zahl 51.581.124 werden die mittleren vier Stellen herauskopiert und als neue Zufallszahl verwendet:

$$\begin{array}{l}
 x_t = 7\ 1\ 8\ 2 \rightarrow 5\ 1\ 5\ 8\ 1\ 1\ 2\ 4 \\
 \quad \quad \quad \searrow \quad \quad \quad \swarrow \\
 x_{t+1} = 5\ 8\ 1\ 1 \rightarrow 3\ 3\ 7\ 6\ 7\ 7\ 2\ 1 \\
 \quad \quad \quad \searrow \quad \quad \quad \swarrow \\
 x_{t+2} = 7\ 6\ 7\ 7 \rightarrow \dots
 \end{array}$$

Dieser Zufallszahlengenerator zeigt sich bald als nicht sehr „zufällig“: Er tendiert dazu, bei kleinen Zahlen viele Nullen weiterzugeben und dann zu Null zu konvergieren.

In den meisten Bibliotheken wird deshalb zur Erzeugung von Zufallszahlen die Methode der „linearen Kongruenzen“ verwendet, bei der durch eine lineare Operation eine Zahl aus einer Menge wieder auf ein Element der selben Menge abgebildet wird. Starten wir mit einem Startwert x_0 , so ist der nachfolgende Wert mit $t = 1, 2, 3, \dots$ durch

$$x_{t+1} = (ax_t + b) \bmod m \quad \text{linearer Kongruenter Generator LCG} \quad (5.20)$$

gegeben. Hierbei ist **mod** die Modulo-Operation mit den positiv-ganzzahligen m , n definiert mit

$$n \bmod m := n - \left\lfloor \frac{n}{m} \right\rfloor m \quad (5.21)$$

Sie gibt den Rest der ganzen Zahl n an, die durch die ganze Zahl m geteilt wird. Beispielsweise ist $16 \bmod 5 = 1$, da $16/5 = 3$ Rest 1. Also wird die Zahl x_t aus dem Intervall $[0 \dots m-1]$ auf x_{t+1} abgebildet, das ebenfalls aus dem Intervall $[0 \dots m-1]$ ist. Allerdings hat der Algorithmus einige Schwächen, die wir kennen sollten:

- Die *Unabhängigkeit* der Zufallszahlen ist nicht gewährleistet, wenn wir sie parallel benutzen. Benötigen wir zwei unabhängige Zufallszahlenreihen, so können wir nicht denselben Zufallszahlengenerator zweimal starten: Wir erhalten dieselben Zahlenreihen, nur an anderer Stelle startend und damit zeitverzögert. Eine Möglichkeit, dies zu vermeiden, besteht darin, die letzte von einer Simulation benutzte Zufallszahl zu speichern und als Saat (Startwert) der nächsten Simulation zu verwenden. Ist die Periode lang genug, so sollten die beiden Zufallszahlenreihen unabhängig sein.

Wenn dies nicht möglich ist, weil beispielsweise zwei parallel ablaufende Simulationen bedient werden müssen, so besteht eine andere Idee darin, für jede Simulation einen extra Generator zu bauen. Dies kann entweder ein Typ mit zwei verschiedenen Parametersätzen sein, oder auf zwei (oder mehreren) völlig unterschiedlichen Konzepten aufbauende Generatoren. Ein zusätzlicher, zentraler Generator entscheidet dann mit seinen Zufallszahlen, von welchem der anderen Generatoren jeweils die Zufallszahl genommen werden soll.

- *Reelle Werte*

Durch die Verwendung ganz-zahliger Werte x_t erreichen wir auch nur ganzzahlige Werte, in diesem Fall maximal m Stück. Wollen wir die Menge $\{0, 1, \dots, m-1\}$ auf das Intervall $[0,1[$ abbilden, so müssen wir *die relativen Werte* bilden, so dass wir die reelle Zahl \tilde{x}

$$\tilde{x}_t = \frac{x_t}{m} \Rightarrow \tilde{x}_t \in \left[0, \frac{m-1}{m}\right] \quad (5.22)$$

erhalten. Dabei muss m sehr groß gewählt werden, üblicherweise eine 2-Potenz $m = 2^k$.

- Ein weiteres Problem ist die *Besetzungsdichte*. Setzen wir z.B. $a = 4$ und $b = 0$, $m = 5$, so ist die Zufallsfolge mit dem Startwert $x_0 = 4$

$$\begin{aligned}x_1 &= (4 \cdot 4 + 0) \bmod 5 = 1 \\x_2 &= (4 \cdot 1 + 0) \bmod 5 = 4 \\x_3 &= (4 \cdot 4 + 0) \bmod 5 = 1\end{aligned}$$

usw. Die Folge 4,1,4,1,... wiederholt sich laufend; die möglichen Zahlen 2,3,5, werden nicht angenommen. Zwar wird bei jeder sequentiellen Methode bei endlich vielen unterschiedlichen Zahlen immer einmal eine Zahl sich wiederholen und damit einen Zyklus erzeugen, aber bei einer ungünstigen Wahl von a , b und m geschieht dies zu früh.: Es werden nicht alle potentiellen Zahlen ausgenutzt; die Besetzungsdichte ist gering.

Es lässt sich zeigen, dass folgende notwendige und hinreichende Bedingungen für eine volle Periode gelten:

- der größte gemeinsame Teiler von b und m ist 1, es gibt also keinen nicht-trivialen gemeinsamen Teiler
- wenn eine Primzahl q in m enthalten ist, dann auch in $a-1$
- wenn 4 in m enthalten ist, dann auch in $a-1$

Beispiel: $a = 5$, $b = 25541$, $m = 7411$ für einen 16-Bit-Rechner, wobei b , m Primzahlen sind.

- *Schnelle Implementierung*

Mit der Darstellung einer reellen Zahl \tilde{x} als Binärzahl von n Bits

$$\tilde{x} = \sum_{i=0}^{n-1} b_i 2^i \quad (5.23)$$

bedeutet dann eine Division durch m zur Modulo-Bildung

$$\frac{\tilde{x}}{m} = \frac{\tilde{x}}{2^k} = \frac{1}{2^k} \sum_{i=0}^{n-1} b_i 2^i = \sum_{i=0}^{n-1} b_i 2^{i-k} \quad (5.24)$$

Der Index i jeder Zweierpotenz wird um k vermindert, das binäre Speicherfeld also um k Stellen nach rechts verschoben und alle Stellen hinter dem Komma fallen weg.

Die anschließende Multiplikation mit m bei der Modulo-Bildung verschiebt wieder alles um k Stellen nach links, so dass als Resultat die k Binärstellen mit kleinstem Index gelöscht sind. Bei der Differenz zu n werden deshalb nur die oberen $n-k$ Stellen gelöscht; die unteren k Stellen bleiben erhalten und bilden den Rest. Zur Modulo-Bildung reicht es also, bei $n = 64$ Bit und $k = 32$ Bit nur die oberen 32 Bit zu löschen bzw. eine lange Zahl (64Bit) in einer kurzen Ganzzahl (32Bit) zu speichern. Die oberen Bits des *overflow*, die dabei verloren gehen, interessieren uns nicht.

Wählen wir $a = 2^p + 1$, so ist $ax_t = 2^p x_t + x_t$. Damit ist die Berechnung des neuen Wertes des LCG insgesamt nur mit Hilfe einer schnellen Linksverschiebung der Bitstellen um p Bits und einer Addition möglich; langsame Multiplikationen benötigen wir nicht.

Den linearen kongruenten Generator können wir verallgemeinern zu der Form

$$x_{t+1} = g(x_t, x_{t-1}, x_{t-2}, \dots) \bmod m \quad (5.25)$$

Dabei dient die Abhängigkeit der Funktion $g(\cdot)$ von früheren Werten dazu, einen eindeutigen Kontext herzustellen, der es erlaubt, nicht unbedingt die Besetzungsdichte zu erhöhen, sondern sehr lange Perioden zu erreichen. Beispiele für die Funktion $g(\cdot)$ sind der *quadratische Kongruenz-Generator*

$$g(\cdot) = ax_t^2 + bx_t + c \quad (5.26)$$

oder der *multiple rekursive Kongruenz-Generator*

$$g(\cdot) = a_0 x_t + a_1 x_{t-1} + a_2 x_{t-2} + \dots + a_s x_{t-s} \quad (5.27)$$

Dieser hat lange Perioden ($m^s - 1$). Eine wesentlich längere Periode (ca. 800 Billionen!) erhalten wir, wenn wir die Koeffizienten a und b ebenfalls als Zufallszahlen nach jeder Generierung neu bestimmen.

5.2.4 Erzeugung einer Verteilung

Damit haben wir eine Methode gefunden, wie wir *uniforme Verteilungen* $U(0,1)$ generieren können. Wie erhalten wir andere Verteilungen? Zunächst erweitern wir die Generierung auf unterschiedliche, nicht normierte Erwartungswerte und Varianzen:

1) *Erwartungswert*

Sei x eine Zufallszahl mit dem Erwartungswert $\langle x \rangle = 0$. Wir erhalten eine neue Zufallsvariable \tilde{x} mit dem Erwartungswert μ durch die Transformation

$$\tilde{x} = x + \mu \quad (5.28)$$

Beweis: Es ist $\langle \tilde{x} \rangle = \langle x + \mu \rangle = \langle x \rangle + \mu = \mu$

2) *Varianz*

Sei x eine Zufallszahl mit Varianz $\langle (x-\mu)^2 \rangle = 1$ und Erwartungswert μ . Wir erhalten eine neue Zufallszahl \tilde{x} mit Varianz σ^2 durch die Transformation

$$\tilde{x} = (x-\mu)\sigma + \mu \quad (5.29)$$

Beweis: Es ist $\langle (\tilde{x} - \mu)^2 \rangle = \langle ((x-\mu)\sigma + \mu - \mu)^2 \rangle = \langle ((x-\mu)^2 \sigma^2) \rangle = \langle (x-\mu)^2 \rangle \sigma^2 = \sigma^2$

Die Transformationen für Erwartungswert und Varianz gelten für alle Verteilungen, auch für die uniforme Verteilung.

3) *Intervallgrenzen*

Wollen wir das Intervall verändern, aus dem eine Zufallszahl gezogen wird, so reicht eine lineare Transformation aus. Sei x eine Zufallszahl aus dem Intervall $[x_0, x_1]$, so erhalten wir eine Zufallszahl y aus dem Intervall $[y_0, y_1]$ durch die Transformation

$$y = y_0 + (x-x_0) \frac{y_1 - y_0}{x_1 - x_0} \quad (5.30)$$

Beim Testen sehen wir mit $x = x_0 \Rightarrow y = y_0$ und $x = x_1 \Rightarrow y = y_1$, dass die Intervallgrenzen stimmen. Erwartungswert und Varianz werden entsprechend transferiert.

In den folgenden Abschnitten sollen nun konkrete Verfahren besprochen werden, wie aus der uniformen Verteilung eine andere, gewünschte erzeugt werden kann.

Generierung einer kontinuierlichen Verteilung

Angenommen, wir möchten eine beliebige Verteilung F aus der uniformen Verteilung generieren. Wie geht das? Hier hat sich die Methode der Inversion der Verteilungsfunktion bewährt. Dazu gehen wir von dem Gedanken aus, dass jede Verteilungsfunktion $F(x)$ nach Gl.(5.8)

$$F(x) = \int_{x_0}^x p(t) dt \quad (5.31)$$

vom Wert $F(x_0) = 0$ bis zum Wert $F(x_1) = 1$ verläuft, also als Wertemenge von $y = F(x)$ das Intervall $[0,1]$ hat und monoton wachsend ist. Wir betrachten darüber

hinaus nur streng monoton wachsende Funktionen, die auch eine Inverse F^{-1} haben. Ziehen wir also mit der uniformen Verteilung $U(0,1)$ zufällig einen Wert $Y = P(y)$ der *cdf* aus dem Intervall $[0,1]$, so entspricht dies einem Wert $F(x)$ und, da $F(x)$ streng monoton wachsend ist, auch eineindeutig einem Wert x . Es existiert also eine inverse Funktion $F^{-1}(y) = x$, mit deren Hilfe wir aus dem uniform verteilten Wert y einen Wert x finden.

$$x = F^{-1}(y) \tag{5.32}$$

In Abb. 5.10 ist dies gezeigt.

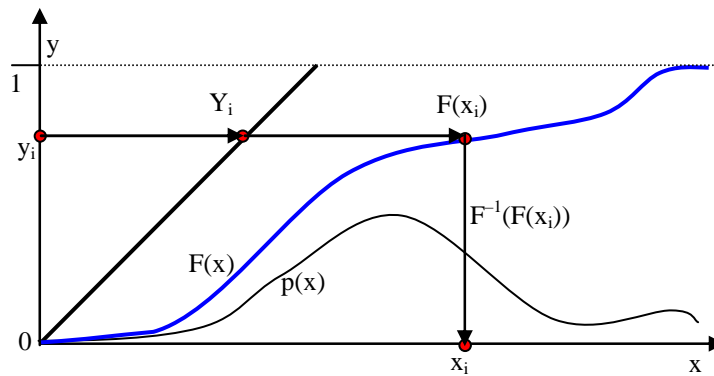


Abb. 5.10 Transformation durch inverse Verteilung

Warum ist dieser so gefundene Wert x nach $F(x)$ verteilt? Schließlich war ja y aus $U(0,1)$ nur uniform verteilt. Der Beweis dafür ist kurz: Die Verteilungsfunktion von x ist

$$\begin{aligned} P(x \leq x_i) &= P(x \leq F^{-1}(Y_i)) = P(F(x) \leq F(F^{-1}(Y_i))) \\ &= P(y \leq Y_i) = Y_i = F(x_i) \end{aligned} \tag{5.33}$$

Inhaltlich bedeutet dies, dass eine Änderung ΔF der Wahrscheinlichkeitsmasse von F gerade einer Änderung ΔY der uniformen Verteilung entspricht. Wenn beide Änderungen gleich sind, so ist auch die Wahrscheinlichkeit von y in ΔY zu sein, genauso groß wie von x in ΔF zu sein; die „Wahrscheinlichkeitsmassen“ sind gleich.

Beispiel Erzeugung von Zufallszahlen nach einer Exponentialverteilung

Die Dichtefunktion pdf

$$p(x) = \begin{cases} \lambda e^{-\lambda x} & \text{für } x \geq 0 \\ 0 & \text{sonst} \end{cases} \quad (5.34)$$

der Exponentialverteilung hat als Verteilungsfunktion

$$F(x) = \int_0^x \lambda e^{-\lambda t} dt = -e^{-\lambda t} \Big|_0^x = 1 - e^{-\lambda x} \quad (5.35)$$

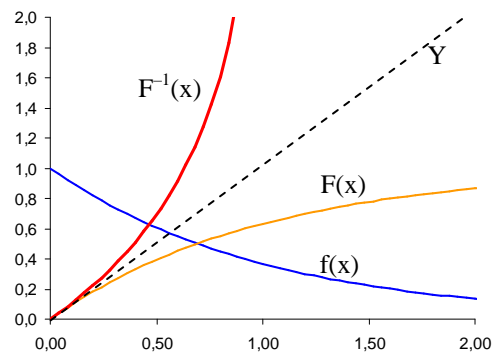
Die inverse Funktion ist

$$y = F(x) = 1 - e^{-\lambda x} \Leftrightarrow e^{-\lambda x} = 1 - y \Leftrightarrow -\lambda x = \ln(1 - y)$$

oder

$$x = F^{-1}(y) = -\frac{\ln(1-y)}{\lambda} \quad (5.36)$$

In Abb. 5.11 sind alle drei Funktionen eingetragen.

**Abb. 5.11** Exponentielle pdf, cdf und inverse Funktion.

Wir sehen in Abb. 5.11, dass die inverse Funktion grafisch gerade die an der Winkelhalbierenden Y (gestrichelt) gespiegelte Funktion $F(x)$ darstellt.

Die Formel (5.36) lässt sich übrigens noch weiter vereinfachen. Mit der Beobachtung, dass $y \in U(0,1)$ dieselbe Verteilungsfunktion hat wie $1 - y$, können wir uns eine Subtraktion sparen und die Transformation vereinfachen zu

$$x = -\frac{\ln(y)}{\lambda} \tag{5.37}$$

Aufgabe 5.2

Man bestimme die Transformation einer uniform verteilten Zufallsvariable auf eine Dreiecksverteilung, definiert durch

$$p(x) = \begin{cases} b + \frac{a}{b}x & \text{für } x < 0 \\ b - \frac{a}{b}x & \text{für } x \geq 0 \end{cases} \tag{5.38}$$

Diskrete Verteilungen

Angenommen, wir haben eine diskrete Verteilung, gemessen als Histogramm. Wie erhalten wir eine danach verteilte Zufallsvariable für die Simulation? Betrachten wir dazu das Histogramm und seine Verteilungsfunktion in Abb. 5.12.

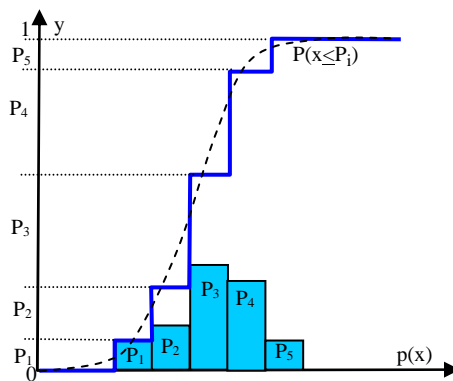


Abb. 5.12 Inversion der diskreten Verteilung

Jede Säule des Histogramms hat eine relative Häufigkeit P_i , wobei

$$\sum_{i=1}^N P_i = 1 \tag{5.39}$$

gilt.

Für die Erzeugung der Zufallszahlen nutzen wir wieder den Gedanken der inversen Funktion: Wir unterteilen das gesamte Intervall $[0,1]$ von y in N Intervalle, wobei jedes Intervall eine Breite von $\Delta y_i = P_i$ hat. Jede Zufallsvariable y , die wir aus der uniformen Verteilung $U(0,1)$ ziehen, fällt mit der Wahrscheinlichkeit P_i in das i -te Intervall und wird als Ereignis $e_i = \text{„Zufallszahl in Intervall } i\text{“}$ registriert. Wenn wir nach dem Ziehen der Zufallszahl dieses Intervall bestimmen, so müssen wir nun nur noch den x -Wert zufällig aus dem x -Intervall ziehen.

Für die Implementierung zur Ereignis- bzw. Intervallbestimmung verwenden wir folgenden Simulationsmechanismus:

Aufgabe: Simuliere das Auftreten eines zufälligen Ereignisses mit Wahrscheinlichkeit P_i .

Lösung: Wir wissen, dass für die uniforme Verteilung $P_i = P(x \leq P_i)$ gilt. Ziehen wir also eine uniform verteilte Zahl aus dem Intervall $[0,1]$, so erhalten wir mit Wahrscheinlichkeit P_i eine Zahl, die kleiner ist als P_i . Also ist der Code

```
x = UniformRandom (0,1)
IF x ≤ Pi THEN SimulateEvent(i)
```

Dies simuliert das Ziehen einer Zufallsvariablen aus einem Intervall. Bei obigem Problem haben wir aber mehrere Intervalle.

Aufgabe: Simuliere das Auftreten eines aus mehreren Ereignissen e_1, e_2, \dots, e_n , wobei jedes Ereignis mit Wahrscheinlichkeit P_i auftritt und die Summe aller Wahrscheinlichkeiten eins ergibt.

Lösung: Dazu unterteilen wir das ganze Intervall $[0,1]$ in n Unterintervalle $[0, P_1], [P_1, P_2], \dots, [P_{n-1}, 1]$ und notieren die Grenzen zwischen den Intervallen in einem $n-1$ dimensionalen Feld $A[1] = P_1, A[2] = A[1] + P_2, \dots, A[n-1] = A[n-2] + P_{n-1}$. Der Algorithmus zieht nun eine uniforme Zufallszahl und vergleicht sie mit den Grenzen um zu entscheiden, in welches Intervall sie eingeordnet werden muss. Da die Breite der Intervalle der uniformen Verteilung den gewünschten Wahrscheinlichkeiten entsprechen, werden die Ereignisse e_i mit der Wahrscheinlichkeit P_i auftreten. Der Pseudocode dafür ist

```
x = UniformRandom (0,1); e = 0; i = 1;
WHILE i < n DO
  IF x < A[i] THEN e = i ; i = n      (*Abbruch*)
  ELSE i=i+1
ENDWHILE
IF e = 0 THEN e = n;                (*letztesIntervall*)
SimulateEvent(e);
```

Die Abfrage von $e = 0$ ist für das letzte Intervall nötig, da die Schleife zu ende geht und nicht explizit „ e_n “ entschieden wurde.

Der oben skizzierte Algorithmus ermöglicht nicht nur die Erzeugung von Zufallszahlen entsprechend einer vorgegebenen Verteilung, sondern ermöglicht auch die Simulation von Einzelereignissen.

Allerdings bleibt dabei eine Frage ungelöst: Wie soll die Zufallszahl innerhalb eines Intervalls verteilt sein? Natürlich könnten wir in den Intervallgrenzen des i -ten Intervalls zwischen den Werten von P_{i-1} und P_{i+1} der Nachbarintervalle interpolieren, etwa durch ein Polynom. In Abb. 5.12. ist eine interpolierte Verteilungsfunktion gestrichelt eingezeichnet.

Aber diese Interpolation ist nicht rational: Woher wissen wir, dass die Interpolation korrekt ist? Warum sollten wir nicht einfach den Wert y mittels der linearen Abbildung aus Gl. (5.30) auf einen x -Wert abbilden? In beiden Fällen erhalten wir das selbe Histogramm, die Ausgangsbasis unserer Überlegungen. Für die Interpolation sind sowohl ein Zusatzwissen über den Entstehungsprozess der beobachteten Zufallszahlen („kontinuierlicher Prozess“) notwendig, als auch die anwendungsabhängige Notwendigkeit, überhaupt eine Interpolation vornehmen zu müssen. Wir müssen uns darüber im Klaren sein, dass bei jeder Interpolation zusätzliche Information in die Simulation einfließt, die falsch sein kann.

Zusammensetzen von Verteilungen

Haben wir einige Grundverteilungen definiert, so können wir sie dazu benutzen, die Generierung von komplizierten Verteilungen aus den einfachen Verteilungen zusammen zu setzen. Ein Beispiel soll dies illustrieren.

Beispiel Laplace-Verteilung

Die doppelt-exponentielle Verteilungsdichte lautet

$$p(x) = \frac{1}{2} e^{-|x|} \quad \text{Laplace-Verteilungsdichte} \quad (5.40)$$

In Abb. 5.13 ist sie verdeutlicht.

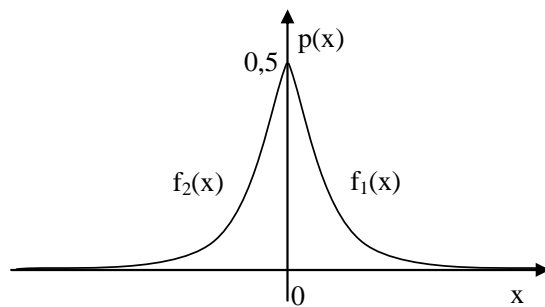


Abb. 5.13 Laplace-Verteilungsdichte

Diese Verteilungsdichte lässt sich auch ausführlich

$$p(x) = \begin{cases} \frac{1}{2}e^{-x} & \text{für } x \geq 0 \\ \frac{1}{2}e^{+x} & \text{für } x < 0 \end{cases} \quad (5.41)$$

als Zusammensetzung aus zwei exponentiellen Verteilungen von Gl. (5.34) formulieren

$$f_1(x) = \begin{cases} e^{-x} & \text{für } x \geq 0 \\ 0 & \text{sonst} \end{cases}, \quad f_2(x) = \begin{cases} e^x & \text{für } x < 0 \\ 0 & \text{sonst} \end{cases} \quad (5.42)$$

Beide werden entsprechend ihren Einzelwahrscheinlichkeiten zur resultierenden Verteilung zusammengesetzt

$$p(x) = \frac{1}{2} f_1(x) + \frac{1}{2} f_2(x) \quad (5.43)$$

Allgemein lässt sich die zusammengesetzte Verteilungsdichte formulieren zu

$$p(x) = \sum_i P_i f_i(x) \quad \text{mit} \quad \sum_i P_i = 1 \quad (5.44)$$

Dabei kann man auch kontinuierliche Verteilungen mit diskreten Verteilungen kombinieren, um Beobachtungen mit Theorie zu einem Gesamtbild zu weben.

Überlagerung von Verteilungen

Eine wichtige und beliebte Methode, Zufallszahlen x einer Verteilung zu erzeugen, besteht in der Überlagerung von mehreren Zufallszahlen X_i von einer anderen Verteilung

$$x = X_1 + X_2 + \dots + X_m \quad (5.45)$$

Dies wird auch als „m-fache Konvolution eines X_i “ bezeichnet.

Beispiel m -Erlang-Verteilung

Sei X_i eine exponentiell verteilte Zufallsvariable mit dem Erwartungswert β/m . Dann ist x eine m -Erlang-verteilte Zufallsvariable mit Erwartungswert β . Da wir die exponentiell verteilten Zufallszahlen X_i nach Gl. (5.37) mit Hilfe der inversen Verteilungsfunktion aus den uniform verteilten Zufallszahlen bilden können, ist

$$x = \sum_{i=1}^m X_i = \sum_{i=1}^m -\frac{\beta}{m} \ln(y_i) = -\frac{\beta}{m} \sum_{i=1}^m \ln y_i = -\frac{\beta}{m} \ln \prod_{i=1}^m y_i \quad (5.46)$$

ist, müssen wir nicht m Logarithmen aus den m uniform verteilten Zufallszahlen y_i bilden, sondern können dies durch eine schnellere Multiplikation ersetzen.

Beispiel Normalverteilung

Die Normalverteilung ist nicht so einfach zu generieren. Nutzen wir die Methode der inversen Verteilungsfunktion, so müssen wir erst analytisch die Verteilungsfunktion der Gauß-Verteilung bilden. Dies ist die Gauß'sche Fehlerfunktion Φ , die als Integral nur in tabellierter Funktion vorliegt. Eine Inversion wäre nur über Tabellen möglich – eine sehr unbefriedigende Situation. Stattdessen können wir uns die Überlagerung von Zufallszahlen zu Nutze machen. Nach dem zentralen Grenzwertsatz wissen wir, dass die Verteilung für die zentrierte Summenvariable x aus n unabhängigen Zufallsvariablen X_i

$$x(n) = \sum_{i=1}^n \frac{(X_i - \mu)}{\sigma} \tag{5.47}$$

zur Normalverteilung $N(\mu, \sigma)$ mit $\mu = 0, \sigma = 1$

$$\lim_{n \rightarrow \infty} x(n) = N(0,1) \tag{5.48}$$

konvergiert, unabhängig von der Art der Verteilungen für X_i . Der Unterschied zu Gl.(5.46) liegt dabei in der Anzahl n der Zufallszahlen: Für große m geht die m -Erlang-Verteilung in eine Normalverteilung über. Wie können wir uns konkret aus der uniformen Verteilung die Normalverteilung generieren?

Angenommen wir verfügen über einen Zufallsgenerator der uns uniform verteilte X_i aus dem Intervall $[-d, +d]$ mit dem Erwartungswert $\mu = 0$ liefert. Dann ist die Wahrscheinlichkeitsdichte $p(X) = 1/(2d)$. Die erwartete Varianz von X wird von den einzelnen Varianzen der X_i gebildet, die sich überlagern. Dazu betrachten wir die n gezogenen Zufallszahlen als Zufallsvektor

$$\mathbf{X} = (X_1, X_2, \dots, X_n)^T \text{ mit dem Erwartungswert } \boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_n)^T \tag{5.49}$$

dessen Varianz sich ergibt zu

$$\begin{aligned} \bar{\sigma}^2 = \langle (\mathbf{X} - \boldsymbol{\mu})^2 \rangle &= \left\langle \sum_{i=1}^n (X_i - \mu_i)^2 \right\rangle = \sum_{i=1}^n \langle (X_i - \mu_i)^2 \rangle \\ &= (\sigma_1^2 + \dots + \sigma_n^2) = n\sigma^2 \end{aligned} \tag{5.50}$$

bei gleicher Varianz σ^2 der Zufallswerte X_i und ist hier bei $\mu_i = 0$

$$\overline{\sigma^2} = n \langle X_i^2 \rangle = n \int_{-d}^{+d} p(X) X^2 dX = \frac{n}{2d} \int_{-d}^{+d} X^2 dX = \frac{n}{2d} \left. \frac{X^3}{3} \right|_{-d}^{+d} = \frac{n}{3} d^2 \quad (5.51)$$

Eine vorgegebene Varianz σ^2 lässt sich also mit einer Summe von $n = 3\sigma^2/d^2$ uniform verteilten, zentrierten Zufallszahlen erreichen. Umgekehrt wird eine Normalverteilung $N(0,1)$ mit uniform verteilten Zufallszahlen X_i aus dem Intervall $[-1/2, 1/2]$ erreicht, wenn $n = 12$ ist. Damit haben wir einen einfachen Algorithmus gefunden, normal verteilte Zufallsvariable zu generieren. In Abb. 5.14 ist dies verdeutlicht.

```

PROCEDURE GaussRandom(mu, sigma:REAL):REAL;
  (* generiert einen normalverteilten Zufallswert
   mit Mittelwert  $\mu$  und Streuung  $\sigma^*$  *)
VAR n, sum:REAL;
BEGIN
  sum:=0;
  FOR i:=1 TO 12 DO
    sum :=sum + UniformRandom() - 0.5
  END;
RETURN sum*sigma + mu;

```

Abb. 5.14 Pseudocode für normal verteilte Zufallszahlen

Möchte man anstelle einer normal verteilten Zufallsvariablen z mit $\sigma = 1$ eine Normalverteilung mit anderer Standardabweichung erhalten,

$$p(z,1) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \quad \rightarrow \quad p(x,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

so ist es ungünstig, dazu die Anzahl der Summanden zu verändern; bei geringer Streuung könnte weniger als ein Summand übrig bleiben. Stattdessen ist es besser, die Transformation von $z = (x-\mu)/\sigma$ auf x mit Gl. (5.29) vorzunehmen, so dass der $N(0,1)$ verteilte Zufallswert z mit der Standardabweichung σ zu $x = z\sigma$ multipliziert wird. Für schmale Verteilungen „stauchen“ wir durch die Skalierung die Normalverteilung oder „ziehen“ sie für breite Verteilungen auseinander.

Beispiel χ^2 -Verteilung und $F_{n,m}$ -Verteilung

Eine χ^2 -Verteilung mit n Freiheitsgraden (χ_n^2 -Verteilung) lässt sich erzeugen, indem man Zufallszahlen erzeugt mittels Überlagerung von n unabhängigen, normal verteilten Zufallszahlen

$$x = \sum_{i=1}^n X_i^2 \quad \text{wobei } X_i \text{ aus } N(0,1) \quad (5.52)$$

Haben wir je eine Zufallszahl x_1 aus einer χ_n^2 -Verteilung und ein x_2 aus einer χ_m^2 -Verteilung, so ist der Quotient y eine Zufallszahl, die einer sog. F(n,m)-Verteilung gehorcht:

$$y = \frac{x_1}{x_2} \quad \text{ist } F_{n,m}\text{-verteilt} \quad (5.53)$$

Sie ist auch als Fisher-Snedecor-Verteilung bekannt und wird gern in der Psychologie verwendet, um herauszufinden, ob sich die Varianzen von zwei betrachteten Systemen signifikant unterscheiden.

5.2.5 Test der Zufallszahlen

Die mit den bisher vorgestellten Methoden generierte Zufallszahlen können gute Qualität haben, oder aber sehr viele Regelmäßigkeiten und damit ein inakzeptables Verhalten zeigen. Verwenden wir solche schlechten Generatoren, so können unsere Simulationen Effekte zeigen, die nicht auf die Eigenschaften der verwendeten Daten, Algorithmen und Simulationen zurück zu führen sind, sondern nur auf die „Macken“ des Generators. Es ist sehr deprimierend, wenn der erhoffte Nobelpreis sich als Schlampigkeit bei der Generatorprogrammierung erweist! Im Unterschied zu „normalen“ Programmen fallen Fehler bei der Programmierung oder Funktion von Zufallszahlen nicht sofort auf, da Menschen Wahrscheinlichkeiten schlecht einschätzen können. Nicht umsonst bestand das erste Programm der ersten, aus sehr unzuverlässigen Röhren gebauten Rechner aus einem Zufallszahlengenerator!

Möchten wir zuverlässige Simulationsergebnisse, so sollten wir nicht nur die Funktionalität der Zufallszahlengeneratoren kennen und einschätzen, sondern auch solide testen. Dazu gibt es verschiedene Tests, von denen die Wichtigsten hier vorgestellt werden.

Grafische Plausibilitätstests

Ein erster Plausibilitätstest lässt sich grafisch realisieren. Dazu tragen wir den Wert x_{t+1} als Funktion des Vorgängers x_t auf. In Abb. 5.15 ist solch ein Diagramm zu sehen.

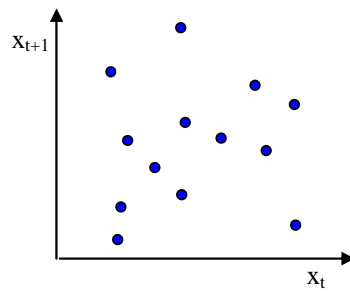


Abb. 5.15 Abhängigkeitsplot der Zufallsvariablen

Jeder Punkt in der Zeichnung ist durch ein Tupel (x_t, x_{t+1}) festgelegt. Für perfekte Unabhängigkeit sollten alle Punkte gleichmäßig über die Fläche verstreut sein. Bilden die Punkte eine „Strasse“, also eine Verdichtung, die durch eine Gerade approximiert werden kann, so sind sie linear korreliert – dieser Generator fällt bereits jetzt durch unseren Test durch; er ist unbrauchbar. Allerdings, auch wenn wir nichts bemerken würden, kann der Generator Probleme haben. Dazu erweitern wir die Betrachtung auf Tripel (x_t, x_{t+1}, x_{t+2}) . Plotten wir die selben Zufallszahlen, aber als Tripel, so finden wir bei dem berühmten FORTRAN Zufallszahlengenerator RANDU mit den LCG-Werten $m = 2^{31}$, $a = 2^{16} + 3$, $b = 0$ und $x_0 = 123.456.789$ regelmäßige Strukturen, siehe Abb. 5.16.

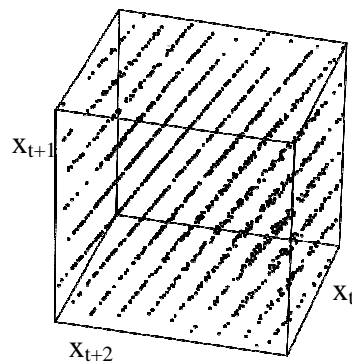


Abb. 5.16 3D-Abhängigkeitsplot der Zufallsvariablen für $n = 2000$ Tripel des RANDU-Generators (nach Law, Kelton 2000)

Run-Tests

Ein weiteres einfaches Testmittel besteht darin, jeden Anstieg oder Abstieg der Zufallszahlen symbolisch mit einem Zeichen s_t zu notieren:

$$s_t = \begin{cases} 1 & x_t \geq x_{t-1} \\ 0 & x_t < x_{t-1} \end{cases} \quad (5.54)$$

Konkateneren wir alle Zeichen zu einer Zeichenkette (*run*) $s = s_1 s_2 \dots s_t$, so sollte die Anzahl der „1“ bzw. „0“, die eine Unterkette „111..“ bzw. „000...“ bilden, stochastischen Proportionen unterworfen sein. Es lässt sich zeigen, dass bei uniformen Zufallszahlen die Verteilung der Kettenlängen r bei N Unterketten zu einer Normalverteilung führt mit

$$\langle r \rangle = \frac{2N-1}{3} \quad \text{und} \quad \sigma^2(r) = \frac{16N-29}{90} \quad (5.55)$$

Aus einer Stichprobe r lässt sich die Wahrscheinlichkeit α ermitteln, mit der die Nullhypothese „Kette gehört einer Normalverteilung an“ stimmt. Dazu bilden wir die normierte Zufallsvariable

$$\hat{r} = \frac{r - \mu}{\sigma} \quad (5.56)$$

und prüfen, ob sie im Intervall zwischen den durch α festgelegten Vertrauensgrenzen liegt.

Wir können auch die Absolutwerte vergleichen und damit solche „1“, „0“-Ketten erzeugen

$$s_t = \begin{cases} 1 & x_t \geq 0,5 \\ 0 & x_t < 0,5 \end{cases} \quad (5.57)$$

Verwenden wir uniforme Zufallszahlen, so sollte

$$P(s_t = „1“) = P(s_t = „0“) = 1/2 \Rightarrow P(n \text{ gleiche Zeichen}) = (1/2)^n$$

gelten. Dies können wir empirisch nachprüfen durch Erzeugung einer Kette s und Auszählen der Unterketten für 0 und 1.

Anpassungstests

Eine übliche Testmethode besteht darin, die gemessenen Daten und die erwarteten Daten miteinander zu vergleichen. In diesem Fall prüfen wir, ob das ermittelte Histogramm der Verteilung signifikant von der erwarteten Verteilung abweicht. Dazu verwenden wir den χ^2 -Test.

Für N Intervalle und n Beobachtungen sollte für die uniforme Verteilung die Anzahl der erwarteten Beobachtungen pro Intervall $S_i = n/N = \text{const.}$ gelten. Stattdessen aber ermitteln wir eine Differenz mit der Varianz

$$\chi^2 = \frac{N}{n} \sum_{i=1}^N \left(S_i - \frac{n}{N} \right)^2 \quad (5.58)$$

Für große N ist dies eine χ^2 -Verteilung mit $N-1$ Freiheitsgraden. Für den Test prüfen wir, ob die ermittelte mittlere Abweichung χ^2 größer als der durch die Akzeptanzwahrscheinlichkeit $1-\alpha$ festgelegte Punkt $\chi^2_{N-1,1-\alpha}$ (*kritischer Punkt*). Ist dies der Fall, so ist die mittlere Abweichung zu groß und die Nullhypothese wird abgelehnt. Der kritische Punkt lässt sich aus dem korrespondierenden kritischen Punkt $x_{1-\alpha}$ einer Normalverteilung

$$\chi^2_{N-1,1-\alpha} \approx (N-1) (1 - a + x_{1-\alpha} a^{1/2})^2 \quad \text{mit } a = \frac{2}{9(N-1)} \quad (5.59)$$

erschließen. Noch schärfer ist der Anpassungstest, wenn wir jeweils d Zahlen zu einem d -dimensionalen Tupel zusammenfassen

$$\mathbf{x}_1 = (x_1, \dots, x_d), \mathbf{x}_2 = (x_{d+1}, \dots, x_{2d}), \dots$$

und dann den d -dimensionalen χ^2 -Test machen

$$\chi^2(d) = \frac{N^d}{n} \sum_{i,j,k,\dots=1}^N \left(P_{ijk\dots} - \frac{n}{N^d} \right)^2 \quad (5.60)$$

Beispiel

Der RANDU-Generator aus Abb. 5.16 erzeugt folgende Werte für χ^2 :

$$\begin{aligned} \chi^2 &= \chi^2(1) = 4202,0 < 4211,4 \quad \text{für } \alpha = 0,1 \\ \chi^2(2) &= 4202,3 \\ \chi^2(3) &= 16.252,3 \end{aligned}$$

Wir sehen deutlich, dass die schlechte Leistung bei $d = 3$ aus Abb. 5.16 sich im Test niederschlägt. Interessanterweise kommt dies aber für kleinere Dimensionen nicht zum Ausdruck.

5.3 Simulation mit Monte Carlo-Verfahren

Die Grundidee der Monte-Carlo-Simulation besteht darin, ein deterministisches Problem mit Hilfe von Zufallsexperimenten zu lösen. Die Bezeichnung „Monte-Carlo“ rührt von dem gleichnamigen Ort her, der ein bekanntes Spielkasino beherbergt und besagt nur, dass zufällige Ereignisse involviert sind.

Beispiel Kaffeautomat

Angenommen, in Ihrer Firma steht ein Kaffeautomat, der bei Einwurf von 1 € einen Becher Kaffee ausgibt. Dazu muss man ein 50¢ - Stück einwerfen und fünf weitere 10¢ - Stücke. Leider aber ist die Münzprüfstufe des Automaten zu scharf eingestellt: 20% aller 50¢ - Stücke und 30% aller 10¢ - Stücke akzeptiert er nicht und lässt sie – unabhängig davon, wie oft man sie einwirft – durchfallen.

Frage: Wie viele Geldstücke jeder Sorte (oder: Wie viel Geldwert) müssen Sie dabei haben, um mit 99%-ger Sicherheit einen Kaffee zu bekommen?

Die Antwort auf diese Frage kann man analytisch mit Kenntnissen der Stochastik erhalten – oder aber durch Simulation. Für die zweite Methode interessieren wir uns in diesem Abschnitt und wollen sie näher ansehen. Dabei können wir die allgemeine Situation der stochastischen Simulation so charakterisieren, dass wir Fragen an das Simulationssystem stellen in Form von Systeminformation und Auswahlregeln („innere Mechanismen“), und durch die Simulation von Zufallsergebnissen als Eingabe \mathbf{x} der Simulation für die Ausgabe T bestimmte Kenngrößen der Simulation als Antworten erhalten. In Abb. 5.17 ist dies visualisiert.

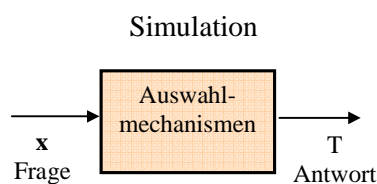


Abb. 5.17 Leistungsschema der Simulation

Betrachten wir dazu zwei Verfahren, das *hit-or-miss* Monte Carlo-Verfahren und das *crude* Monte Carlo-Verfahren.

5.3.1 *Hit-or-miss* Monte Carlo-Verfahren

Ein typischer Ansatz besteht darin, alle möglichen Ereignisse zu simulieren und dann mit Hilfe einer Selektionsregel die „günstigen“ auszusuchen. Der relative Anteil der Ereignisanzahl ist dann die relative Häufigkeit für „günstig“. Verdeutlichen wir dies an dem Problem der Berechnung eines Integrals, siehe (Hammersley, Handscomb 1965).

Integration einer Linearen Funktion

Angenommen, wir haben eine lineare Funktion $f(x) = x$. Dann können wir die Fläche unter Kurve, das Integral im Intervall $[0, x_0]$, durch die Erzeugung von Zufallszahlen abschätzen. Dazu erzeugen wir uns zwei-dimensionale Zufallszahlen (x, y) mit $x \in [0, x_0]$ und $y \in [0, y_0]$. Die so erzeugten Ereignisse sortieren wir durch die Abfrage

```
IF  $y \leq f(x)$  THEN  $H = H+1$ 
```

Die Anzahl H der Ereignisse, die die Bedingung erfüllen, ist bei n Ereignissen im Erwartungswert $n/2$.

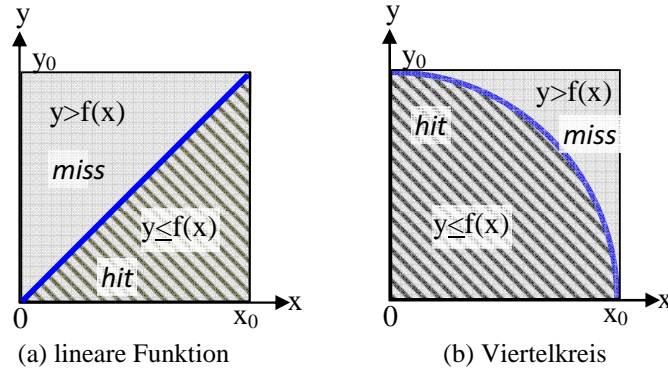


Abb. 5.18 Die *hit-and-miss* Simulation zur Integralmessung

Den Selektionsvorgang können wir uns dabei so vorstellen, als ob wir n Kugeln von oben auf die Gesamtfläche aller möglichen Ereignisse fallen lassen würden. In Abb. 5.18 (a) ist die Situation abgebildet. Die graue Fläche deckt das 2-dim Intervall ab, aus der das Ereignis stammen kann; es entspricht der Fläche, auf welche die Kugel fallen kann. Die schraffierte Fläche entspricht dem Anteil, auf den die Bedingung $y < f(x)$ zutrifft, die also durch die Kugeln „getroffen“ wurde (*hit*). Für die restliche graue Fläche trifft die Bedingung nicht zu („*miss*“). Der Anteil der Treffer an der Gesamtzahl von Ereignissen ist gerade das gesuchte Integral.

Integration eines Kreises

Diese Idee ist leicht auf andere, problematischere Integrale anwendbar. Beispielsweise gilt für die Ermittlung der Kreisfläche F eines Einheitskreises mit $R = 1$

$$F = \pi R^2 = \pi \quad (5.61)$$

Betrachten wir einen Viertelkreis so wie in Abb. 5.18 (b), so ist das Verhältnis von Kreisfläche $F = \pi/4$ (schraffiert) zur Gesamtfläche $A = 1$ (grau)

$$T = \frac{F}{A} = \pi/4 \quad (5.62)$$

Dieses Verhältnis gilt auch für die Treffer eines Zufallsexperiments, wenn wir Kugeln uniform verteilt auf eine solche Fläche fallen lassen und die Auswahlregel für Treffer aufstellen

$$y < f(x) \quad \text{wobei } x_1^2 + x_2^2 = 1, \text{ also } x_2 = f(x_1) = (1-x_1^2)^{1/2} \quad (5.63)$$

Wir erhalten so nicht nur ein Integral, sondern können über die Kenntnis der analytischen Lösung sogar die Zahl π durch Zufallsexperimente bestimmen!

Buffons Problem

Ein anspruchsvolleres Problem ist folgende Situation (*Buffons Problem*): Ein Stab der Länge L fällt auf den Boden, der mit sehr langen Streifen der Breite $B > L$ bedeckt ist (Abb. 5.19). Wie groß ist die Wahrscheinlichkeit P , dass der Stab eine Streifengrenze kreuzt und nicht innerhalb eines Streifens zur Ruhe kommt?

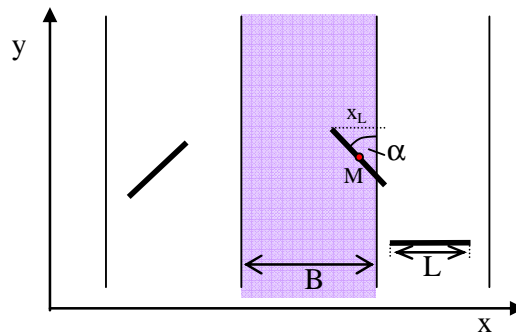


Abb. 5.19 Buffons Problem

Dieses Problem können wir durch Simulation lösen. Legen wir die Lage des Stabes durch die Lage seines Mittelpunkts M und den Winkel α zwischen Stab und Grenze fest, so modellieren wir uns den Abstand des Stabmittelpunkts M von der nächstgelegenen Grenze durch eine uniforme Zufallsvariable x aus $U(0, B/2)$ und den Winkel α durch die uniforme Zufallsvariable α aus $U(0, \pi/2)$. Sei x_L die Ausdehnung des Stabes in x -Richtung. Dies lässt sich errechnen aus der Definition des Sinus im Dreieck

$$\sin(\alpha) = \frac{x_L}{L}, \quad x_L = L \sin(\alpha) = f(\alpha) \quad (5.64)$$

Ist der Abstand vom Mittelpunkt zur Grenze kleiner als $x_L/2$, so wird der Stab die Grenze schneiden, so dass die Trefferregel lautet

$$\text{IF } x \leq L \cdot \sin(\alpha) / 2 \text{ THEN } H = H + 1 \quad (5.65)$$

In Abb. 5.20 ist dies visualisiert.

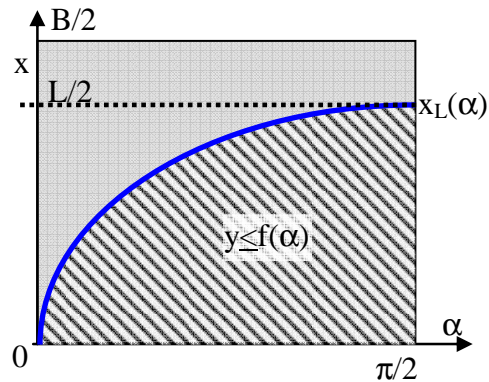


Abb. 5.20 Auswahlregel und Lage des Stabes

Die Wahrscheinlichkeit P lässt sich ausdrücken als Verhältnis des Integrals unter der Kurve $f(\alpha)$ zur Gesamtfläche

$$P = \frac{\int_0^{\pi/2} \frac{L}{2} \sin \alpha \, d\alpha}{\frac{B}{2} \frac{\pi}{2}} = \frac{-\frac{L}{2} \cos \alpha \Big|_0^{\pi/2}}{\frac{B\pi}{4}} = \frac{2L}{B\pi} \quad (5.66)$$

Auch hier lässt sich durch die Monte-Carlo-Simulation entweder die Wahrscheinlichkeit P feststellen, oder, wenn sie analytisch bekannt ist, einen der Parameter messen, falls er nicht gegeben ist. Dies ist beispielsweise dann interessant, wenn wir B oder L nicht direkt messen können, weil sie etwa im mikroskopischen Bereich sind, aber die physikalischen, makroskopischen Auswirkungen von sehr vielen "Nadeln" in einem Kristallgitter beobachten können.

Test der Resultate

Die Simulation der Integralsberechnung ergibt Resultate, die allerdings etwas schwanken. Wir müssen uns deshalb die Frage stellen: Wie genau sind die Ergebnisse eines solchen Verfahrens? Können wir die Ergebnisse testen?

Verwenden wir n Zufallsexperimente mit einer Trefferrate $T_i = H/n$, so ist das arithmetische Mittel \hat{T}

$$\hat{T} = \frac{1}{n} \sum_{i=1}^n T_i \quad (5.67)$$

als Schätzung für den Erwartungswert daraus normal verteilt. Jede Schätzung der Trefferrate T_i weicht mit der Varianz σ^2 vom wahren Wert T_μ ab; die resultierende

Varianz von \hat{T} ist gerade $n\sigma^2$. Die Einzelvarianz σ^2 lässt sich dabei abschätzen durch die experimentell beobachtete Varianz s^2

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (T_i - \hat{T})^2 \quad (5.68)$$

Mit diesen Werten können wir nun prüfen, ob ein beobachtetes Ergebnis T_i aus dem Konfidenzintervall einer Normalverteilung um den Erwartungswert \hat{T} stammt. Ist $[-N_\alpha, N_\alpha]$ das Konfidenzintervall einer Normalverteilung, in dem mit der Wahrscheinlichkeit $1-2\alpha$ die Zufallszahlen aus $N(0,1)$ liegen, so ist das mit unserer Varianz entsprechend skalierte Konfidenzintervall $[\hat{T} - x_\alpha, \hat{T} + x_\alpha]$ durch die Grenzen

$$x_\alpha = N_\alpha \frac{s}{\sqrt{n}} \quad (5.69)$$

gegeben. Liegt T_i innerhalb des Intervalls, so gilt die Nullhypothese "T_i ist das Ergebnis des hit-and-miss-Verfahrens" nach dem beidseitigen Test mit Wahrscheinlichkeit $1-2\alpha$.

Dabei müssen wir aber beachten, dass eine Normalverteilung erst mit guter Genauigkeit ab etwa $n > 40$ zu beobachten ist. Darunter sollte man eher den t-Test machen, also anstelle der Normalverteilung eine t-Verteilung mit $n-1$ Freiheitsgraden annehmen und einem Konfidenzintervall, das mit der Varianz zu

$$x_\alpha = t_\alpha \frac{s}{\sqrt{n}} \quad (5.70)$$

skaliert wird, benutzen.

5.3.2 Crude Monte Carlo-Verfahren

Die bisher besprochene Methode bestimmt Integrale durch Simulation von Zufallsvariablen und ihrer Einschränkung durch eine Auswahl mittels zu integrierender Funktionen. Dabei haben wir die Situation, dass sowohl die Treffermenge als auch die Gesamt-Menge Zufallszahlen sind, so dass der Quotient eine Varianz aufweist, die von beidem abhängt. Wir können nun die Varianz des Ergebnisses verringern, indem wir nur die Gesamtmenge verwenden, um das Integral als Erwartungswert zu ermitteln.

Betrachten wir dazu ein Beispiel, etwa das Integral

$$I = \int_0^1 f(x) dx = \int_0^1 \frac{e^x - 1}{e - 1} dx \quad (5.71)$$

Die Funktion $f(x)$ nimmt im Intervall $[0,1]$ Werte an zwischen null und eins, so dass die Fläche unter der Kurve durchaus mit der *hit-and-miss*-Methode ermittelt-

bar wäre. Wir ermitteln den Wert aber diesmal anders. Dazu formen wir das Integral etwas um, ohne seinen Wert zu verändern

$$I = \int_0^1 \frac{e^x - 1}{e - 1} \cdot 1 \, dx = \int_0^1 \tilde{f}(x)p(x)dx = \langle \tilde{f}(x) \rangle \text{ mit } p(x) = 1 \quad (5.72)$$

Wir haben die Funktion $f(x)$ so verändert, dass sie als Produkt einer Funktion \tilde{f} und einer Dichtefunktion $p(x)$ geschrieben werden kann, wobei in diesem konkreten Fall $f = \tilde{f}$ und die uniforme Verteilung $p(x) = 1$ gelten. Dabei bedeutet Gl. (5.72), dass wir den Wert des Integrals I auch als Erwartungswert der Funktion \tilde{f} durch Simulation ermitteln können

$$I = \langle \tilde{f}(x) \rangle = \frac{1}{n} \sum_{i=1}^n \tilde{f}(x_i) \quad \text{wobei } x_i \in U(0,1) \quad (5.73)$$

Haben wir ein Integral in einem allgemeinen Intervall, so ändern wir $\tilde{f}(x)$ entsprechend:

$$\begin{aligned} I &= \int_a^b f(x)dx = \int_a^b f(x)(b-a) \frac{1}{b-a} dx \\ &= \int_a^b \tilde{f}(x)p(x)dx = \langle \tilde{f}(x) \rangle \text{ mit } p(x) = \frac{1}{b-a} \text{ und } \tilde{f}(x) = f(x)(b-a) \end{aligned} \quad (5.74)$$

Das Integral erhalten wir wieder wie in Gl. (5.73) durch Simulation und Bildung des diskreten Erwartungswertes, wobei allerdings diesmal $x_i \in U(a,b)$ gilt. Die Methode können wir ausweiten auf andere Integrale, etwa solche mit einer unendlichen Obergrenze. Beispielsweise ist

$$I = \int_0^{\infty} x \lambda e^{-\lambda x} dx = \int_0^{\infty} \tilde{f}(x)p(x)dx \text{ mit } \tilde{f}(x) = x \text{ und } p(x) = \lambda e^{-\lambda x} \quad (5.75)$$

Dabei verwenden wir als Dichtefunktion die Exponentialverteilung, die wir die wir uns über inverse Verteilungsfunktion generieren. Allgemein können wir dies formulieren zu

$$\begin{aligned} I &= \int_{-\infty}^{\infty} f(x)dx = \int_{-\infty}^{\infty} \tilde{f}(x)p(x)dx = \langle \tilde{f}(x) \rangle \\ \text{mit } \tilde{f}(x) &= f(x) \frac{1}{\lambda} e^{\lambda x} \quad \text{und } p(x) = \lambda e^{-\lambda x} \text{ Exponentialverteilung} \end{aligned} \quad (5.76)$$

Das Integral erhalten wir wieder wie in Gl. (5.73) durch Simulation und Bildung des diskreten Erwartungswertes, wobei allerdings diesmal x_i aus der exponentialen Verteilung gezogen wird.

Wir können auch den Ansatz derart modifizieren, dass wir anstelle der exponentiellen Dichte die Normalverteilung nehmen:

$$I = \int_{-\infty}^{\infty} f(x) dx = \int_{-\infty}^{\infty} \tilde{f}(x) p(x) dx = \langle \tilde{f}(x) \rangle \quad (5.77)$$

mit $\tilde{f}(x) = f(x) \sqrt{2\pi} e^{x^2/2}$ und $p(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ *Normalverteilung*

Entsprechend wird hier der Erwartungswert I mit den x_i aus $N(0,1)$ gebildet.

5.4 Der Simulationsrahmen

Nachdem wir die verschiedenen Eigenschaften, Erzeugung und Test der Zufallszahlen näher betrachtet haben, wollen wir in diesem Abschnitt den Simulationsprozess selbst genauer unter die Lupe nehmen. Üblicherweise wird der Auftrag für eine Simulation erteilt, um eine bestimmte Fragestellung zu beantworten. Dabei ist die richtige Beschreibung des Problems entscheidend für eine sachgerechte Lösung.

5.4.1 Problemidentifikation

Der realitätsgetreuen Problemdefinition stellen sich häufig viele sachfremde Faktoren in den Weg, auf die man sich vorbereiten sollte:

- Die Terminologie des Anwenders ist komplex und undurchsichtig.
- Die Wahrnehmung des Problems ist stark von den sehr unterschiedlichen Interessen innerhalb der Gruppe des Auftraggebers bestimmt, wie etwa von Vorurteilen, Erwartungen, Werten, Schuldzuweisungen, Egozentrik, Machtansprüchen usw.
- Es ist schwer, von dem, was die Auftraggeber als Problem ansehen, zu dem wirklichen Problem durchzudringen. Gibt man zu schnell auf und übernimmt die subjektive, meist egozentrische Position eines Anwenders ("Vorstand"), so können leicht andere wichtige Aspekte anderer Anwender vernachlässigt werden.
- Sind wir zu selbstsicher, so führt ein "Schnellschuss" leicht zu einer realitätsfernen Lösung.
- Auch übermäßige Betroffenheit bewirkt eine "betriebsblinde" Lösung.

Das Finden einer Problemlösung für eine Gruppe von Menschen (Firma, Verband, Behörde,...) hat wichtige soziale und gruppenpsychologische Implikationen, vor deren Hintergrund die technische Seite oft die einfachste ist und die eine sachge-

rechte Lösung manchmal sehr schwer machen. Dieser Tatsache muss man sich bewusst sein, bevor man mit einer Problemdefinition und –modellierung beginnt.

5.4.2 Modellspezifikation und –detaillierung

Der allgemeine Prozess zur Spezifikation einer Simulation lässt sich leicht mit dem allgemeinen Prozess zur Softwareerstellung vergleichen. In beiden Fällen durchlaufen wir verschiedene Phasen:

1. Am Anfang muss ein Glossar erstellt werden, das die Grundbegriffe des zu modellierenden Systems enthält und die Bedeutung und Wechselwirkung informell beschreibt.
2. In einem zweiten Schritt werden nun formale Modelle erstellt, um die Wechselwirkungen abzubilden. Dazu dienen sowohl UML-Diagramme als auch eine Modellierung der Eingaben, wie sie näher im nächsten Abschnitt erläutert wird.
3. Nach der Modellierung folgt die Evaluation des Modells.

Bei der Spezifikation der Modellgenauigkeit geht man von den Fragen aus, die an das Modell bei der Simulation gestellt werden sollen. Dabei geht man im *top down*-Modus vor, das heißt das Modell wird nur soweit detailliert, soweit es zur Beantwortung der Fragen (Implementierung alternativer Hypothesen, Messung von Effekten) notwendig ist. Die Gefahr einer zu starken Detaillierung sind vielfältig:

- Errichten wir ein zu detailliertes Modell ("*brute force*"), so ist die Anzahl der möglichen Simulationsläufe aus Rechenzeitgründen begrenzt und führt zu unsignifikanten stochastischen Ergebnisabschätzungen.
- Das Verhalten des Gesamtsystems wird bei hoher Detaillierung schwer verständlich und validierbar, so dass Fehler in der Modellierung schwer erkannt werden können.
- Ein zu hoher Detaillierungsgrad kann die Fertigstellung des Modells so stark verzögern, dass das Projekt erfolglos abgebrochen werden muss.

Gerade bei Simulationen für noch nicht existierende Realsysteme, die noch erhebliche Modifikationen erfahren können, ist es sinnvoll, den Detaillierungsgrad anfangs nicht zu groß zu wählen und erst später an die Fragestellung anzupassen.

5.4.3 Modellintegration realen Systemverhaltens

Für eine sachgerechte Modellierung ist es notwendig, die Beziehungen und Wechselwirkungen der einzelnen Objekte des Systems untereinander zu messen. Typischerweise können wir dies durch eine Regression erreichen (s. Kapitel 2 und 4), bei der wir die Einflüsse der verschiedenen Variablen testen und bewerten. Typische messbare Eingabegrößen einer diskreten, ereignisorientierten Simulation sind

- die Frequenz oder der Zeitabstand auftretender Ereignisse, etwa zwischen verschiedenen Eingabeereignissen oder zwischen der Eingabe und der Ausgabe
- eine nachgefragte Menge (Ausgabe) pro Zeiteinheit
- die Zeitdauern für die Dienstleistung in Warteschlangen, etwa bei der Bedienung von Menschen oder der Fertigung von Waren

Haben wir Warteschlangen im System, so müssen wir ermitteln, ob und welche Abweichung von dem klassischen Warteschlangenmodell hier vorliegt:

- Gibt es eine zentrale Warteschlange für mehrere Bedienstationen, oder hat jede Station ihre eigene?
- Ist es möglich, sich vorzudrängeln (*cheating*), die Warteschlange plötzlich zu wechseln (*swopping*), das Warten vorzeitig abubrechen (*reneging*) oder angesichts einer vollen Warteschlange sich gar nicht erst anzustellen (*balking*)?

Beachtete man diese Nebenbedingungen nicht, so schätzt man den Durchsatz vollkommen falsch ein.

Modellierung der Eingabeverteilungen

Zur Bestimmung der Verteilung der Eingabegrößen können wir auf eine Vielzahl möglicher Verteilungen zurückgreifen. Dabei sollten wir nicht nur die Verteilung wählen, die sich am Besten den gemessenen Daten anpassen lässt, sondern bei gegebener Auswahl diejenige wählen, deren Annahmen inhaltlich am Besten dem Modell entspricht:

- Beobachten wir eine reelle Variable, die als *Summe* von unabhängigen Zufallsgrößen entsteht, so ist eine Normalverteilung angebracht.
- Ist die beobachtete Variable x als *Produkt* unabhängiger Zufallsgrößen y_i zustande gekommen, so ist der Logarithmus der beobachteten Variablen

$$\ln x = \ln \prod_{i=1}^m y_i = \sum_{i=1}^m \ln y_i = \sum_{i=1}^n X_i \quad (5.78)$$

die Summe von Zufallsvariablen X_i und damit eine normalverteilte Zufallsvariable. Die Zufallsvariable x wird deshalb als "log-normal"-verteilt bezeichnet.

- Angenommen, es sind keine Daten über die fragliche Eingabevariable verfügbar. Etwa, weil das zu simulierende System in der Realität noch gar nicht gebaut wurde oder die Ereignisse zu selten sind und die Beobachtungszeit nicht ausreicht, dann fragen wir uns: was nun? In diesen Fällen können wir versuchen, durch Befragung von Experten eine Verteilung zu schätzen. Eine Möglichkeit besteht darin, eine Untergrenze a , eine Obergrenze b , und ein erwarteten Wert μ für die Variable abzuschätzen. Typisches Beispiel dafür ist die Zeit, eine Arbeit zu erledigen, oder die Zeit, in der ein Bauteil ausfällt oder seine Reparaturzeit. Die vermutete Verteilungsdichte kann dann als Dreiecksfunktion oder Trapezfunktion angesetzt werden, siehe Abb. 5.21.

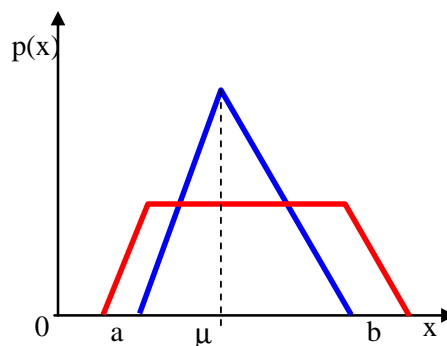


Abb. 5.21 Heuristische Verteilungen

Sind die Ober- und Untergrenzen zu vage, so lässt sich die Verteilung auch über die subjektive Einschätzung des Vertrauensintervalls $[x_a, x_{1-\alpha}]$ mit $\alpha = 5\%$ erreichen, in das nach Meinung der befragten Experten 90% aller Ereignisse fallen werden. Es ist meist einfacher, solche Grenzen abzuschätzen als die maximale oder minimale Grenze. Mit Hilfe der Angabe des Vertrauensintervalls können wir dann die Intervallgrenzen a, b rekonstruieren. Allerdings kann die heuristische Methode nur als ein Akt der Verzweiflung gewertet werden für den Fall, wenn keine Daten verfügbar sind. Haben wir Daten, so sollten wir andere Methoden anwenden.

Modelle von Ankunftsprozessen

Bisher betrachteten wir nur die Verteilung von kontinuierlich oder diskret verteilten zufälligen Ereignissen. Dabei nahmen wir an, dass die Einzelbeobachtun-

gen zwar einer bestimmten Verteilung gehorchen, aber trotzdem unabhängig von einander sind. In einer Vielzahl von Anwendungen ist diese Annahme aber nicht gerechtfertigt: Ein Liefertermin mag zufallsbehaftet sein, ist aber über die Lagerzeit und Bestellzeit abhängig vom Termin der vorherigen Lieferung. Ignorieren wir die Abhängigkeit der zufälligen Ereignisse von den Vorgängern, so ignorieren wir wichtige Modellierungsinformation. Im Folgenden betrachten wir deshalb mehrere Arten derartiger Abhängigkeiten.

Autoregressive Prozesse

Eines der bekanntesten Erzeugungsarten beschreibt den Einfluss früherer Beobachtungen (Autokorrelation) auf die jetzige Beobachtung x_i einer Zufallsvariablen X

$$x_i = \mu + a_1(x_{i-1} - \mu) + \dots + a_N(x_{i-N} - \mu) + \varepsilon_i \quad AR(N) \quad (5.79)$$

wobei μ der Erwartungswert der Zufallsvariablen X und ε_i eine $N(0, \sigma^2)$ -verteilte Zufallsvariable ist. Dieser *autoregressive Prozess* $AR(N)$ kann auch einen gleitenden Erwartungswert (*Moving Average*) haben und wird dann als **ARMA**-Prozess bezeichnet.

Möchte man nicht nur normal verteilte Zufallseinflüsse modellieren, sondern auch andere Zufallsprozesse mit zeit-konstanten Parametern ("stationäre Prozesse"), so kann man obigen Prozess an die Realität anpassen. Im Unterschied zu den Koeffizienten a_i , die man mit Hilfe der linearen Regression aus Kapitel 2 bestimmen kann, geht es hier um die stochastischen Proportionen von ε_i . Beobachten wir eine Verteilungsfunktion $F(x_i)$, so können wir über die Methode der inversen Verteilungsfunktion die Verteilung der x_i simulieren:

$$x_i = F^{-1}(\Phi(z_i)) \quad ARTA\text{-Prozess} \quad (5.80)$$

mit dem $N(0,1)$ -normalverteilten ARMA-Prozess $\{z_i\}$. Wir erhalten so die gewünschte marginale Verteilung für x_i . Dieser Prozess wird "*Autoregressive To Anything*" ARTA genannt, wobei die Hauptarbeit allerdings in der geeigneten Spezifikation und Anpassung der Koeffizienten a_i der autoregressiven Komponenten liegt, siehe (Carion und Nelson 1998).

Ankunftsprozesse

In vielen Simulationen wollen wir Ereignisse simulieren, deren Auftrittszeiten t_0, t_1, \dots diskret sind und die dabei einer bestimmten Verteilung unterliegen. Kann man das Auftrittseignis als Ankunft einer Ware in einem Lager, eines Kunden in einem Geschäft oder dgl. interpretieren, so spricht man von *Ankunftseignissen*.

DEF Ein stochastischer Prozess, der durch die Anzahl $N(t) = \max\{i \mid t_0 < t_i \leq t\}$ der Ereignisse gekennzeichnet ist, die vom letzten Zeitpunkt t_0 bis zum Zeitpunkt t auftreten, nennen wir einen *Ankunftsprozess*.

Ein Ankunftsprozess wird meist als Modell verwendet für Kunden, die für eine Dienstleistung anstehen müssen, etwa in Warteschlangen. Die Zeit $t_i - t_{i-1}$ zwi-

schen zwei Ankunftszeiten wird als "Zwischenankunftszeit" (*inter arrival time*) bezeichnet.

Der bekannteste und am meisten für Warteschlangen aller Art benutzte Ankunftsprozess ist der **Poisson-Prozess**. Er resultiert aus der Betrachtung von n Zeitunterteilungen. In jedem Zeitabschnitt Δt kommt mit Wahrscheinlichkeit p ein Kunde an oder nicht. Die resultierende Verteilung von k Kunden in n Zeitabschnitten ist eine *Binomialverteilung*. Für $\Delta t \rightarrow 0$ und $n \rightarrow \infty$ erhalten wir eine Poisson-Verteilung.

Der Poisson-Prozess ist dabei definiert durch folgende Bedingungen:

Sei die Anzahl der Kunden, die im Zeitabschnitt $[t, t+s]$ ankommen, mit $k = N(t+s) - N(t)$ bezeichnet. Dann gilt

- (1) Die Kunden kommen nach einander an.
- (2) Die Anzahl k ist unabhängig von der Anzahl $N(u)$, $0 \leq u \leq t$, die vorher angekommen waren.
- (3) Die Verteilung der Kundenzahl k ist unabhängig von t (Stationarität).

Diese drei Bedingungen sind nicht selbstverständlich. Beispielsweise trifft (1) nicht zu, wenn die Kunden in Gruppen eintreffen. Bedingung (2) kann verletzt werden, wenn die Kunden nicht so lange anstehen wollen und bei der Ankunft vor der Warteschlange bereits resignierend abrechnen und woanders hingehen. Die Proportion (3) schließlich besagt beispielsweise, dass am Vormittag genauso viele Kunden kommen wie am Nachmittag oder am Abend – ein in der Realität selten beobachtetes Phänomen. Trifft Bedingung (3) nicht zu, so sprechen wir von einem *nicht-stationären Poisson-Prozess*.

Trotzdem ist es sinnvoll, für kleine Zeitabschnitte die obigen drei Annahmen zu machen. Damit ergibt sich folgender Satz (s. Çinlar 1975)

SATZ Für einen Poisson-Prozess gilt die Wahrscheinlichkeit P für k Kunden

$$P(k) = \frac{(\lambda s)^k e^{-\lambda s}}{k!} \quad \text{mit} \quad \begin{array}{l} k = N(t+s) - N(t) \\ k = 0, 1, 2, \dots, \quad t, s \geq 0 \end{array} \quad (5.81)$$

mit dem Erwartungswert $\langle N(s) \rangle = \lambda s$.

In einer Zeiteinheit kommen $N(1) = \lambda$ Kunden an, also mit einer Zeit pro Kunde von $1/\lambda$. Damit gilt

SATZ Ist $\{N(t), t \geq 0\}$ ein Poisson-Prozess mit der Rate λ , so sind die Zwischenankunftszeiten Δt_i exponentiell verteilt mit dem Erwartungswert $1/\lambda$.

Die Umkehrung des Satzes ist ebenfalls korrekt:

Sind die Δt_i exponentiell verteilt, so liegt ein Poisson-Prozess vor.

Bei nicht-stationären Poisson-Prozessen ist die Rate eine Funktion der Zeit

$$\lambda = \lambda(t) \text{ mit der Stammfunktion } \Lambda(t) = \int_{t_0}^{t_1} \lambda(y) dy \quad (5.82)$$

Definieren wir uns

$$a(t,s) = \Lambda(t+s) - \Lambda(t) = \int_t^{t+s} \lambda(y) dy \quad (5.83)$$

Bei konstantem λ wird $a(t,s) = \lambda s$. Im nicht-stationären Fall müssen wir aber λs ersetzen durch $a(t,s)$

$$P(k) = \frac{a(t,s)^k e^{-a(t,s)}}{k!} \quad \text{mit} \quad \begin{array}{l} k = N(t+s) - N(t) \\ k = 0, 1, 2, \dots, \quad t, s \geq 0 \end{array} \quad (5.84)$$

Um die Funktion $\lambda(t)$ für die Realität abschätzen zu können, müssen geeignete Histogramme aufgestellt werden. Dabei muss die Zeit in kleine, aber auch wiederum nicht zu kleine Teilintervalle zerlegt werden.

Beispiel Ladengeschäft

Für ein Ladengeschäft soll $\lambda(t)$ geschätzt werden. Dazu messen wir an mehreren Tagen die Anzahl der Kunden, die innerhalb eines 10-Minuten-Intervalls eintreffen. Für jedes Intervall jeden Wochentags gibt es unterschiedliche Anzahlen. Wir mitteln die Anzahl für das Intervall einer Tageszeit über alle Wochentage und erhalten so einen tageszeitabhängigen Erwartungswert.

Wie modellieren wir einen Prozess, bei dem auch die Forderung (1) verletzt wird? Solche Prozesse, bei denen die Kunden in Gruppen erscheinen, beobachten wir bei allen Sportveranstaltungen, Buffet-Warteschlangen etc. bei denen ein fester Zeitpunkt, der für alle Kunden vorgegeben wird. Ein Ansatz, auch solche Situationen zu beschreiben, besteht darin, nicht die Einzelpersonen, sondern die Gruppen als Ereignisse zu betrachten. Die Zwischenankunftszeiten der Gruppen kann dann wie vorher mit einer exponentiellen Verteilung modelliert werden. Die Anzahl der Mitglieder einer Gruppe ist hier eine zweite Zufallsvariable, deren Verteilung über diskrete Histogramme selbst wieder approximiert werden kann. Die Anzahl der eintreffenden Kunden bis zum Zeitpunkt t_i ist also nach $N(t)$ Gruppen der Größe B_i

$$x(t) = \sum_{i=1}^{N(t)} B_i \quad t \geq 0 \quad (5.85)$$

Der Verbund der Zufallsvariable $\{x(t), t \geq 0\}$ mit dem Poisson-Prozess $\{N(t), t \geq 0\}$ wird "verbundener Poisson-Prozess" genannt.

5.5 Auswertung der Simulationsergebnisse

Alle Simulationen benötigen vor der Auswertung eine Definition oder Spezifikation, wie man die ursprüngliche Fragestellung mittels der Simulationsergebnisse beantworten will. Es ist wichtig, sich vor Simulationsbeginn darüber im Klaren zu werden, weil die gesamte Simulation davon beeinflusst wird. Üblicherweise definiert man sich dafür eine oder mehrere Leistungsvariable, und Leistungsmaße, die in der Simulation ermittelt und deren Größe bei Simulationsende im Sinne der Fragestellung interpretiert werden.

Bei diskreten, deterministischen Simulationen ist dies nicht schwer. Simuliert man beispielsweise Verkehrsströme, so ist die Auslastung der Verkehrsmittel bzw. Verkehrsstrassen eine klare Kenngröße für die Güte der Verkehrsregelung. Beachten wir aber die stochastische, zeitabhängige Natur der Verkehrsströme, so ist die Auswertung nicht mehr ganz so einfach: Neben dem Erwartungswert müssen wir auch die Varianzen und die Verteilung der Auslastung berücksichtigen. Dazu müssen wir die üblichen statistischen Tests anwenden.

Betrachten wir als Beispiel den Vergleich zweier alternativer Systeme von Warteschlangen.

5.5.1 Vergleich zweier Systeme

Das Ziel einer Simulation besteht üblicherweise nicht darin, eine absolute Aussage über beispielsweise die mittlere Zahl der Kunden in einer Warteschlange zu ermitteln, sondern darin, zwei verschiedene Strategien der Arbeitsbewältigung darin zu vergleichen, welche von beiden im Mittel die kleinere Anzahl von Kunden warten lässt. Die Maßzahlen für den Erfolg eines Systems sind also immer nur relativ zu einem anderen System.

Wollen wir die Leistung zweier Systeme vergleichen, so müssen wir sie zuerst ermitteln. Das Ergebnis ist aber von der Anzahl der Versuche abhängig.

Beispiel *Warteschlangen*

Ein System A aus zwei Servicestationen, z. B. Geldautomaten, soll mit einer doppelt so schnellen Servicestation B verglichen werden, die auch doppelt so teuer ist. Welche Lösung sollte man kaufen? Aus Datenerhebungen zu Stosszeiten wird sichtbar, dass die Kunden wie ein Poisson-Prozess mit einer Rate λ von 1 Kunde pro Minute eintreffen. Die Servicezeit eines der Server von A ist exponentiell verteilt mit dem Erwartungswert 1,8 Minuten, die von B mit 0,9 Minuten. Die Warteschlange wird im System A vom Pool beider Stationen bedient. Welches der beiden Systeme erzeugt die kürzere Warteschlange?

Entsprechend der Warteschlangentheorie lässt sich dies bei unserem einfachen Beispiel durch die mittleren Verzögerungen $d_A = 3,7$ min und $d_B = 4,13$ min ausrechnen. Die Erklärung dafür ist einfach: Die Bedienung durch zwei unabhängige Server ist kürzer, da beim Eintreffen eines Kunden in System A die

Bearbeitung auch beginnen kann, wenn noch ein anderer bearbeitet wird, was in System B nicht geht. Wissen wir nicht, dass hier Parallelarbeit möglich ist, und versuchen, auf Basis von Simulationen eine Entscheidung zwischen beiden Systemen zu treffen, so beobachten wir hohe stochastische Abweichungen und können dabei leicht eine falsche Entscheidung treffen. In Abb. 5.22 ist ein solcher Fall zu sehen, bei dem die mittleren Wartezeiten von 100 Kunden für beide Systeme eingetragen wurden. Hohle Kreise symbolisieren die Werte für System A, schwarz ausgefüllt wurden die Punkte für B. Dabei sind die Werte für verschiedene Anzahl n von Wiederholungen vertikal aufgetragen. Die theoretischen Erwartungswerte der Wartezeiten sind gepunktet für System A und durchgezogen für System B gezeigt.

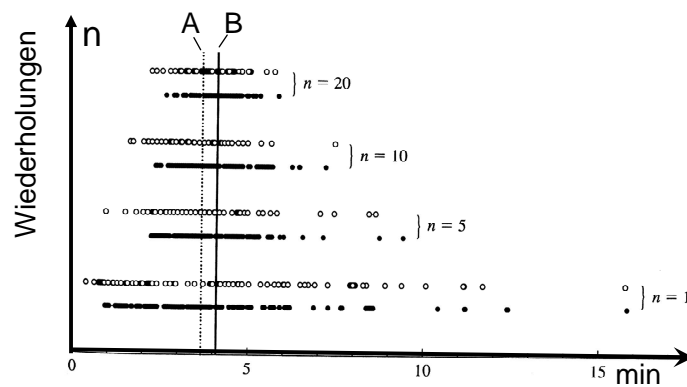


Abb. 5.22 Streubreiten der Wartezeiten von 100 Kunden

Mitteln wir die wiederholten Beobachtungen der Wartezeit pro Kunde, so reduziert sich die Varianz, da sich die Abweichungen der Beobachtungen gegenseitig aufheben und zum Erwartungswert der einzelnen Beobachtungen konvergiert.

Wir sehen, dass auch bei größerer Zahl von Wiederholungen durchaus nicht klar ist, welches der beiden Systeme besser ist, da die Varianz der Beobachtungen zu groß ist. Wie können wir die Varianz reduzieren? Dazu bilden wir die Differenz der Beobachtungen, im Beispiel die beobachteten Wartezeiten x_A und x_B eines Kunden i . Die Varianz der Differenz $z = x_A - x_B$ ist dann

$$\text{var}(z) = \langle (z - \bar{z})^2 \rangle = \langle z^2 \rangle - 2 \langle z \rangle \bar{z} + \bar{z}^2 = \langle z^2 \rangle - \bar{z}^2 \tag{5.86}$$

$$\begin{aligned} \text{var}(x_A - x_B) &= \langle (x_A - x_B)^2 \rangle - (\bar{x}_A - \bar{x}_B)^2 \\ &= \langle x_A^2 \rangle - \bar{x}_A^2 + \langle x_B^2 \rangle - \bar{x}_B^2 - 2 \langle x_A x_B \rangle + \bar{x}_A \bar{x}_B \end{aligned} \tag{5.87}$$

Mit der Definition der Kovarianz

$$\text{cov}(x_A, x_B) = \langle (x_A - \bar{x}_A)(x_B - \bar{x}_B) \rangle = \langle x_A x_B \rangle - \bar{x}_A \bar{x}_B \tag{5.88}$$

und der Definition der Varianz nach Gl. (5.86) erhalten wir für Gl.(5.87)

$$\text{var}(x_A - x_B) = \text{var}(x_A) + \text{var}(x_B) - 2\text{cov}(x_A, x_B) \quad (5.89)$$

Ist die Kovarianz größer als null, so wird die Varianz reduziert, etwa bei korrelierten Variablen. Wie erhalten wir korrelierte Ergebnisse? Beispielsweise dadurch, indem wir für alle Wiederholungen gemeinsame Zufallszahlen (*common random numbers* CRN) wählen: Alle Simulationen starten mit demselben Startwert des Zufallszahlengenerators.

5.5.2 Antithetische Zufallsvariable

Um die Varianz der Ergebnisse bei nur einem System mit Wiederholungen zu verkleinern kann man bereits bei der Erzeugung der Zufallszahlen ansetzen. Die Grundidee besteht darin, Paare von Wiederholungen für dasselbe System zu bilden und die Zufallszahlen dabei so zu wählen, dass die resultierenden Abweichungen bei einem Lauf genau entgegengesetzt sind von denen des Wiederholungslaufs. In der Summe beider Ausgaben zur Mittelung sollten dann die Abweichungen sich kompensieren. Mit Gl. (5.89) ist die Varianz der Summe bei $y_A = -x_B$

$$\text{var}(x_A + y_A) = \text{var}(x_A) + \text{var}(y_A) + 2 \text{cov}(x_A, y_A) \quad (5.90)$$

Sind die Zufallszahlen für x bzw. y *antithetisch*, so wird die Kovarianz der Zufallszahlen und damit der Ergebnisse x und y negativ und die Varianz der Summe wird reduziert.

Soweit, so gut – aber wie erhalten wir negativ korrelierte Zufallsvariable? Eine einfache Maßnahme besteht darin, für jede Zufallsvariable $u \in U(0,1)$ beim Wiederholungslauf die Variable $1-u$ zu verwenden. Beide Variablen sind uniform verteilt bei gleicher Wahrscheinlichkeitsdichte, aber sie sind negativ korreliert:

$$\text{cov}(u, 1-u) = \langle (u - \bar{u})((1-u) - (1-\bar{u})) \rangle = \langle u(1-u) \rangle - \bar{u}(1-\bar{u}) \quad (5.91)$$

Da $\bar{u} = 0,5$ für uniforme Zufallsvariablen aus $[0,1]$ ist

$$\begin{aligned} \text{cov}(u, 1-u) &= \int_0^1 u(1-u)p(u)du - \frac{1}{4} = \int_0^1 u - u^2 du - \frac{1}{4} \\ &= \frac{1}{2}u^2 - \frac{1}{3}u^3 \Big|_0^1 - \frac{1}{4} = \frac{1}{2} - \frac{1}{3} - \frac{1}{4} = -\frac{1}{12} \end{aligned} \quad (5.92)$$

Die Hoffnung ist, dass sich die negative Korrelation durch die Kette der Simulationsmechanismen durchzieht bis zum gemessenen Ergebnis und damit die Varianz des Ergebnisses geringer wird..

5.5.3 Synchronisierung der Zufallsvariablen

Sowohl bei der Differenzbildung bei zwei Systemen als auch bei der Summenbildung bei antithetischen Zufallsvariablen für paarweise Wiederholungen von einem System gehen wir davon aus, dass die CRN-Zufallsvariablen paarweise korrespondieren. Sowohl bei der Differenzbildung als auch bei den antithetischen Variablen nehmen wir an, dass die beobachteten Ergebnisse von jeweils den gleichen Mechanismen hervorgebracht wurden und zu positiv oder negativ korrelierten Messwerten führen. Diese Annahme ist aber nicht automatisch erfüllt, sondern muss inhaltlich gerechtfertigt werden. Implementieren wir eine alternative Strategie zu System A in System B, so kann dies aber zu einer größeren oder kleineren Anzahl von benötigten Zufallszahlen für Ankunftszeiten etc. führen. Würden wir nun beispielsweise für die antithetischen Zufallsvariablen im Code von System B bei gleichem Startwert für den uniformen Zufallszahlengenerator überall die Zufallsvariable u durch $1-u$ ersetzen, so würden korrespondierende Ereignisse nicht mehr die korrespondierenden Zufallszahlen erhalten und die resultierenden Messwerte wären nicht mehr negativ korreliert, so dass auch keine Verkleinerung der Varianz erfolgt. In Abb. 5.23 ist eine solche Situation abgebildet.

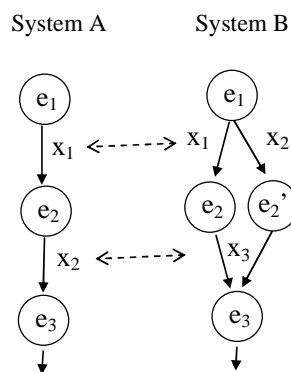


Abb. 5.23 Das Problem korrespondierender Zufallszahlen

Hierbei wird Ereignis e_2 mit zwei Zufallszahlen generiert; eine x_1 für die Auftrittszeit und eine x_2 für die Entscheidung über ein alternatives Ereignis. Damit wird der gesamte Ereignisstrom, der danach folgt, auf anderen Zufallszahlen beruhen. Wie lassen sich zwei Ströme von Zufallszahlen synchronisieren? Dafür gibt es mehrere Techniken:

- **Zusätzliche Zufallszahlen:** Man kann im Ablauf der Generatoren zusätzliche, nicht benutzte Zahlen erzeugen und darauf achten, dass korrespondierende Ereignisse auch korrespondierende Zufallszahlen erhalten und überflüssige Zufallszahlen nicht genutzt werden. Die Programmierung der Systeme A und B müssen also auf einander abgestimmt werden.

- Stromzuordnung (*stream dedication*): Man kann für jede Art von Zufallszahlen, z.B. uniforme und exponentiell verteilte Zufallszahlen, einen getrennten, von einander unabhängigen Generator wählen. Werden dann in Version B zusätzliche Entscheidungen und damit zusätzliche Zufallszahlen fällig (etwa, welcher Art das Ereignis sein soll), so wird die Erzeugung der korrespondierenden Ankunftszeit nicht beeinflusst. Auch gesonderte Generatoren für Servicezeiten sind sinnvoll, um sowohl die Synchronität zu erhalten als auch zu verhindern, dass bei sehr vielen Ereignissen die Periode des Generators überschritten wird und unbeabsichtigte Wiederholungen auftreten.
- Wir können für Ereignisse beim Lauf von System A die korrespondierenden Zufallszahl speichern (z.B. als Attribut des Ereignisses) und sie für den Lauf von System B wieder verwenden.

Bei all diesen Techniken ist der Erfolg aber nicht garantiert, da alternative Strategien über Zeitverzögerung von Ereignissen etc. die Korrelationen beeinflussen und damit teilweise sogar den gegenteiligen Effekt, eine Vergrößerung der Varianz, erzeugen können.

5.5.4 Auswertung mit Konfidenzintervallen

Für den Vergleich der Leistungsmaße zweier alternativer Systeme bietet es sich an, nicht die Einzelmessungen y_A und y_B , sondern die Variable aus der Differenz beider Systeme

$$z = y_A - y_B$$

zu analysieren. Ist der Erwartungswert null („Nullhypothese“), was wir mit unserem Konfidenzintervall $[z_{\alpha}, z_{1-\alpha}]$ des statistischen Tests nachprüfen können, so gibt es keinen Unterschied. Der Test sagt uns dann nicht nur, ob die Systeme mit Wahrscheinlichkeit $(1-2\alpha)$ gleich sind, sondern bei Ungleichheit auch, welches besser ist. Die direkte Differenzbildung verhindert auch die Notwendigkeit, vorher Erwartungswerte zu bilden, die ebenfalls verglichen werden müssen.

Das Konfidenzintervall für zwei Systeme mit gleicher Wiederholungsanzahl $n_1 = n_2 = n$ lässt sich folgendermaßen bilden. Mit dem geschätzten Erwartungswert

$$\bar{z}(n) = \frac{1}{n} \sum_{i=1}^n z_i \quad (5.93)$$

erhalten wir die geschätzte Einzelvarianz s^2/n der n unabhängigen Zufallszahlen z_i

$$\frac{s^2}{n} = \frac{1}{n} \frac{1}{n-1} \sum_{i=1}^n (z_i - \bar{z}(n))^2 \quad (5.94)$$

Da bei kleinen Werten von n noch keine Normalverteilung der Erwartungswerte vorliegt, verwenden wir eine t-Verteilung mit $n-1$ Freiheitsgraden und bilden die kritischen Punkte $t_{n-1,\alpha}$ und $t_{n-1,1-\alpha}$ an den Rändern der zentrierten t-Verteilung mit Varianz 1. Wir erhalten dann das Konfidenzintervall durch eine Skalierung auf Varianz s^2/n

$$z_{1,2} = \bar{z}(n) \pm t_{n-1,1-\alpha} \sqrt{s^2/n} \quad (5.95)$$

Vergleichen wir ein simuliertes System mit einem realen System, so haben wir aber beim realen System meist eine sehr viele kleinere Anzahl von Beobachtungen. Mit $n_1 \neq n_2$ ist die Voraussetzung für Gl. (5.95) nicht erfüllt. Das Konfidenzintervall für diesen Fall müssen wir deshalb anders bilden.

Ausgehend von den geschätzten Erwartungswerten \bar{x}_A und \bar{x}_B der beobachteten Messwerte x_A und x_B , die wir analog zu Gl.(5.93) bilden, können wir auch die Varianzen $\sigma_A^2 = s_A^2/n_1$ und $\sigma_B^2 = s_B^2/n_2$ der Ergebnisse beider Simulationsarten analog zu Gl.(5.94) bilden. Allerdings ist hier die Anzahl der Freiheitsgrade der t-Verteilung nur approximativ bestimmbar (Welch 1928) zu

$$\hat{f} = \frac{(\sigma_A^2 + \sigma_B^2)^2}{\sigma_A^4/(n_1 - 1) + \sigma_B^4/(n_2 - 1)} \quad (5.96)$$

und das Konfidenzintervall der Akzeptanz mit Wahrscheinlichkeit $(1-2\alpha)$ ist

$$z_{1,2} = (\bar{x}_A(n_1) - \bar{x}_B(n_2)) \pm t_{\hat{f},1-\alpha} \sqrt{\sigma_A^2 + \sigma_B^2} \quad (5.97)$$

und wird als *Welch Konfidenzintervall* bezeichnet.

Literaturverzeichnis

- Armstrong J. S. (ed.), Principles of Forecasting. Norwell, MA: Kluwer Academic Press 2001
- Alexandridis A., Sarimveis H., Bafas G.: *A new algorithm for online structure and parameter adaption of RBF networks*, Neural Networks 16, pp.1003-1017 (2003)
- Banks J, Carson JS, Nelson B, Nicol D : Discrete –Event System Simulation; Prentice-Hall, Englewood Cliffs, NJ, 4th ed. 2004
- Benninga S: Financial Modeling, The MIT Press , 2000, 2nd, 640 S.
- Biethahn J, Hummeltenberg W, Schmidt B, Stähly P, Witte T, Simulation als betriebliche Entscheidungshilfe. *State of the Art und neuere Entwicklungen*, Physica-Verlag Heidelberg, 1999
- Bossel, H: Systeme, Dynamik, Simulation. Modellbildung, Analyse und Simulation komplexer Systeme. BoD GmbH, Norderstedt (Mai 2004)
- Bossel, H: Systemzoo 1. Elementarsysteme, Technik und Physik. BoD GmbH, Norderstedt (Juni 2004)
- Bossel, H: Systemzoo 2. Klima, Ökosysteme und Ressourcen. BoD GmbH, Norderstedt (August 2004)
- Bossel, H: Systemzoo 3 Wirtschaft, Gesellschaft und Entwicklung BoD GmbH, Norderstedt (September 2004)
- Bossel, H: Modellbildung und Simulation. Konzepte, Verfahren und Modelle zum Verhalten dynamischer Systeme, Vieweg Braunschweig 1994, 2. Aufl.
- Bomberger J.D., Seborg D.E., Ogunnaike B.A.: RBFN Identification of an industrial polymerization reactor model, in: Mujtaba I.M., Hussain M.A.: Application of Neural Networks and Other Learning Technologies in Process Engineering, Imperial College Press, London 2001, pp. 23-48
- Broy M, Steinbrüggen R.: Modellbildung in der Informatik, m. CD-ROM, Springer Verlag
- Cardaliaguet P., Euvrard G.: *Approximation of a Function and its Derivative with a Neural Network*; Neural Networks, Vol.5, pp.207-220 (1992)
- Distelkamp M, Hohmann F, Lutz C, Meyer B., Wolter M.: Das IAB/INFORGE-Modell, in: Beiträge zur Arbeitsmarkt- und Berufsforschung BeitrAB275, Institut für Arbeitsmarkt- und Berufsforschung der Bundesanstalt für Arbeit (IAB), Nürnberg 2003
- Fair, Ray C: Specification, Estimation and Analysis of Macroeconomic Models, Cambridge
- Forrester, Jay: Industrial Dynamics. MIT Press, Cambridge MA 1961
- Forrester, Jay: Urban Dynamics; 1969, ISBN 1-56327-058-7
- Forrester, Jay: World Dynamics (second edition), Wright-Allen Press, 1971
- Forrester, Jay: The collected papers of Jay W. Forrester, Wright-Allen Press, 1975

- Franses, P. H. & D. J. C. Van Dijk (2000), *Non-linear Time Series: Models in Empirical Finance*. Cambridge, U. K.: Cambridge University Press.
- Francois j: Flexible Estimation and Interference Within General Equilibrium Systems, 4th Ann. Conf. Global Economic Analysis, Purdue 2001. (<http://www.intereconomics.com/francois/Text2000-010.pdf>)
- Gaab W; Heilemann U; Wolters J (Eds.) *Arbeiten mit ökonomischen Modellen*, Series: Studies in Contemporary Economics, Physica Verlag, 2004, VI, 275 S. ISBN: 3-7908-0154-2
- Greene, W.H.: *Econometric Analysis*, 5th Ed., Prentice Hall, Upper Saddle River NJ, USA 2003
- Gilli M: Causal Ordering and Beyond. *Int. Economic Review* 33, 957-971, 1992
- Gilli M: The Logical Structure of a Model, in: (Gaab, Heilemann, Wolters 2004), 233-257
- Göldner: *Systemanalyse (Bd.1-3)*. Harri Deutsch Verlag
- Hinton G., Williams C., Revow M.: *Combining Two Methods of Recognizing Hand-Printed Digits*; in: I. Aleksander, J. Taylor (Eds), *Art. Neural Systems* 2, Elsevier Sc. 1992, pp.53-60
- Ism: *Introducing Statistical Methods*, ISBN 0761951415
- Jetschke, Gottfried: *Mathematik der Selbstorganisation*, Verlag Harri Deutsch, Frankfurt 1989
- Kreft I., de Leeuw J: *Introducing Multilevel Modeling*, Sage Publications 1998
- Law A., Kelton D: *Simulation Modeling and Analysis*, McGraw Hill Companies, Singapore 2000
- Lee S. and Kil R.: *A Gaussian Potential Function Network With Hierarchically Self-Organizing Learning*; *Neural Networks*, Vol. 4, pp. 207-224, 1991
- Liebl F: *Simulation*, Oldenbourg Verlag, München 1995, 2.Aufl, ISBN 3-486-23373-4
- Ljung L, Glad T: *Modeling of Dynamic Systems*, Prentice Hall 1994, ISBN: 0135970970
- Luke D: *Multilevel Modeling*, Sage University Papers Series, Sage Publications, Thousand Oaks, CA 91320, USA 2004
- Meadows D.H., Meadows Donella, Zahn Erich, Milling Peter: *Die Grenzen des Wachstums; Bericht des Club of Rome zur Lage der Menschheit*, Rowohlt Verlag, Reinbek 1973
- Meadows, D.H., Meadows, D.L., Randers, J.,: *Beyond the limits*. Chelsea Green, Post Mills VT, 1992
- Meadows, D.H., Meadows, D.L., Randers, J.,: *Die neuen Grenzen des Wachstums*, DVA, Stuttgart 1992.
- Meinhard, Hans: *The Algorithmic Beauty of Sea Shells*, Springer Verlag Heidelberg, 3rd ed. 2003 (*dtsch: Wie Muscheln sich in Schale werfen*)
- Meyer B. et al. (1999): *Marktkonforme Umweltpolitik. Wirkungen auf Luftschadstoffemissionen, Wachstum und Struktur der Wirtschaft*. Heidelberg.
- Molnár P: *Modellierung und Simulation der Dynamik von Fußgängerströmen*, Shaker Verlag 1996
- Scherf, Helmut E.: *Modellbildung und Simulation dynamischer Systeme*, Oldenbourg (Oktober 2004)
- H. Shioler, U. Hartmann: *Mapping Neural Network Derived from Parzen Window Estimator*; *Neural Networks*, Vol.5, pp.903-909 (1992)

- Schnur P., Zika G.: Längerfristige Arbeitskräftebedarfsprojektion und Politiksimulation, in: Beiträge zur Arbeitsmarkt- und Berufsforschung BeitrAB250, Institut für Arbeitsmarkt- und Berufsforschung der Bundesanstalt für Arbeit (IAB), Nürnberg 2002
- Specht D.: *A General Regression Neural Network*; IEEE Transactions on Neural Networks, Vol.2, No.6, Nov. 1991, pp.568-567
- Specker A: Modellierung von Informationssystemen. Ein methodischer Leitfaden zur Projektabwicklung
- Uebe G: World of Economic Models; Aldershot 1995
- Vu, Esfandiari: Dynamic Systems. MacGraw Hill, 1998
- Welch B.L.: The Significance of the Difference Between Two Means When the Population Variances are Unequal, *Biometrika*, 25:350-362 (1938)
- Zobrist George W., Leonard James V. (Hrsg).: Simulation Systems CRC Press, 2000, 332 Seiten, ISBN 9056996665

In der Bibliothek:

Giloi, Wolfgang K.: Principles of continuous system simulation
Teubner, Stuttgart 1975, 172 S.

ISBN : 3-519-02336-9

I.6-3

Payne, James A.: Introduction to simulation
programming techniques and methods of analysis
McGraw-Hill New York [u.a.], 1982, 324 S.

ISBN : 0-07-048945-9

I.6-5

Simulation technischer Systeme 1, 1974, 291 S.

ISBN : 3-446-11883-7

I.6-8

Simulation technischer Systeme 2, 1976, 663 S.

ISBN : 3-446-11884-5

I.6-9

Fishman, George S.: Principles of discrete event simulation

Wiley: New York [u.a.], 1978, 514 S.

Wiley series on systems engineering and analysis

ISBN : 0-471-04395-8

I.6-11

Bossel, Hartmut: Simulation dynamischer Systeme

Grundwissen, Methoden, Programme

Vieweg, Braunschweig 1989, 310 S.

ISBN : 3-528-04746-1

I.6-13

Bossel, Hartmut: Modellbildung und Simulation

Konzepte, Verfahren und Modelle zum Verhalten dynamischer Systeme ; ein Lehr- und Arbeitsbuch

Vieweg Braunschweig 1994, 2., veränd. Aufl., mit verb. Simulationssoftware, 402 S. : graph. Darst. + 1 Diskette

ISBN : 3-528-15242-7

I.6-19 I.6-19 DSK normal; I.6-19 ex 2 R7.09.2028; I.6-19 ex 2 DSK R26.05.2030;

Canty, Morton J.: Chaos und Systeme

Eine Einführung in Theorie und Simulation dynamischer Systeme

Vieweg, Braunschweig [u.a.] 1995, 174 S. + Diskette

ISBN : 3-528-05469-7

I.6-15 R27.07.2023;

Zeigler, Bernard P.: Theory of modelling and simulation

Krieger, Malabar, Fl. 1984, 435 S.

ISBN : 0-89874-808-9

I.6-16

Law, Averill M. | Kelton, W. David: Simulation modeling and analysis

McGraw-Hill Boston [u.a.] 2000, 3. ed., 760 S

ISBN : 0-07-100803-9

I.6-20 R20.10.2028; I.6-20 ex 2;

Liebl, Franz: Simulation

problemorientierte Einführung

Oldenbourg, München [u.a.] 1995, 273 S. 2., überarb. Aufl,

ISBN : 3-486-23373-4

I.6-18; I.6-18 ex 2 R7.09.2028;

Links

<http://www.pestel-institut.de>

Methodische Klammer über alle Forschungsbereiche und Themen ist die **Systemanalyse** und als Werkzeug das (Computer-) Modell. Mit dessen Hilfe wird zunächst die Entwicklung der Vergangenheit beschrieben. Bildet das Modell diese durch Daten gut ab, können Prognosen mit einem Zeithorizont bis zu 15 Jahren vorgenommen werden.

Beispiele für Forschungsarbeiten mit dieser Methode sind:

- Regionalstudien:

Entwicklung regionaler Prognosen für Bevölkerung, Wohnungsbedarf, Wirtschaftsentwicklung und Arbeitsmarkt für einzelne Städte, Landkreise oder Bundesländer.

- Abschätzung der Perspektiven von Großinvestitionsprojekten:

Wie entwickelt sich die Wirtschaftlichkeit von z. B. multifunktionalen Stadien, von Altenheimen oder von Klinikprojekten? Diese Modellrechnungen gehen über die üblichen Wirtschaftlichkeitsrechnungen weit hinaus und berücksichtigen auch die Einflußfaktoren auf die zukünftige Nachfrage nach den jeweils angebotenen Dienstleistungen, auf das Konkurrenzumfeld sowie auf den Kapitalmarkt.

- Analyse von Ursache- und Wirkungs-Zusammenhängen:

Was bestimmt die Preisentwicklung von Wohnbauland? Welche Einflüsse wirken auf die Nachfrage nach Einfamilienhäusern ein? Mit Hilfe der Systemanalyse können Marktbewegungen transparent gemacht werden und die Auswirkungen der Veränderungen bis in die Betriebswirtschaft der Unternehmen verfolgt werden.

Club of Rome

<http://www.clubofrome.de/sektion/dscor.html>

HAUS RISSEN, Internationales Institut für Politik und Wirtschaft, in Hamburg

<http://www.hausrissen.org/>

Frederic Vester: Die Kunst vernetzt zu denken

Welt-Wirtschaftsmatrix

<http://www.hwwa.de/wmatrix/Home.html>

Datenbanken

<http://www.hwwa.de/migration/>

<http://www.hwwa.de/EU-Beitrittslaender/>

Götz Uebe's List Of Macroeconomic Models,

<http://www.unibw-hamburg.de/WWEB/math/uebe/modelle/titelseite.html>

Arbeit und Steuern-Gleichgewichtsmodell

http://www.hwwa.de/Projekte/Forsch_Schwerpunkte/FS/Grundlagen%20Konjunkturanalyse/FP%20611.htm

Zika: <http://www.iab.de/asp/internet/publikationenByMit.asp?Mitarbeiter=167>

Publikationen-Modellierung:

<http://www.iab.de/asp/internet/dbprjDoku.asp?pkyProjekt=i880215a02>

Statistik-Bundesamt: <http://www.destatis.de/allg/d/veroe/inoutputueb.htm>

Websites Modellierung

<http://clear.eawag.ch/>

Klimamodellierung für die Schweiz

Diskrete Modellierung

Axel Jantsch: Modeling Embedded Systems and Soc's, concurrency and time in models of computation, Morgan Kaufmann, San Francisco 2004, ISBN 1-55860-925-3, *Stark theorielastige Einführung in diskreter Systeme, Markovketten usw.*

Abbildungsverzeichnis

Abb. 1.1 black box, white box und grey box -Verhalten.....	1-3
Abb. 2.1 Methoden vorausschauender Modellierung (nach [Arm01]).....	2-1
Abb. 2.2 Makroökonom. Input-Output-Tabelle der Konsumverflechtung.....	2-2
Abb. 2.3 Lineare Approximation von Daten.....	2-4
Abb. 2.4 Lineare Approximation einer verrauschten Funktion.....	2-7
Abb. 2.5 Unterschiede in der Parameterverteilung.....	2-12
Abb. 2.6 Rechts- und linksseitiger Hypothesentest.....	2-15
Abb. 2.7 Formales Neuron.....	2-18
Abb. 2.8 Ein zweischichtiges Netz.....	2-19
Abb. 2.9 Grundstruktur des Backpropagation-Netzwerks.....	2-21
Abb. 2.10 Fehlerrückführung im Mehrschichtensystem.....	2-22
Abb. 2.11 Die Gradientensuche.....	2-22
Abb. 2.12 Propagierung der Aktivität von Neuron i zu Neuronen k.....	2-24
Abb. 2.13 Generalisierung und Überanpassung bei einer Funktionsapproximation.....	2-26
Abb. 2.14 Überanpassung beim Lernen.....	2-27
Abb. 2.15 Das Netzwerk normierter Basisfunktionen.....	2-29
Abb. 2.16 Einfache (a) und normierte Netzaktivität (b) aus drei RBF Neuronen.....	2-31
Abb. 2.17 Transformation auf Normalverteilung.....	2-33
Abb. 2.18 Überdeckung des Eingaberaums.....	2-36
Abb. 2.19 Annäherung von Anzahl und Überdeckung von Glockenfunktionen.....	2-37
Abb. 2.20 Die Reaktorsimulation.....	2-39
Abb. 2.21 RUS Strategie: Ausgabe (oben) und Eingabe (unten).....	2-40
Abb. 2.22 RDS Strategie: Ausgabe (oben) und Eingabe (unten).....	2-41
Abb. 2.23 Ergebnisse der RBF-Approximation mit unterschiedlich breiten Zentren.....	2-41
Abb. 2.24 Ergebnisse der RBF-Approximation.....	2-42
Abb. 3.1 Der Wirkungsgraph des Antarktis-Energiemodells.....	3-3
Abb. 3.2 Wirkungsgraph des Weltmodells.....	3-6
Abb. 3.3 Gewichteter Wirkungsgraph des Weltmodells und seine Wirkungsmatrix.....	3-7
Abb. 3.4 Entwicklung der Umweltbelastung U mit der Zeit.....	3-11
Abb. 3.5 Systemzustandsänderung durch direkte Rückkopplung.....	3-12
Abb. 3.6 Systemgraph der differentiellen Kopplung des Weltmodells.....	3-13
Abb. 3.7 Entwicklung der Umweltbelastung U mit der Zeit.....	3-14
Abb. 3.8 Das Systemdiagramm des Subsystems 1 und seine Teilgleichungen.....	3-17
Abb. 3.9 Das Systemdiagramm beider Gleichungen.....	3-18
Abb. 3.10 Visualisierung der Entscheidung sowie ihre Definition.....	3-18
Abb. 3.11 Das Subsystem 2 und seine Zusammenfassung.....	3-19
Abb. 3.12 Systemdiagramm des Subsystems 3.....	3-20
Abb. 3.13 Das aus den Subsystemen zusammengekoppelte Gesamtsystem.....	3-21
Abb. 3.14 Das veränderte Gesamtsystem.....	3-22
Abb. 3.15 Systementwicklung bei (a) $C_F = 0.0$ und (b) $C_F = 0.03$	3-23
Abb. 3.16 Systementwicklung bei $C_F = 1.0$	3-23
Abb. 3.17 Allgemeines Diagramm eines dynamischen Systems.....	3-25

Abb. 3.18 Wasserstand und Zustandsvariable bei einer Badewanne.....	3-27
Abb. 3.19 Das integrative Element im Systemdiagramm	3-27
Abb. 3.20 Systemdiagramm eines Systems ohne Zustandsvariable	3-30
Abb. 3.21 Systemzustandsänderung durch verzögerte Rückkopplung.....	3-30
Abb. 3.22 Visualisierung der Differenzialgleichung	3-32
Abb. 3.23 Logistische Wachstumskurve für $z^*=1$	3-33
Abb. 3.24 Lineares System für zwei Zustandsvariable.....	3-35
Abb. 3.25 2-dim Phasendiagramm eines linearen Systems	3-37
Abb. 3.26 2-dim Phasendiagramme eines linearen Systems	3-39
Abb. 3.28 Populationsentwicklung bei Räuber-Beute Abhängigkeit	3-40
Abb. 3.29 Phasendiagramme bistabiler Schwinger	3-41
Abb. 3.30 Pigmentierung bei der Muschel <i>Lyria planicostata taiwanica</i>	3-42
Abb. 3.31 Ausprägung der Aktivität entlang einer Wachstumslinie x	3-44
Abb. 3.32 Die Zufluss, Abfluss und Gleichgewicht beim Speicher	3-48
Abb. 3.33 Der Gleichgewichtszustand als Folge gegenläufiger Mechanismen	3-49
Abb. 3.34 Verhalten des Systems bei Preisanstieg bei $c_1 = 1, c_2 = 2$	3-50
Abb. 3.35 Beispiel eines Gleichgewichtsmodells (aus [IAB253])	3-51
Abb. 3.36 Die Newton-Raphson Approximation.....	3-53
Abb. 3.37 Eingabe und Ausgabe der Papierherstellung (nach Ljung 1994).....	3-55
Abb. 3.38 Verlauf der Zielfunktion in Abhängigkeit des Parameters c	3-56
Abb. 3.39 Arbeitsschritte einer Gleichgewichtsanalyse (aus [IAB253]).....	3-57
Abb. 4.1 Der hierarchische Aufbau eines globalen Wirtschaftssimulationsmodells	4-2
Abb. 4.2 Das INFORGE-Modell (aus Schnur 2002)	4-3
Abb. 4.3 Aufteilung der Produktion in Untergruppen.....	4-3
Abb. 4.4 Die Wirkungsstruktur der Ländermodelle (nach Distelkamp 2003).....	4-4
Abb. 4.5 Die Input-Output-Tabellenstruktur für die Güternachfrage/Produktion	4-5
Abb. 4.6 Die vielschichtige Zusammensetzung gesellschaftlichen Verhaltens	4-7
Abb. 4.7 Der mittlere Spendeneinfluss je nach US-Staat (nach Luke 2004).....	4-9
Abb. 4.8 Der mittlere Spendeneinfluss je nach Tabakernte (in Tausend acres)	4-11
Abb. 4.9 (a) Graph und (b) korrespondierende Adjazenzmatrix	4-15
Abb. 4.10 Die streng zusammenhängende Komponente mit V_2	4-15
Abb. 4.11 Ermittlung der nächsten streng zusammenhängenden Komponente.....	4-16
Abb. 4.12 Reduzierter Graph der Blöcke und seine Adjazenzmatrix.....	4-16
Abb. 4.13 (a) angeordneter Graph (b) Adjazenzmatrix.....	4-17
Abb. 4.14 (a) Minimalgraph (b) Erreichbarkeitsmatrix.....	4-19
Abb. 4.15 Essentielle Mengen und bipartite Graphen.....	4-20
Abb. 4.16 Der Gesamtgraph aus Kopien des Graphen zu verschiedenen Zeitpunkten. ...	4-22
Abb. 4.17 Erreichbarkeitsmatrizen C_0, C_1, C_2	4-22
Abb. 4.18 Variablensubstitution (a).....	4-23
Abb. 4.19 Variablensubstitution (b).....	4-23
Abb. 4.20 Zusammenfassung zweier rückgekoppelter Variabler i und k	4-24
Abb. 4.21 Vereinfachung der Einflüsse bei Submodellen.....	4-24
Abb. 5.1 Zeitlicher Ablauf von kausalen Eingaben.....	5-4
Abb. 5.2 Diskretisierung in der Zeit.....	5-5
Abb. 5.3 Ereignisanordnung in der Zeit.....	5-8
Abb. 5.4 Modell des Supermarkts.....	5-9
Abb. 5.5 Kausalkette der Kundenbedienung.....	5-10
Abb. 5.6 Datenstruktur für die Ereignisanordnung in der Zeit.....	5-11
Abb. 5.7 Datenstruktur für die Ereignisanordnung in der Zeit.....	5-11
Abb. 5.8 Die Wahrscheinlichkeitsdichte einer uniformen Verteilung.....	5-13
Abb. 5.9 Histogramm einer Beobachtung einer uniformen Verteilung.....	5-15

Abb. 5.10 Transformation durch inverse Verteilung.....	5-21
Abb. 5.11 Exponentielle pdf, cdf und inverse Funktion.....	5-22
Abb. 5.12 Inversion der diskreten Verteilung	5-23
Abb. 5.13 Laplace-Verteilungsdichte.....	5-25
Abb. 5.14 Pseudocode für normal verteilte Zufallszahlen	5-28
Abb. 5.15 Abhängigkeitsplot der Zufallsvariablen	5-30
Abb. 5.16 3D-Abhängigkeitsplot der Zufallsvariablen	5-30
Abb. 5.17 Leistungsschema der Simulation	5-33
Abb. 5.18 Die hit-and-miss Simulation zur Integralmessung.....	5-34
Abb. 5.19 Buffons Problem	5-35
Abb. 5.20 Auswahlregel und Lage des Stabes	5-36
Abb. 5.21 Heuristische Verteilungen	5-42
Abb. 5.22 Streubreiten der Wartezeiten von 100 Kunden.....	5-47
Abb. 5.23 Das Problem korrespondierender Zufallszahlen.....	5-49

Anhang A Der Fixpunktalgorithmus für TMSE

Behauptung:

Die stochastische Iterationsgleichung

$$\mathbf{a}(t+1) = \tilde{\mathbf{a}}(t+1) / |\tilde{\mathbf{a}}(t+1)| \quad \text{mit} \quad \tilde{\mathbf{a}}(t+1) = \mathbf{a}(t) - \mathbf{x}\mathbf{x}^T \mathbf{a}(t) \quad (\text{A.1})$$

bzw. ihr Erwartungswert

$$\mathbf{a}(t+1) = \tilde{\mathbf{a}}(t+1) / |\tilde{\mathbf{a}}(t+1)| \quad \text{mit} \quad \tilde{\mathbf{a}}(t+1) = \mathbf{a}(t) - \mathbf{A}_{\mathbf{x}\mathbf{x}} \mathbf{a}(t) \quad (\text{A.2})$$

mit dem Erwartungswert $\langle \mathbf{x}\mathbf{x}^T \rangle = \mathbf{A}_{\mathbf{x}\mathbf{x}}$, der Autokorrelationsmatrix, konvergiert zum Eigenvektor der Matrix $\mathbf{C}_{\mathbf{x}\mathbf{x}}$ mit dem kleinsten Eigenwert.

Beweis:

Die Iteration bewirkt, dass bei jedem Iterationsschritt der Parametervektor normiert bleibt

$$\hat{a}_i = \frac{a_i(t)}{|\mathbf{a}(t)|} \Rightarrow \sum_i \hat{a}_i^2 = \frac{1}{|\mathbf{a}(t)|^2} \sum_i a_i^2 = \frac{\mathbf{a}^2}{|\mathbf{a}|^2} = 1 \quad (\text{A.3})$$

Was ist das Konvergenzziel der Lernregel (A.2) mit normiertem \mathbf{a} ? Betrachten wir dazu eine Zielfunktion R , die die Lernregel (A.2) als stochastische Gradientenabstiegsregel hat.

$$\mathbf{a}(t+1) = \mathbf{a}(t) - \mathbf{A}_{\mathbf{x}\mathbf{x}} \mathbf{a}(t) = \mathbf{a}(t) - \frac{\partial R(\mathbf{a})}{\partial \mathbf{a}} \quad (\text{A.4})$$

Wie wir aus Kapitel 2.6.2 wissen, konvergiert der Gradientenabstieg immer. In diesem Fall ist mit dem Gradientenabstieg das Konvergenzziel durch das Minimum der Zielfunktion gegeben. Wie man durch partielles Ableiten leicht nachprüfen kann, lautet mit dem Gradienten

$$\frac{\partial R(\mathbf{a})}{\partial \mathbf{a}} = \mathbf{A}\mathbf{a} \quad (\text{A.5})$$

die Zielfunktion mit der Nebenbedingung des normierten Parametervektors

$$R(\mathbf{a}) = \frac{1}{2} \mathbf{a}^T \mathbf{A}\mathbf{a} \quad \text{mit} \quad |\mathbf{a}| = \text{const} = 1 \quad (\text{A.6})$$

Was ist das Minimum dieser Zielfunktion, das mit dem Gradientenalgorithmus erreicht wird? Das Problem, den Extremwert einer Funktion unter Nebenbedingungen zu finden, wird durch die Methode der „Lagrange'schen Parameter“ erreicht. Dazu gehen wir folgendermaßen vor:

Zuerst bilden wir eine Hilfsfunktion L , die aus der zu maximierenden Funktion und der mit einem Hilfsparameter μ gewichteten Nebenbedingung besteht. Die Nebenbedingung ist in einer Form aufgeschrieben, die null ergibt, hier: $\mathbf{a}^2 - 1 = 0$.

$$L(\mathbf{a}, \mu) = R(\mathbf{a}) + \mu (\mathbf{a}^2 - 1) \quad (\text{A.7})$$

Die für ein Extremum notwendigen Bedingungen sind die Ableitungen dieser Hilfsfunktion

$$\frac{\partial L}{\partial \mathbf{a}} = 0 \quad \text{und} \quad \frac{\partial L}{\partial \mu} = 0 \quad (\text{A.8})$$

In unserem Fall ist dies

$$\frac{\partial L}{\partial \mathbf{a}}(\mathbf{a}, \mu) = \frac{\partial R(\mathbf{w})}{\partial \mathbf{w}} + \mu 2\mathbf{a} = \mathbf{A}\mathbf{a} + \mu 2\mathbf{a} = 0$$

$$\text{oder} \quad \mathbf{A}\mathbf{a} = \lambda \mathbf{a} \quad \text{mit} \quad \lambda := -2\mu \quad (\text{A.9})$$

Dies ist eine *Eigenwertgleichung* für \mathbf{A} ; die Lösungen \mathbf{a}^* sind die Eigenvektoren der Autokorrelationsmatrix \mathbf{A} . Die Zielfunktion (A.6) wird bei dem Eigenvektor \mathbf{a}^* mit dem größten Eigenwert maximal und beim Eigenvektor

mit dem kleinsten Eigenwert minimal. Damit ist das Konvergenzziel beim Gradientenabstieg durch den Eigenvektor mit dem kleinsten Eigenwert gegeben, was zu beweisen war.

Index

A

Abbaurrate 3-17, 3-45
activation function 2-19
Adjazenzmatrix 4-13, 4-14, 4-15, 4-16, 4-17, 4-18
affine Funktion 2-43
Aggregation 4-1
Aktivator 3-42, 3-43, 3-44, 3-45, 3-46
Aktivatorkonzentration 3-45
Aktivitätsfunktion 2-18, 2-43
Albedo 3-2
ANCOVA-Modell 4-8
Ankunftsprozesse 5-42
ANOVA-Modell 4-8, 4-9, 4-10
Anpassungstests 5-31
Anti-Hebb-Regel 2-38
Antithetische Zufallsvariable 5-48
ARMA-Prozess 5-43
ARTA-Prozess 5-43
Ausgabeaktivität 2-31
Ausgabefunktion 2-19, 2-24, 2-43
Ausgabegleichung 3-25, 3-29
Ausgleichsparabel n -ter Ordnung 2-18
Autokorrelationsmatrix 2-9
autoregressive Prozess 5-43
Axon 2-18

B

Backpropagation-Netz 2-20, 2-21, 2-28
Basisverbindung 4-19
Bayes-Klassifikation 2-30
Besetzungsdichte 5-16, 5-18, 5-19

Bestimmtheitsmaß 2-14
Binomialverteilung 5-44
bipartite Graphen 4-20
bistabile Schwinger 3-41
black box 1-2, 1-3
Blöcke 4-14, 4-16, 4-17, 4-18, 4-19, 4-20
Buffons Problem 5-35

C

charakteristisches Polynom 2-9
cluster 2-34, 2-35
Cobb-Douglas-Produktion 2-11
common random numbers 5-48
Competitive Learning 2-39
cross validation 2-28
cumulative distribution function 5-12

D

Delta-Regel 2-23
deterministische Simulation 5-1
deterministisches Verhalten 2-3
direkte Einflüsse 3-6
Dummy-Variable 2-11

E

Eigenvektor 2-8, 2-9, 2-38
Eigenwert 2-8, 2-9, 2-38
Endknoten 4-14
endogene Variable 2-3, 2-14
Erreichbarkeitsmatrix 4-18, 4-19, 4-21
error back-propagation 2-20, 2-21
Erwartungswert 2-4, 5-6, 5-13, 5-14, 5-20, 5-26, 5-27, 5-34, 5-36,

5-37, 5-38, 5-39, 5-43, 5-44, 5-45, 5-46, 5-47, 5-50
erwartungswerttreu 2-11, 2-12
essentielle
 Rückkopplungskantenmenge 4-20
essentielle
 Rückkopplungsknotenmenge 4-20
Euklid'scher Abstand 2-32
Euklid'scher Abstand 2-30
Euler-Cauchy-Integration 3-14
exogenen Einflüsse 2-14
exponentielle Verzögerung 3-30
exponentielles Wachstum 3-30, 3-35, 3-36

F
feedforward 2-20
Fisher-Snedecor-Verteilung 5-29
forecasting 2-1
Futterkonkurrenz 3-38, 3-39, 3-40

G
Gauß'sche Fehlerfunktion 5-27
Gauß'sche Proportion 2-5
Gauß'sches Eliminationsverfahren 2-17
Gaußsche Fehlerfunktion 2-15
Gauß-Verteilung 5-27
Generalisierung 2-3, 2-26, 2-31
glass box 1-3
Gleichgewichtslösung 3-49
Gleichungssystem 3-34
Glockenfunktion 2-28, 2-29, 2-30, 2-32, 2-34, 2-35, 2-36, 2-37, 2-38
GLODYM 4-1
Glossar 3-2, 3-4
Gradient 2-22, 2-23
grey box 1-3

H
Hesse'sche Normalform 2-7
hidden units 2-19
Histogramm 5-14, 5-15, 5-23, 5-25, 5-31

hit-and-miss Simulation 5-34
homogene Form 2-3
Hypothesentest 2-10, 2-15

I

In A-2
INFORGE 4-2, 4-3, 4-4, 4-5
INFORUM 4-1, 4-2
Inhibitor 3-42, 3-43, 3-44, 3-45, 3-46
input units 2-19
Input-Output-Analyse 4-5
Input-Output-Tabelle 2-2

K

k-mean-Clustering 2-35, 2-41, 2-42
knotendisjunkt 4-14
Konfidenzintervall 5-37, 5-50, 5-51
Kongruenz-Generator
 multiple-rekursiv 5-19
 quadratisch 5-19
konsistent 2-11, 2-12
Kovarianz 2-4
Kovarianzmatrix 2-8, 2-9, 2-16, 2-32
Kreis 4-14, 4-16, 4-20
Kreuzvalidierung 2-28

L

Lagrange'sche Funktion 2-8
Laplace-Verteilung 5-25
LCG Siehe linearer Kongruenter Generator,
Lernrate 2-23, 2-25
Lernregel 2-23, 2-38
likelihood-Wahrscheinlichkeit 3-53
linearer Kongruenter Generator 5-17
Lineares System 3-35
Linearisierung 2-10
logistisches Wachstum 3-32, 3-40

M

Mahalanobis Abstand 2-32
Mahalanobis-Abstand 2-33
maximum likelihood-Approximation 4-11

- maximum likelihood*-Schätzung 3-52
- m-Erlang-Verteilung 5-26
- m-fache Konvolution 5-26
- Minimum des erwarteten quadratischen Fehlers 2-6
- Missing values 5-6
- Mittelwert 2-4
- Modellbildung 2-1, 2-2
- Modellintegration 5-41
- Modellneuron 2-18
- Modellspezifikation 5-40
- Modulo-Bildung 5-18, 5-19
- Monte-Carlo-Simulation 5-32, 5-36
- Multikollinearität 2-13, 2-14
- multilayer 2-20
- Multilevel-Modelle 4-8
- multiple Regression 2-5, 2-16
- Muster 2-9, 2-18, 2-27, 2-28, 2-35, 2-38, 2-44
- N**
- Newton-Raphson 3-53, 3-54
- Newton-Raphson-Iteration 5-3
- Nullhypothese 2-14, 2-15
- Nyquist-Theorem 5-5
- O**
- OFF-Line-Lernen 2-25
- ON-Line-Lernen 2-25
- overfitting 2-26
- P**
- pdf* Siehe Wahrscheinlichkeitsdichte
- Pfad 3-7, 3-10, 4-14
- Phasendiagramm 3-37, 3-40, 3-50
- Poisson-Prozess 5-44, 5-45, 5-46
- Polynome 2-17
- probability density function 5-12, *Siehe* Wahrscheinlichkeitsdichte
- Pulsfortpflanzung 3-8, 3-9
- Q**
- Quelle 3-7
- R**
- Radiale Basisfunktion 2-28
- radialen Basisfunktionen 2-44
- RANDU-Generator 5-30, 5-32
- RAS 4-6
- Räuber-Beute 3-39, 3-40
- RBF *Siehe* Glockenfunktion
- RBF-Approximation 2-41, 2-42
- RBF-Funktion 2-41
- RBF-Netz 2-41
- RBF-Netzwerk 2-34
- RBF-Neuronen 2-31, 2-34, 2-37, 2-38
- RBF-Zentren 2-42
- reduzierter Graph 4-16
- Replication check 3-58
- Resampling 5-6, 5-7
- Ridge-Regression 2-12
- Rückkopplung 3-3, 3-10, 3-11, 3-12, 3-30, 3-37
- S**
- Samplingfrequenz 5-4
- Sättigungsmechanismus 3-44
- Schicht 2-19, 2-21, 2-23, 2-24, 2-25, 2-29, 2-30, 2-31, 2-34, 2-37, 2-38, 2-39, 2-41, 2-42, 2-43, 2-44
- Schleife 3-7
- Schwellwert 2-13, 2-18
- Senke 3-7
- Sensitivitätsanalyse 3-6
- Sigma-Funktionen 2-43, 2-44
- Sigma-unit 2-18
- signifikant 2-15
- Startknoten 4-14
- Stichprobenvarianz 2-16
- stopped training 2-27
- Störterm 2-13
- Störungsanalyse 3-8
- stratifizierte Kreuzvalidierung 2-28
- streng zusammenhängend 4-14, 4-16, 4-17
- streng zusammenhängende Komponenten 4-15
- Systemdiagramm 3-16, 3-17, 3-18, 3-19, 3-20, 3-22, 3-27, 3-29, 3-30

Systemmatrix 3-7, 3-8, 3-9, 3-10, 3-11, 3-13, 3-15, 3-20

T

Taylorentwicklung 2-17

Testmuster 2-26

TLMSE *Siehe* Total Least Mean Squared Error

Total Least Mean Squared Error 2-6

Trainingsmuster 2-22, 2-23, 2-25, 2-27

U

Übergangsknoten 3-7

unbiased 2-11, 2-12

V

Validierungsmenge 2-26, 2-27, 2-28

Variablensubstitution 4-22, 4-23

Varianz 2-4, 5-7, 5-14, 5-20, 5-27, 5-28, 5-31, 5-36, 5-37, 5-47, 5-48, 5-49, 5-50, 5-51

Verteilungsfunktion 5-12, 5-20, 5-21, 5-22, 5-23, 5-25, 5-26, 5-27, 5-38, 5-43

Volterra 3-39

W

Wahrscheinlichkeitsdichte 5-12, 5-13, 5-14, 5-27, 5-48

Welch Konfidenzintervall 5-51

Weltmodell 3-11, 3-12, 3-15, 3-20, 3-23

Wert-Diskretisierung 5-4

white box 1-2, 1-3

Wirkungsgraph 3-2, 3-3, 3-6, 3-7, 3-8, 3-12, 3-13, 3-28

Wirkungsmatrix 3-7

Z

Zeit-Diskretisierung 5-4

Zeitreihen 2-1, 2-2

zentraler Grenzwertsatz 5-27

Zielfunktion 2-5, 2-7, 2-22, 2-23

Zufallszahlen 5-12, 5-14, 5-15, 5-16, 5-17, 5-19, 5-22, 5-24, 5-25, 5-26, 5-27, 5-28, 5-29, 5-30, 5-31, 5-33, 5-37, 5-39, 5-48, 5-49, 5-50

Zustand 3-4, 3-9, 3-11, 3-12, 3-20, 3-24, 3-25, 3-26, 3-30, 3-32, 3-48

Zustandsdynamik 3-8, 3-12

Zustandsgleichung 3-25, 3-29

Zustandsnormierung 3-28

Zustandsvariablen 3-12, 3-23, 3-26, 3-28, 3-30, 3-32, 3-33, 3-34, 3-35, 3-46

Zyklen 3-9

Zyklus 3-3, 3-7, 3-10, 4-14

χ^2 -Test 5-31, 5-32

χ^2 -Verteilung 5-28, 5-32

Modelle zur biologischen Musterbildung: Turings Theorie und die spätere Entdeckung der Rolle von lokaler Selbstverstärkung und lang-reichweitiger Inhibition

Hans Meinhardt

Die Entwicklung eines höheren Organismus beginnt in der Regel mit einer befruchteten Eizelle und geschieht unter der Kontrolle der Gene. Ein Blick auf eineiige Zwillinge zeigt, bis in welche Details die Strukturbildung in den Genen kodiert ist. Der Verweis auf die Gene gibt aber noch keinen Aufschluss darüber, wie diese Strukturbildung erreicht wird. Die genetische Information wird bei jeder Zellteilung sorgfältig an beide Tochterzellen weitergegeben und ist somit, von wenigen Ausnahmen abgesehen, in allen Zellen die gleiche. Eine zentrale Frage in der Entwicklungsbiologie ist daher, wie sich ein höherer Organismus mit seiner komplexen aber reproduzierbaren Anordnung von Organen und differenzierten Zellen aus einer einzigen Zelle entwickeln kann.

Biologen haben in den letzten ca. 250 Jahren wichtige Informationen über die zugrundeliegenden Mechanismen zusammengetragen. Dazu wurde die normale Entwicklung eines Organismus gestört; z. B. hat man Teile weggeschnitten oder man hat bestimmte Teile an andere Stellen transplantiert. Das Vorgehen ist ähnlich wie in der Physik üblich: Man stört ein System und beobachtet das Verhalten nach der Störung. Aus diesen Beobachtungen kann man aber nicht direkt auf die zugrundeliegenden Mechanismen schließen. Zum Beispiel kann man aus den Spektrallinien, wie man sie nach der Anregung von Atomen beobachten kann, keinen direkten Rückschluss auf den Bau der Atome ziehen. Aber man kann, wie von Bohr begonnen, Modelle aufstellen und dann prüfen, wie weit die Modelle die Beobachtung wiedergeben. Dabei ist wichtig, dass die Modelle in einer mathematisch präzisen Weise angegeben werden. Das Bohr'sche Modell wurde trotz unplausibler Annahmen ernst genommen, weil

eine Vielzahl von Spektrallinien durch eine einzige Formel sehr genau wiedergegeben wurde.

Generell zeigte sich, dass die Embryonalentwicklung eines Organismus ein erstaunlich robuster Prozess ist. Viele Embryonen kann man in frühen Entwicklungsstadien in mehrere Teile zerlegen, und jedes dieser Fragmente ist in der Lage, einen kompletten Organismus zu bilden. Der frühe Seeigel-Embryo ist ein Beispiel. Bei der Maus kann jede Zelle, die einem Embryo im Acht-Zell-Stadium entnommen wurde, einen kompletten Organismus bilden. Diese Experimente zeigten, dass hier in frühen Stadien das weitere Schicksal der Zellen noch nicht endgültig festgelegt ist. Es muss eine Kommunikation zwischen den Zellen stattfinden, durch die Signale aufgebaut werden, die den Zellen ein positionsgemäßes Verhalten ermöglichen. Wenn eine bestimmte Struktur entfernt wurde, so werden offenbar die notwendigen Signale gebildet, durch die die entfernte Struktur ersetzt werden kann. Dabei hat sich gezeigt, dass kleine Nester von Zellen eine organisierende Funktion übernehmen. Die Organisator-Region im Amphibien-Keim, für deren Entdeckung Hans Spemann 1935 den Nobelpreis bekam, ist ein klassisches Beispiel. Eine Transplantation dieser kleinen Region auf die gegenüberliegende Seite des Keimes führt zur Bildung eines zweiten Embryos; beide Embryonen sind an der Bauchseite miteinander verwachsen. Die Idee ist,

DOI 10.1007/s00287-012-0625-4
© Springer-Verlag 2012

Hans Meinhardt
Max-Planck-Institut für Entwicklungsbiologie,
Spemannstr. 35, 72076 Tübingen
E-Mail: hans.meinhardt@tuebingen.mpg.de

Zusammenfassung

Die genetische Information ist in allen Zellen gleich. Für die Strukturbildung in einem sich entwickelnden Organismus müssen während der Entwicklung Signalsysteme aufgebaut werden, die ein positionsabhängiges Verhalten der Zellen ermöglichen. Alan Turing hat in seiner Arbeit von 1952 „The Chemical Basis of Morphogenesis“ [14] gezeigt, dass durch die Wechselwirkung von zwei unterschiedlich schnell diffundierenden Substanzen entsprechende Konzentrationsverteilungen aufgebaut werden können. Wir haben später gezeigt, dass eine solche Musterbildung dann und nur dann möglich ist, wenn eine lokal sich selbst verstärkende Reaktion mit einer lang-reichweitigen antagonistischen Reaktion gekoppelt ist. Heute kennen wir eine Reihe von solchen Signalsystemen, wie sie von der Theorie vorausgesagt wurden. Systeme mit drei beteiligten Komponenten wurden schon in Turings Arbeit erwähnt, aber wenig beachtet. Dieser Reaktionstyp, der zu Wanderwellen und zu Oszillationen in Gegenphase führen kann, spielt in verschiedenen Bereichen der Biologie auch eine wichtige Rolle.

dass in einer solchen Organisator-Region Signalmoleküle produziert werden. Durch Ausbreitung, zum Beispiel durch Diffusion, sowie durch den Zerfall der Moleküle entstehen gradierte Konzentrationsverteilungen. Die lokale Konzentration, der eine Zelle ausgesetzt ist, ist ein Maß für ihre Entfernung vom Signalzentrum, wodurch sich Zellen gemäß ihrer Position korrekt verhalten können. Damit verschiebt sich das Problem der räumlichen Musterbildung auf die Frage: Wie entsteht eine lokal begrenzte Organisator-Region an einer ganz bestimmten Stelle? Und wie kann eine solche Region neu gebildet werden, wenn die vorhandene Region entfernt wurde?

Mathematisch formulierte Modelle in der Biologie waren mehr die Ausnahme als die Regel. Von ihrer Geschichte her ist die Biologie mehr eine beobachtende und weniger eine deduktive Wissenschaft. 1952 publizierte Alan Turing eine wichtige Arbeit mit dem durchaus provokanten Titel „The Chemical Basis of Morphogenesis“ [14]. Er zeigte darin, dass durch die Wechselwirkungen von zwei Substanzen, die sich

durch Diffusion unterschiedlich schnell ausbreiten, eine instabile Situation entstehen kann. Kleine lokale Störungen oder sogar statistische Schwankungen sind ausreichend, um die Bildung eines Konzentrationsmusters zu initiieren. Es bilden sich hohe Konzentrationen in einigen Regionen und niedrige in anderen. Die von Turing angegebene Beispielgleichung für eine Reaktion mit dieser Eigenschaft lautet wie folgt:

$$\frac{\partial X}{\partial t} = 5X - 6Y + 1 \quad (1a)$$

$$\frac{\partial Y}{\partial t} = 6X - 7Y + 1 \quad [+ \text{ Diffusion}] \quad (1b)$$

Wie leicht zu sehen ist, hat eine solche Reaktion einen Gleichgewichtszustand bei $X = 1$ und $Y = 1$. Turing hat nun eine Kette von Zellen angenommen, die sich alle bis auf kleine Schwankungen in diesem Gleichgewichtszustand befinden. Mit einer kleinen Tabellen-Rechnung hat er gezeigt, dass die Konzentrationsunterschiede im Laufe der Zeit weiter ansteigen, was natürlich das Ziel seiner mathematischen Überlegungen war. Aus einer gegebenen Anfangsverteilung kann man mithilfe einer solchen Gleichung die späteren Konzentrationen berechnen. Das Anwachsen der Konzentrationsunterschiede ist in der Simulation Abb. 1 leicht zu erkennen.

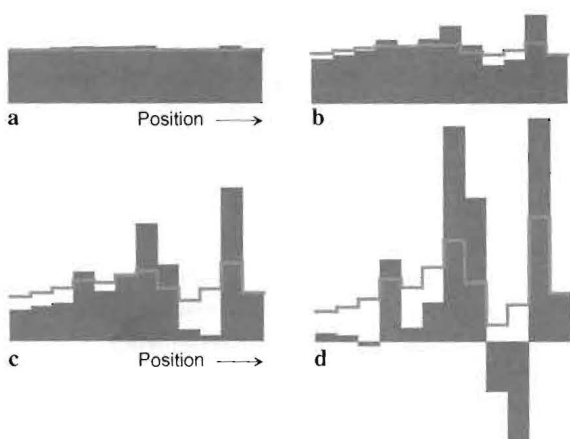


Abb. 1 Simulation einer Musterbildung mit der von Turing angegebenen (1a,b). Kleine Abweichungen vom Gleichgewichtszustand können die Musterbildung anstoßen. Die Konzentrationsunterschiede werden im Laufe der Zeit immer größer. Eine Eigenschaft von Turings Gleichung ist, dass „negative Konzentrationen“ entstehen können. a angenommener Ausgangszustand; b–d spätere Verteilungen; schwarze Balken: Substanz X; graue Linien: Substanz Y

Obw
Problem
Arbeit zu
ist verm
tiker sch
nicht der
klärung
dass Stru
stand he
von Wad
Entwickl
Überleg
wurde, b
Musterb.
initiiert
Struktur
gensatz
der beob
Ein
angegeb
sinnvoll
schwind
von X ab
küle abg
mehr vo
sehen ur
einfach z
setzen. C
angegeb
zentric
Konzent
steigt di
sie in de
Schließl
tratione
immer r

Mu Sel lan

Zusam
Arbeit
eine M
ist, we
tion m
ist, die
sere A
wir ni
es leic
tem (1

Obwohl sich damit eine Lösung für ein zentrales Problem in der Biologie abzeichnete, fand Turings Arbeit zunächst wenig Beachtung. Ein Grund dafür ist vermutlich, dass die Arbeit für Nichtmathematiker schwer zu verstehen ist. Auch wurde von ihm nicht der Versuch unternommen, eine intuitive Erklärung für das faszinierende Phänomen zu geben, dass Strukturen aus einem zuvor strukturlosen Zustand heraus entstehen können. Ein Argument, das von Waddington [15], einem sehr einflussreichen Entwicklungsbiologen mit Sinn für theoretische Überlegungen, gegen Turings Ansatz vorgebracht wurde, betrifft die Reproduzierbarkeit. Wenn die Musterbildung durch statistische Schwankungen initiiert wird, so sollten auch die sich bildenden Strukturen eine große Variabilität zeigen, im Gegensatz zu der unglaublichen Reproduzierbarkeit der beobachteten Entwicklung.

Ein weiteres Problem war, dass die von Turing angegebene Reaktionsgleichung molekular nicht sinnvoll ist. Die Rate, mit der die Moleküle X verschwinden ($-6Y$ in (1a)), hängt nur von Y , nicht von X ab. Das bedeutet, dass auch dann noch Moleküle abgebaut werden können, wenn keine Moleküle mehr vorhanden sind. Turing hat das Problem gesehen und vorgeschlagen, negative Konzentrationen einfach zu ignorieren, d. h., diese ad hoc auf null zu setzen. Ohne dieses Abschneiden führt die oben angegebene Reaktion nie zu einer stabilen Konzentrationsverteilung. In den Zellen, in denen die Konzentration einen bestimmten Wert übersteigt, steigt die Konzentration immer weiter an, während sie in den anderen Zellen immer weiter absinkt. Schließlich werden in einigen Zellen die Konzentrationen negativ und nehmen im weiteren Verlauf immer negativere Werte an.

Musterbildung erfordert lokale Selbstverstärkung und lang-reichweitige Inhibition

Zusammen mit Alfred Gierer habe ich 1972 eine Arbeit publiziert, in der wir gezeigt haben, dass eine Musterbildung dann und nur dann möglich ist, wenn eine sich lokal-selbstverstärkende Reaktion mit einer antagonistischen Reaktion gekoppelt ist, die eine lange Reichweite hat [4]. Als wir unsere Arbeit zur Publikation einreichten, wussten wir nichts von Turings Arbeit. Im Nachhinein ist es leicht zu zeigen, dass Turings Gleichungssystem (1a,b) unsere Bedingung erfüllt: X hat eine

positive Rückwirkung auf seine eigene Produktion. Außerdem katalysiert X die Produktion der schneller diffundierenden Komponente Y ; Y kontrolliert den Abbau von X , funktioniert also als Antagonist. Obwohl die beiden Turing'schen Gleichungen nahezu identisch aussehen, ist die Fähigkeit dieser Reaktion zur Musterbildung in der Selbstverstärkung von X und in der Inhibition des durch X produzierten Antagonisten Y begründet. In Turings Arbeit sind diese Bedingungen aber nicht erwähnt, auch nicht in versteckter Form. Vermutlich sind ihm aber Überlegungen in dieser Richtung vor seinem Tod durch den Kopf gegangen. In einer nachgelassenen Aufzeichnung wurde die folgende, etwas kryptische Notiz gefunden: "The amplitude of the waves is largely controlled by the concentration V of 'poison'" (zitiert nach [5], S. 494). Die Äquivalenz von Turings und unserem Ansatz wurde von Gierer [3] gezeigt. Eine Ideengeschichte der biologischen Musterbildung ist bei Roth [11] zu finden.

Eine denkbare molekulare Realisierung des von uns vorgeschlagenen Mechanismus besteht aus einem Aktivator a , der seine eigene Produktion katalysiert. Unter Kontrolle des Aktivators wird ein Inhibitor h produziert, der die Produktion des Aktivators hemmt:

$$\frac{\partial a}{\partial t} = \frac{\rho a^2}{h} - \mu_a a + D_a \frac{\partial^2 a}{\partial x^2} + \rho_0 \quad (2a)$$

$$\frac{\partial h}{\partial t} = \rho a^2 - \mu_h h + D_h \frac{\partial^2 h}{\partial x^2} \quad (2b)$$

Nach dieser Gleichung ist der Abbau des Aktivators ($-\mu_a a$) proportional zu a , d. h., die Abbaurate wird null, wenn die Konzentration null wird. Damit sind negative Konzentrationen unmöglich. Ein solcher linearer Abbau erfordert aber zwingend einen nichtlinearen Produktionsterm, sonst würde sich ein indifferentes Gleichgewicht ergeben. Diese Bedingung ist erfüllt, wenn zwei Aktivator-Moleküle zunächst ein Doppelmolekül, ein Dimer, bilden müssen. Für diese theoretische Voraussage gibt es heute viele Beispiele. Der Term ρ_0 in (2a) beschreibt eine Basis-Produktion des Aktivators. Diese ist z. B. notwendig um eine erneute Aktivierung während einer Regeneration zu beschreiben (Abb. 2).

Im Nachhinein mag es als verwunderlich erscheinen, warum es so lange gedauert hat, bis dieses Prinzip in der Entwicklungsbiologie gefunden wurde. Auch in der unbelebten Natur entstehen strukturierte Gebilde aus homogenen Anfangsver-

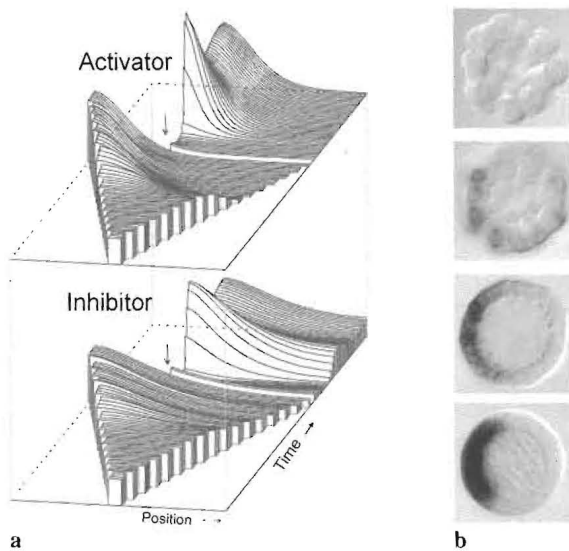


Abb. 2 Musterbildung durch die Wechselwirkung zwischen einem autokatalytischen Aktivator und einem lang-reichweitigen Inhibitor [4]. Links: Simulation mit (2a,b) in einer wachsenden Kette von Zellen. Wenn eine bestimmte Größe (Reichweite des Aktivators) überschritten ist, bildet sich auf einer Seite eine hohe, auf der anderen Seite eine niedrige Konzentration aus. Systeme dieser Art sind selbstregulierend. Entfernt man die aktivierte Region (Pfeil), so bildet sich eine neue. Rechts: Biologisches Beispiel: Die Bildung eines Signals im Seeigel durch die lokale Nodal-Aktivierung; es initiiert die Bildung der oralen Öffnung (Bild freundlicherweise von T. Lepage zu Verfügung gestellt; siehe [2]). Die Wechselwirkung von Nodal (Aktivator) und Lefty2 (Inhibitor) hat alle Eigenschaften, die von unserer Theorie vorhergesagt wurden [12]

teilungen, natürlich ohne dass dabei die Gesetze der Physik verletzt werden: Naiv würde man erwarten, dass der Sand in den Wüsten durch den Wind gleichverteilt wird. Aber das Gegenteil ist der Fall. Der Wind ist die treibende Kraft für die Bildung von Dünen. Aus einer diffusen Wolke kann ein feinstrukturierter Blitz schlagen oder gleichmäßig über das Land verteilter Regen kann im Laufe der Zeit durch Erosion scharf begrenzte Flüsse auswaschen. Diese Struktur erzeugenden Prozesse beruhen auf dem gleichen oben erwähnten Prinzip: Eine minimale lokale Abweichung von einer Gleichverteilung hat eine so starke Rückwirkung auf sich selbst, dass diese Abweichung weiter anwächst. Bei einer Sanddüne erzeugte vielleicht ein Stein einen Windschatten, hinter dem sich verstärkt Sand ablagerte. Der abgelagerte Sand vergrößerte den Windschatten, was die weitere Ablagerung beschleunigte; noch mehr Sand wurde abgelagert, usw.

Damit Strukturbildung möglich ist, muss neben der Selbstverstärkung auch noch die zweite von unseren Bedingungen erfüllt sein; Selbstverstärkung allein würde nur zu einer immer weiteren Ausbreitung und zu einem explosiven Anstieg der Reaktion führen. Dadurch würde nur ein Zustand in einen anderen, wieder strukturlosen Zustand übergehen. Strukturbildung impliziert aber, dass an einem Ort etwas geschieht, was in der Umgebung *nicht* geschieht. Eine Inhibition (Hemmung) muss von einem sich bildenden Maximum ausgehen, die dieser Selbstverstärkung effektiv entgegenwirkt. Diese Inhibition muss sich schneller ausbreiten als die autokatalytische Reaktion selbst. Dadurch ist die lokale Selbstverstärkung nur auf Kosten einer größeren Umgebung möglich und wird dadurch begrenzt. Diese Hemmung kann auch auf der Verarmung einer für die Reaktion notwendigen Komponente beruhen, die aus der Umgebung rekrutiert wird. Das ist im Dünenbeispiel der Fall. Wenn sich Sand abgelagert hat, trägt der Wind weniger Sand mit sich und die Wahrscheinlichkeit für eine weitere Dünenbildung in der Umgebung sinkt. Neben einem großen Fluss wird sich kein zweiter Fluss bilden. Auch in sozialen Auseinandersetzungen bilden sich so Strukturen und Ungleichheiten heraus. Bewerben sich mehrere Personen z. B. um ein Präsidentenamt, so wird jeder der Bewerber versuchen, seine eigene Position zu stärken und die der Rivalen zu schwächen. Am Ende wird einer diesen Kampf um die Spitze gewinnen. Im sich entwickelnden Organismus wird die Siegergruppe nicht Präsident, sondern ein organisierendes Zentrum.

Musterbildung ist nur möglich, wenn die verschiedenen Ausbreitungsgeschwindigkeiten der Komponenten ins Spiel kommen. Darum kann eine Musterbildung erst dann einsetzen, wenn ein wachsendes System eine gewisse Ausdehnung erreicht hat. In diesem Stadium kann sich eine hohe Konzentration nur am Rand ausbilden, da nur in dieser Anordnung die aktivierten und die nichtaktivierten Zellen den maximalen Abstand voneinander haben (Abb. 2). Dadurch entstehen gradierte Verteilungen der beteiligten Substanzen, die als Signalmoleküle verwendet werden können. Die Zellen sind positionsabhängig verschiedenen Konzentrationen ausgesetzt. Damit können in reproduzierbarer Weise in verschiedenen Regionen verschiedene Gene aktiviert werden. Dieses ist ein essenzieller Schritt in der Bildung der embryonalen Hauptachsen wie

Kopftische das e. gensa Wadc E weite aktivi Die A noch ist [4, Selbst Ir der g räumli Konze len. Ir muste Eines Abb. 2 Seeige Wirbe Links-

Tu ui Ein be der Tu Komp dass in Schrit Ein an phase hat di beigen kenne wäre. beit a angeg B pisch auf D ande ten h sind hoch zu v sche an d wird

KopfSchwanz oder RückenBauch. Auch wenn statistische Fluktuationen ein solches Muster initiieren, das entstehende Muster ist reproduzierbar, im Gegensatz zu der oben erwähnten Vermutung von Waddington.

Ein System, wie in (2) angegeben, ist in einem weiten Bereich selbstregulierend. Entfernt man eine aktivierte Region, so bildet sich eine neue (Abb. 2). Die Autokatalyse wird erneut gestartet, sobald der noch in dem Gewebe verbliebene Inhibitor abgebaut ist [4, 6], in Übereinstimmung mit der beobachteten Selbstregulation in frühen Entwicklungsstadien.

In anderen Parameterbereichen erlaubt der gleiche Mechanismus die Bildung von räumlich-periodischen und zeitlich-periodischen Konzentrationsverteilungen sowie von Wanderwellen. In der Zwischenzeit wurden mehrere solcher musterbildenden Reaktionen entdeckt (siehe [8]). Eines davon ist das Nodal/Lefty2 System [1, 12]. In Abb. 2 ist die Ausbildung der Nodal-Aktivierung im Seeigel-Keim gezeigt [2]. Das gleiche System ist in Wirbeltieren entscheidend an der Ausbildung der Links-rechts-Asymmetrie beteiligt.

Turings Dreikomponenten-Systeme und was dahinter steckt

Ein bemerkenswerter aber fast vergessener Teil der Turing'schen Arbeit beschäftigt sich mit Dreikomponenten-Systemen. Turing zeigte darin, dass in solchen Systemen spontan d. h., ohne Schrittmacher-Region Wellen auftreten können. Ein anderer Modus sind Oszillationen in Gegenphase in benachbarten Regionen. Turing selbst hat diesem Teil seiner Arbeit keine Bedeutung beigemessen. Er schrieb, dass er kein System kenne, wo ein solcher Mechanismus realisiert wäre. Mit dieser Begründung hat er in seiner Arbeit auch kein Gleichungssystem für diesen Modus angegeben.

Bei dem Versuch, die Pigment-Muster auf tropischen Schneckenschalen zu verstehen, sind wir auf Drei-Komponenten-Systeme gestoßen, die unter anderem die von Turing benannten Eigenschaften haben [10]. Die Muster auf Schneckenschalen sind ein wunderbares Übungsbuch der Natur, um hochdynamische Musterbildungssysteme besser zu verstehen (Abb. 3). Eine Schnecke oder Muschel kann ihr Gehäuse nur am Rand, d. h., nur an der Wachstumskante vergrößern. In der Regel wird auch nur hier Pigment in die Schale eingelagert.

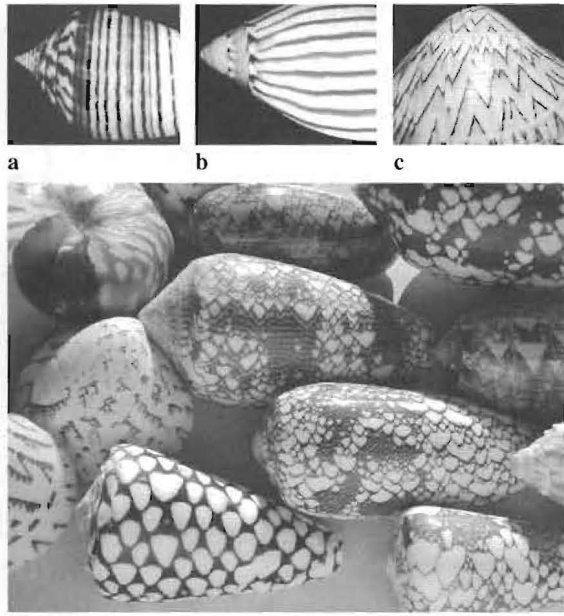


Abb. 3 Elementare und komplexe Muster auf tropischen Schneckenschalen. Pigmente werden nur an der Wachstumskante eingelagert. Deshalb stellen die zweidimensionalen Muster auf der Schale eine zeitliche Aufzeichnung eines dynamischen Prozesses dar, der entlang der linear ausgedehnten Wachstumskante stattgefunden hat. Diese kompletten Aufzeichnungen ermöglichen eine Dekodierung der zugrundeliegenden dynamischen Prozesse. Beispiele sind in den Abb. 4 und 5 zu finden

Die zweidimensionalen Muster entstehen also durch eine zeitliche Aufzeichnung von Prozessen, die an der (eindimensionalen) Wachstumskante stattgefunden haben. Viele dieser Prozesse haben chaotische Eigenschaften, was dazu führt, dass die komplexen Muster im Detail immer verschieden sind. Die Muster dienen weder zur gegenseitigen Erkennung noch zur Warnung von Fressfeinden, sondern entstehen vermutlich bei der Ausscheidung von Abbauprodukten. Durch die Einlagerung in die Schale werden diese Ausscheidungsprodukte sicher und dauerhaft entsorgt. Es ist für die Schnecke unerheblich, ob diese Ausscheidungen an bestimmten Stellen oder in gewissen zeitlichen Abständen erfolgen. Damit hatte die Natur die Möglichkeit zu spielen; es gab keinen selektiven Druck auf ein bestimmtes Muster. Elementare Muster sind Streifen senkrecht, parallel oder schräg zur Wachstumskante. Sie sind das Protokoll einer Pigmentablagerung, die entweder in stabilen räumlichen Abständen, in einer zeitlich-periodischen oder in einer sich wel-

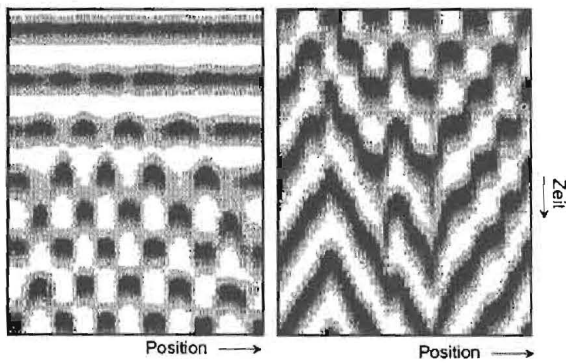


Abb. 4 Oszillationen in Gegenphase und Wanderwellen durch Drei-Komponenten-Systeme – ein schon von Turing antizipierter Mustertyp. Hier ist eine sich selbstverstärkende Reaktion angenommen, die durch zwei antagonistische Reaktionen geregelt wird. Durch einen lang-reichweitigen Antagonisten wird die räumliche Strukturbildung erreicht. Ein zweiter lokal-wirkender Antagonist mit langer Zeitkonstante bewirkt das Verschwinden der Aktivierung nach einiger Zeit. Eine synchrone Oszillation kann in eine Oszillation im Gegentakt übergehen. In der zeitlichen Aufzeichnung entsteht ein schachbrettartiges Muster. Durch eine kleine Parameter-Änderung geht das System dazu über, Wanderwellen auszubilden, was in der zeitlichen Aufzeichnung zu den schrägen Linien führt (links). Auf der Schale von *Bankivia fasciata* findet man den gleichen Übergang [9, 10]. Simulation mit (3a-c); das Bild wurde freundlicherweise von J. Campbell zur Verfügung gestellt

lenartig ausbreitenden Weise erfolgt war (Abb. 3, oben). Diese elementaren Muster sind mit Gleichungen vom Typ (2a,b) beschreibbar. Daneben gibt es aber Schalen, die mit dem Standard-Mechanismus nicht zu erklären sind (Abb. 3, unten). Viele dieser komplexen Muster kann man unter der Annahme simulieren, dass in jeder Zelle zwei oszillierende Systeme vorhanden sind. Ein System steuert die Pigmentablagerung. Ein zweites System, das in der Regel keine sichtbaren Spuren auf der Schale hinterlässt, verändert die Parameter des Pigment-systems. Die Natur des zweiten Systems musste also aus dem ungewöhnlichen Verhalten des pigmentbildenden Systems erschlossen werden [9]. In

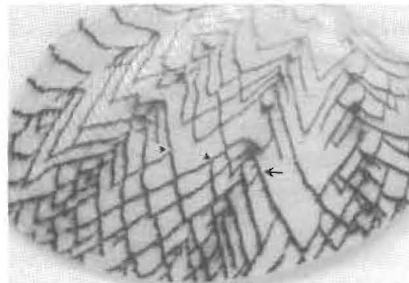
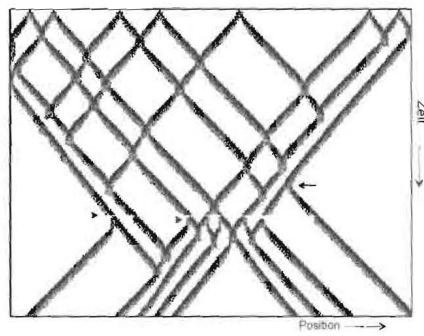


Abb. 5 Ungewöhnliche Eigenschaften von Wanderwellen, die durch Drei-Komponenten-Systeme generiert wurden. Wellen können sich solitonaartig durchdringen, was auf der Schale Kreuzungen von schrägen Linien hinterlässt. Normalerweise löschen sich Wellen in erregbaren Medien bei einer Kollision gegenseitig aus, was zu V-artigen Mustern auf der Schale führt. Ebenso ungewöhnlich ist, dass sich spontan rückwärts laufende Wellen bilden (sich verzweigende schräge Linien; Pfeil). Nach einer globalen Störung wurden gleichzeitig mehrere solche rückwärtslaufenden Wellen gebildet (Pfeilspitzen). Diese Muster-Elemente findet man auf der Schale von *Tapes literatus* wieder und sie werden ohne weitere Zusatz-Annahmen durch die in (3a-c) angegebenen Reaktionen beschrieben (aus [9, 10])

benachbarten Zellen sind die Oszillationen durch den Austausch von Signalmolekülen miteinander gekoppelt.

Etwas leichter zu durchschauen waren Muster, wie sie in den Abb. 4 und 5 gezeigt sind. Eine Art Schachbrett-Muster entsteht, wenn benachbarte Regionen entlang der Wachstumskante in Gegenphase oszillieren und diese Oszillationen aufgezeichnet werden (Abb. 4). Das ist ungewöhnlich, denn in der Literatur wird meist angenommen, dass sich gekoppelte nichtlineare Oszillatoren synchronisieren [13]. Diese und mehrere andere Muster können durch Drei-Komponenten-Systeme beschrieben werden. Dabei wird durch zwei der Komponenten, wie oben beschrieben, zunächst ein räumlich-periodisches Muster aufgebaut. Durch eine dritte Komponente, einem lokal wirkenden Antagonisten

mit langer Zeitkonstante, werden die entstandenen lokalen Maxima quasi „vergiftet“. Je nach Parametern kann das zwei Konsequenzen haben: Erstens, die Maxima verschwinden und erscheinen erneut an einer anderen, nichtvergifteten Stelle. Das kann zu Oszillationen in Gegenphase führen (Abb. 4). Zweitens, Maxima werden in eine Nachbar-Region verschoben, dort wieder vergiftet, usw. Es entstehen spontan, d. h., ohne Schrittmacher-Region, Wanderwellen. Diese beiden Mustertypen, Oszillationen in Gegenphase und Wanderwellen, sind aber genau die Muster, die Turing in dem entsprechenden Teil seiner Arbeit erwähnt hat. Es ist anzunehmen, dass Turings und unsere Drei-Komponenten-Systeme mathematisch äquivalent sind. Dieses wurde aber bisher noch nicht gezeigt. Da Wanderwellen und Oszillationen in Gegenphase auf dem gleichen Prinzip beruhen, können kleine Änderungen in den Parametern ausreichen, um einen Übergang von dem einen in den anderen Modus einzuleiten. Ein solcher Übergang ist in Abb. 4 sowohl in der Simulation als auch auf einem Schneckenhaus zu sehen.

Wanderwellen, die durch solche Wechselwirkungen entstehen, können ungewöhnliche Eigenschaften haben. Zum Beispiel können sich Wellen durchdringen. Normalerweise löschen sich kollidierende Wellen in erregbaren Medien aus. Das gegenseitige Auslöschen von aufeinander-treffenden Feuerfronten bei Waldbränden ist ein anschauliches Beispiel: Eine Feuerfront kann nicht in eine kürzlich abgebrannte Region weiterlaufen. Ebenso ungewöhnlich ist die spontane Bildung rückwärts-laufender Wellen. Auf Schalen findet man beide Prozesse (Abb. 5). Wie die Simulationen zeigen, sind diese durch Drei-Komponenten-Systeme beschreibbar.

Die Simulationen wurden mit dem folgenden Gleichungssystem durchgeführt:

$$\frac{\partial a}{\partial t} = \frac{sb(a^2 + b_a)}{(s_b + s_c)(1 + s_a a^2)} - r_a a + D_a \frac{\partial^2 a}{\partial x^2} \quad (3a)$$

$$\frac{\partial b}{\partial t} = b_b - \frac{sb(a^2 + b_a)}{(s_b + s_c)(1 + s_a a^2)} - r_b b + D_b \frac{\partial^2 b}{\partial x^2} \quad (3b)$$

$$\frac{\partial c}{\partial t} = r_c a - r_c c + D_c \frac{\partial^2 c}{\partial x^2} \quad (3c)$$

Dabei ist a die selbstverstärkende Komponente. Für die selbstverstärkende Reaktion wird die Substanz b benötigt und verbraucht; die Verarmung von b ist

also die erste antagonistische Reaktion. Ein weiterer Inhibitor c wird durch a produziert; er verlangsamt die Selbstverstärkung von a . Die indizierten Parameter $r_a, r_b, \dots, s_a, s_b, \dots$ beschreiben Zerfallsraten, Sättigungsterme usw. In den Simulationen in Abb. 4 und 5 ist angenommen, dass b der lokale und c der lang-reichweitige Antagonist ist.

In der Zwischenzeit hat sich herausgestellt, dass solche hochdynamischen Systeme, die nie zu einem stabilen Endzustand führen, häufig in der Biologie anzutreffen sind. Sie sind, zum Beispiel, involviert in der Bildung von Vogelfedern, in der Bestimmung der Zell-Mitte von E.-coli-Bakterien, damit dort die Zellteilung eingeleitet werden kann, oder in der Messung minimaler Konzentrationsunterschiede, was z. B. für weiße Blutkörperchen oder für auswachsende Nerven wichtig ist, damit diese ihre Zielregionen erreichen können [7–9].

Schlussbemerkung

In seiner Morphogenese-Arbeit von 1952 [14] hat Alan Turing gezeigt, dass eine Musterbildung durch die Wechselwirkung von zwei Komponenten möglich ist, wenn diese sich unterschiedlich schnell ausbreiten. Das war ein wichtiger Schritt. Von da aus war es noch ein langer Weg, die Musterbildung in der Biologie besser zu verstehen. Die Entdeckung, dass lokale Selbstverstärkung zusammen mit einer lang-reichweitigen Inhibition die eigentlich treibende Kraft ist, hat die Entwicklung realistischer Modelle sehr erleichtert. Diese Modelle sind Minimal-Modelle; sie zeigen auf, was mindestens notwendig ist, um einen bestimmten entwicklungsbiologischen Schritt fertigzustellen, ohne dass die Logik dieser Prozesse durch eine überbordende Komplexität unverständlich wird. Viele Fragen sind dabei noch immer offen. Zum Beispiel ist noch weitgehend unverstanden, wie das regional unterschiedliche Wachstum reguliert wird. Wie in anderen naturwissenschaftlichen Disziplinen erlaubt eine mathematisch formulierte Theoriebildung das Auffinden von elementaren Prinzipien, die nicht direkt aus der Beobachtung erschlossen werden können. Durch eine Theorie werden Diskrepanzen zwischen einem angenommenen Mechanismus und den beobachteten Phänomenen meist sehr viel deutlicher, was einen Wechsel in den Paradigmen erleichtert. Es ist zu erwarten, dass der Theoriebildung in der Entwicklungsbiologie eine wachsende Bedeutung zukommen wird,

um die immer komplexer erscheinenden molekulargenetischen Beobachtungen besser verstehen zu können.

Danksagung

Alfred Gierer danke ich für eine langjährige freundschaftliche Zusammenarbeit. Viele von den hier vorgestellten Ideen sind aus der gemeinsamen Arbeit hervorgegangen.

Literatur

1. Chen Y, Schier AF (2002) Lefty proteins are long-range inhibitors of squint-mediated Nodal signaling. *Curr Biol* 12:2124–2128
2. Duboc V, Rottinger E, Besnardeau L, Lepage T (2004) Nodal and BMP2/4 signaling organizes the oral-aboral axis of the sea urchin embryo. *Dev Cell* 6:397–410
3. Gierer A (1981) Some physical, mathematical and evolutionary aspects of biological pattern formation. *Proc R Soc Lond B* 295:429–440
4. Gierer A, Meinhardt H (1972) A theory of biological pattern formation. *Kybernetik* 12:30–39
5. Hodges A (1983) *Alan Turing: the enigma*. Simon and Schuster, New York
6. Meinhardt H (1982) *Models of biological pattern formation*. Academic Press, London (available for download at <http://www.eb.tuebingen.mpg.de/meinhardt/82-book>)
7. Meinhardt H (1999) Orientation of chemotactic cells and growth cones: Models and mechanisms. *J Cell Sci* 112:2867–2874
8. Meinhardt H (2008) *Models of biological pattern formation: from elementary steps to the organization of embryonic axes*. *Curr Top Dev Biol* 81:1–63
9. Meinhardt H (2009) *The Algorithmic Beauty of Sea Shells*, 4. Aufl (mit PC-Programmen und animierten Simulationen auf einer CD). Springer, Heidelberg, New York
10. Meinhardt H, Klingler M (1987) A model for pattern formation on the shells of molluscs. *J Theor Biol* 126:63–89
11. Roth S (2011) Mathematics and biology: a Kantian view on the history of pattern formation theory. *Dev Genes Evol* 221:255–279
12. Schier AF (2009) Nodal Morphogens. *Cold Spring Harb Perspect Biol*, doi:10.1101/cshperspect.a003459
13. Stewart I (1991) All together now. *Nature* 350:557–557
14. Turing A (1952) *The Chemical Basis of Morphogenesis*. *Phil Trans B* 237:37–72
15. Waddington CH (1956) *Principles of Embryology*. Georg Allen and Unwin Ltd, London

1
Andr
dass
erhie
them
Ausz
Studi
Grun
Acke
dung
epoc
an Ap
verev
siche
ferun

Im S
begi
land
er z
Jahr
nem
Rac
tori
Zen
nem
mi
sity
nac
ter
de
Ve
sch