
Verfahren der stochastischen Signalverarbeitung

1–semestrige Vorlesung für Studenten der Physik

BERND POMPE

Institut für Physik der Universität Greifswald

Studentenfassung vom 29. Januar 2007

Inhaltsverzeichnis

1	Lineare Systeme, Fourieranalyse und Abtasttheorem	7
1.1	Einleitung	7
1.2	Signale und Systeme	8
1.3	Fourier–Transformation	9
1.4	Abtasttheorem	15
1.5	Diskrete Fourier–Transformation	21
2	Interpolation (engl.)	29
2.1	Introduction	29
2.2	Interpolation	29
2.3	Differentiation	39
3	Korrelation	45
3.1	Definition und graphische Interpretation	45
3.2	Eigenschaften	48
3.3	Korrelationsfunktion	50
3.4	Korrelationsfunktion für zeitkontinuierliche Signale	52
3.4.1	Kreuzkorrelationsfunktion für periodische Signale	53
3.4.2	Autokorrelationsfunktion für periodische Signale	54
4	Stochastischer Prozess	57
4.1	Grundlegenden Begriffe	57
4.2	Zusammenhang von Korrelationsfunktion und Spektraldichte	59
4.3	Beispiele	61
4.3.1	Vollkommen unabhängiger Prozess	61
4.3.2	Korrelation normalverteilter Zufallsgrößen	62
4.3.3	Vollkommen abhängiger δ –korrelierter Prozess	64
4.3.4	Gleitmittel–Prozess	64
4.3.5	Autoregressiver Prozess	66
5	Autoregressives Prädiktor–Modell	71
5.1	Lineares Modell	71
5.2	Nichtlineares Modell	72
5.3	The General Prediction Problem	73
5.4	Best Least Mean Square Predictor	74
5.5	Nearest Neighbor Predictor	75

6	Nichtparametrische Schätzung von Wahrscheinlichkeitsdichten	77
6.1	Motivation	77
6.2	Histogramm	78
6.3	Naiver Schätzer	78
6.4	Kern-Schätzer	78
7	Schätzung von Zufallsgrößen	81
7.1	Problemstellung und grundlegende Begriffe	81
7.1.1	Mittelwertschätzung	81
7.2	Definitionen	84
7.2.1	Erwartungstreuer Schätzer (unbiased estimator)	84
7.2.2	Konsistenter Schätzer (consistent estimator)	86
7.2.3	Effizienter Schätzer (efficient estimator)	87
7.3	Zentraler Grenzwertsatz	87
7.4	Grundlegende Schätzer	88
7.5	Maximum Likelihood Schätzung (MLS)	89
7.5.1	Beispiel: MLS des Mittelwertes bei ungleicher Messgenauigkeit	90
7.6	Maximum Entropie Schätzung (MES)	91
7.6.1	Beispiel: MES bei gegebenen Erwartungswerten — ein-dimensionaler Fall	91
7.6.2	MES bei gegebenen Erwartungswerten — mehr-dimensionaler Fall	93
7.7	Spektralschätzung	94
8	Statistisches Schließen	97
8.1	Grundlegende Begriffe	97
8.2	Wichtige Verteilungen	99
8.2.1	χ^2 -Verteilung (HELMERT-PEARSON-Verteilung)	99
8.2.2	Student-Verteilung (t -Verteilung)	100
8.3	Signifikanztests	101
8.3.1	χ^2 -Streuungstest	103
8.3.2	χ^2 -Unabhängigkeitstest	104
8.4	χ^2 -Anpassungstest	106
8.5	χ^2 -Homogenitätstest	106
9	Singulärwertzerlegung (SWZ)	107
9.1	Problemstellung	107
9.2	Normierung	107
9.3	SWZ	108
10	Vektorquantisierung	113
10.1	Problemstellung	113
10.2	Prinzip	113
10.3	Optimales Codebuch bei gegebener Wahrscheinlichkeitsdichte	114
10.4	Optimales Codebuch aus Monte-Carlo-Methode: LBG-Algorithmus	115
10.5	Schnelle Suche	117

11 Künstliche Neuronale Netze (KNN)	121
11.1 Neuronen als Elementarbausteine natürlicher Neuronaler Netze	121
11.2 Die Modellierung natürlicher Neuronaler Netze	122
12 Bayes Statistik	125
A Fourier–Entwicklungen	127
A.1 Reihenentwicklung einer symmetrischen Rechteckimpulsfolge	127
A.2 Reihenentwicklung eines symmetrischen Dreiecksignals	128
A.3 Spektraldichte eines Rechteckfensters	129
A.4 Zeitfunktion eines spektralen Rechteckfensters	129
B Programm: Fast–Fourier–Transformation	131
C Programm: Interpolation	133
D Grundlagen der Fehlerrechnung	137
D.1 Problemstellung	137
D.2 Einteilung der Fehler	138
D.3 Abschätzung und Fortpflanzung des Größtfehlers	139
D.3.1 Allgemeine Betrachtung	139
D.3.2 Beispiel	140
D.4 Statistische Analyse und Fortpflanzung des zufälligen Fehlers	141
D.4.1 Fehlerstatistik	141
D.4.2 Fortpflanzung zufälliger Fehler	145
D.5 Beispiel	147
D.6 Gewichtete Mittelung	150
E Regression	151
E.1 Lineare Regression	151
E.1.1 1. Fall: x fehlerfrei, y mit festem Fehler σ_y	151
E.1.2 2. Fall: x_n und y_n mit Fehler σ_x bzw. σ_y	153
E.1.3 3. Fall: x_n fehlerfrei, y_n mit variablem Fehler $\sigma_{y;n}$	153
E.1.4 4. Fall: x_n und y_n mit variablem Fehler $\sigma_{x;n}$ bzw. $\sigma_{y;n}$	154

Kapitel 1

Lineare Systeme, Fourieranalyse und Abtasttheorem

1.1 Einleitung

In diesem Kapitel wird auf die Bedeutung des Abtasttheorems bei der Signalerfassung und –rekonstruktion eingegangen sowie das Verständnis für die Übertragungsfunktion linearer zeitinvarianter Systeme vertieft.

Nichtelektrische Messgrößen wie z.B. die Temperatur, die Schall- und Lichtintensität können durch einen entsprechenden Sensor (Thermoelement, Mikrofon bzw. Phototransistor) in eine elektrische Messgröße (meist die elektrische Spannung) umgewandelt werden. Ist die Messgröße zeitlich veränderlich, so ist dies auch die Spannung am Sensor, welche das *Messsignal* darstellt. Eine messtechnische Grundaufgabe besteht darin, dieses Messsignal zu erfassen, zu bearbeiten, um es dann z.B. graphisch anzuzeigen bzw. zu analysieren oder als Steuergröße in einem Regelkreis zu verwenden. Die Verarbeitung des Messsignals kann dabei auf zwei wesentlich verschiedenen Wegen erfolgen: analog oder digital.

Im analogen Fall wird das Messsignal z.B. verstärkt, analog gefiltert und auf einem Oszilloskop angezeigt. Mathematische Operationen mit dem Signal wie z.B. die Verschiebung oder Quadrierung werden durch elektronische Bauelemente, meist in Verbindung mit Operationsverstärkern, realisiert. Vorteile einer analogen Messwertverarbeitung gegenüber einer digitalen bestehen darin, dass das Signal praktisch ohne Zeitverzögerung mit einer vergleichsweise simplen elektronischen Schaltung verarbeitet werden kann. Allerdings ist die analoge Signalverarbeitung empfindlicher gegenüber verschiedenen Störungen etwa in Form des Temperaturrauschens von Halbleitern oder elektromagnetischen Einstreuungen. Darüber hinaus ist die analoge Ausführung der gewünschten Operationen meist ungenauer als auf digitalem Wege. Die Reproduzierbarkeit der Ergebnisse ist demzufolge im analogen Fall geringer.

Werden die Messsignale in der Zeit und Amplitude diskretisiert, so können sie auch mit einem Digitalrechner verarbeitet werden. Die digitale Signalverarbeitung setzt sich gegenüber der analogen weithin durch, sind doch digitale Systeme bedeutend störsicherer. Mathematische Operationen können mit vergleichsweise großer Genauigkeit ausgeführt werden. Die Verarbeitungsergebnisse sind praktisch exakt und mit einer äußerst geringen Fehlerwahrscheinlichkeit reproduzierbar. Durch die Verwendung frei programmierbarer Digitalrechner ist eine bedeutend größere Flexibilität der Apparatur gegeben als im analogen Fall. Der meist höhere schaltungstechnische Aufwand digitaler Systeme sowie ihre zunächst geringere Verarbeitungsgeschwindigkeit verlieren mit den laufenden technologischen Fortschritten — größere Integrationsdichten,

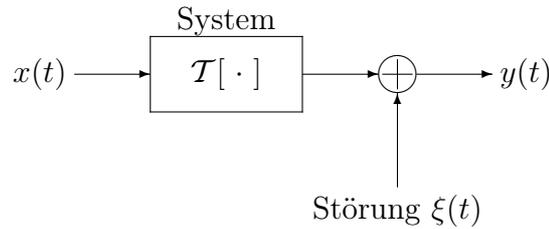


Bild 1.1: Allgemeines Schema der Signalübertragung

digitale Signalprozessoren u.a. — zunehmend ihre Bedeutung.

Eine digitale Signalverarbeitung setzt wegen der endlichen Speicherkapazität von digitalen Anlagen sowie wegen der Forderung nach einer möglichst raschen Signalverarbeitung eine Diskretisierung in der Zeit sowie eine Quantisierung der Signalamplituden voraus. Diese Diskretisierungen führen zu einer Signalverfälschung (Jitter, Shimmer, Quantisierungsrauschen) und somit letztlich zu Ungenauigkeiten in den Ergebnissen der Signalverarbeitung, der Signaldarstellung bzw. –rekonstruktion. Werden jedoch gewisse Regeln bei dieser Diskretisierung eingehalten, so kann man davon ausgehen, dass die durch Diskretisierung entstehenden Fehler praktisch vernachlässigbar sind.

1.2 Signale und Systeme

Wird ein Signal $x(t)$ über einen Nachrichtenkanal übertragen oder in einem allgemeineren Sinne „verarbeitet“, so entsteht als Ergebnis dieser Verarbeitung ein Ausgangssignal $y(t)$, das vom Eingangssignal $x(t)$ abhängt. Der Operator, welcher $x(t)$ nach $y(t)$ überführt, sei mit \mathcal{T} bezeichnet. Bei einer Signalübertragung bzw. –verarbeitung treten jedoch immer zufällige Fehler auf. Vermeintlich gleiche Eingangssignale liefern im Allgemeinen unterschiedliche Ausgangssignale. Der Zusammenhang zwischen Ein- und Ausgangssignal ist also nicht „rein“ deterministisch. Aus systemtheoretischer Sicht besteht deshalb im Allgemeinen der Zusammenhang

$$y(t) = \mathcal{T}[x(t)] + \xi(t) . \quad (1.1)$$

Hierin beschreibt \mathcal{T} den deterministischen Zusammenhang zwischen Ein- und Ausgangssignal, zufällige Übertragungs- oder Verarbeitungsfehler werden durch die Zufallsgröße $\xi(t)$ berücksichtigt. In der Abbildung 1.1 ist dies illustriert.

Eine wichtige Aufgabe besteht darin, zu einer gegebenen technischen Apparatur den *Operator* \mathcal{T} zu bestimmen, was einer Modellbildung entspricht, und \mathcal{T} auf relevante Signale $x(t)$ anzuwenden. (Eine Charakterisierung der Störung $\xi(t)$ ist ebenso von Interesse. Dies wird hier aber nicht weiter verfolgt.) Die Bestimmung von \mathcal{T} kann auf vielfältige Weise erfolgen. So können z.B. zu verschiedenen Eingangssignalen $\{x_i(t)\}$ die entsprechenden Ausgangssignale $\{y_i(t)\}$ beobachtet (gemessen) oder theoretisch bestimmt werden, um dann auf den gesuchten Operator \mathcal{T} zu schließen, meist unter Vernachlässigung zufälliger Störungen ($\xi(t) \approx 0$). Die genannte Aufgabe ist jedoch im Allgemeinen nicht lösbar, da es keine universelle Theorie hierzu gibt. Beschränkt man sich aber auf eine gewisse Klasse von Systemen, die sog. *linearen zeitinvarianten Systeme*, so steht eine ausgebaute Signaltheorie zur Verfügung.

Der Operator \mathcal{T} eines *linearen* Systems genügt für zwei beliebige Eingangssignale $x_1(t)$ sowie $x_2(t)$ sowie willkürlich gewählte reelle Zahlen a_1, a_2 der *Überlagerungsbedingung*

$$\mathcal{T}[a_1x_1(t) + a_2x_2(t)] = a_1\mathcal{T}[x_1(t)] + a_2\mathcal{T}[x_2(t)] . \quad (1.2)$$

Diese Beziehung besagt, dass die Antwort des linearen Systems auf eine Summe von Signalen gleich der Summe von Systemantworten ist, die man erhalten hätte, wenn die Signale einzeln übertragen worden wären. Darüber hinaus ist es gleich, ob das Eingangssignal zunächst mit einem beliebigen Faktor gestreckt und dann übertragen oder in umgekehrter Reihenfolge verfahren wird.

Bezeichne $\mathcal{S}^\tau: x(t) \rightarrow x(t + \tau)$ den (linearen) Verschiebungsoperator. Ein System heißt *zeitinvariant* (allgemeiner: *verschiebungsinvariant*), wenn zu einer beliebigen Verschiebung τ die Operatoren \mathcal{T} und \mathcal{S}^τ kommutativ sind,

$$\mathcal{T} \circ \mathcal{S}^\tau[x(t)] = \mathcal{S}^\tau \circ \mathcal{T}[x(t)]. \quad (1.3)$$

Das bedeutet, die Reihenfolge der Operationen „Zeitverschiebung“ und \mathcal{T} sind vertauschbar bzw. \mathcal{T} hängt hier nicht von der Zeit ab (wohl aber das transformierte Signal y).

Eine allgemeine Form des Operators \mathcal{T} eines linearen zeitinvarianten Systems ist das *Faltungsintegral*

$$\mathcal{T}[x(t)] = \int_{-\infty}^{+\infty} x(\tau)h(t - \tau)d\tau \equiv x(t) * h(t) . \quad (1.4)$$

Die Funktion h charakterisiert das System und wird *Gewichtsfunktion* genannt.

Lineare zeitinvariante Systeme sind häufig recht gute Modelle realer Systeme. So sind z.B. alle elektrischen Übertragungsnetzwerke, die nur aus passiven Bauelementen (Ohmschen Widerständen R , Kapazitäten C , Induktivitäten L) bestehen, linear und zeitinvariant. Ebenso sind aktive Filter, sofern keine Übersteuerungen infolge zu großer Signalpegel auftreten, in guter Näherung als linear und zeitinvariant anzusehen. Andererseits stellt z.B. eine Gleichrichterschaltung oder ein übersteuerter Verstärker ein nichtlineares System dar. Streng genommen sind alle realen Systeme nichtlinear und zeitvariant, betreibt man jedoch z.B. eine elektronische Schaltung mit Eingangssignalen, die in einem bestimmten eingegrenzten Bereich liegen, so kann man häufig in guter Näherung das System als linear modellieren. Werden darüber hinaus z.B. alterungsbedingte Änderungen des Übertragungsverhaltens des realen Systems vernachlässigt, so kann man das System als zeitinvariant betrachten.

1.3 Fourier–Transformation

Häufig kennt man den Operator \mathcal{T} eines linearen zeitinvarianten Systems nur in seiner Wirkung auf eine spezielle Familie $\{x_i(t)\}$ von Signalen. Setzt man z.B. in einem passiven RCL–Netzwerk die komplexen Widerstände $R_C = 1/(j\omega C)$ und $R_L = j\omega L$ für Kapazitäten C bzw. Induktivitäten L , so erhält man unter Anwendung der Kirchhoffschen Regeln für feste Werte von R, C und L eine komplexwertige *Übertragungsfunktion* $H(j\omega)$. Ein Beispiel ist in der Abbildung 1.2 dargestellt. Hierbei handelt es sich um einen Bandpass mit der Übertragungsfunktion

$$H(j\omega) \equiv |H(j\omega)|e^{j\varphi_H(\omega)} = \frac{j\omega RC}{1 + j\omega RC - \omega^2 LC} , \quad (1.5)$$

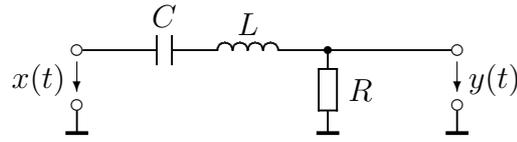


Bild 1.2: Lineares System: LRC-Bandpass

mit dem Betrag

$$|H(j\omega)| = \frac{(\omega RC)^2}{(1 - \omega^2 LC)^2 + (\omega RC)^2}$$

und der Phase

$$\varphi_H(\omega) = \arctan \frac{1 - \omega^2 LC}{\omega RC} .$$

Die Übertragungsfunktion (1.5) beschreibt, wie *harmonische* Schwingungen

$$x(t) = |x| \cdot \cos(\omega t + \varphi_x) \tag{1.6}$$

in Abhängigkeit von der Kreisfrequenz ω durch den Bandpass modifiziert werden:

$$y(t) = |H(j\omega)| \cdot |x| \cdot \cos[\omega t + \varphi_x + \varphi_H(\omega)] .$$

Somit hat man einen expliziten Ausdruck für den Operator \mathcal{T} : $x(t) \rightarrow y(t)$ für harmonische Eingangssignale.

Fourier-Zerlegung

Ist das Eingangssignal $x(t)$ nicht harmonisch, so kann aus der Übertragungsfunktion (1.5) nicht unmittelbar auf $\mathcal{T}[x(t)]$ geschlossen werden. Approximiert man jedoch $x(t)$ durch additive Überlagerung harmonischer Funktionen verschiedener Frequenzen, so kann unter Beachtung der Linearität (1.2) zunächst \mathcal{T} auf die harmonischen Signale angewandt werden und anschließend das gesuchte Ausgangssignal $y(t) = \mathcal{T}[x(t)]$ durch Überlagerung der übertragenen harmonischen Komponenten von $x(t)$ erhalten werden. Die Approximation eines Signals mit harmonischen Signalen heißt *Fourier-Zerlegung*.

Für *zeit-kontinuierliche periodische* Signale $x(t)$ lautet die Fourier-Zerlegung

$$x(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cdot \cos n\omega_0 t + b_n \cdot \sin n\omega_0 t. \tag{1.7}$$

Hierbei ist $T_0 \equiv 2\pi/\omega_0$ die (triviale) Periode von $x(t)$ und die *Fourier-Koeffizienten*

$$a_n = \frac{2}{T_0} \int_{-T_0/2}^{T_0/2} x(t) \cdot \cos n\omega_0 t \, dt , \quad b_n = \frac{2}{T_0} \int_{-T_0/2}^{T_0/2} x(t) \cdot \sin n\omega_0 t \, dt . \tag{1.8}$$

Die Fourier-Zerlegung ist nicht immer möglich. So ist z.B. die periodische Funktion $x(t) = \sin \frac{1}{\sin \omega t}$ nicht zerlegbar. Hinreichende (im Allgemeinen jedoch nicht notwendige) Bedingungen für die Existenz der Fourier-Zerlegung eines Signals $x(t)$ sind:

1. $x(t)$ ist absolut integrierbar, d.h. $\int_0^{T_0} |x(t)| dt < \infty$.
2. In einer Periode $[t, t + T_0]$ treten nur endlich viele Unstetigkeiten und relative Minima und Maxima auf.

Für reale periodische Signale können diese Bedingungen zumeist als erfüllt angesehen werden. Die Fourier-Reihe (1.7) kann unter Verwendung der Abkürzungen

$$\begin{aligned} A_0 &= \frac{a_0}{2}, \\ A_n &= \sqrt{a_n^2 + b_n^2} \quad \text{und} \\ \varphi_n &= \arctan\left(-\frac{b_n}{a_n}\right) \end{aligned}$$

folgendermaßen dargestellt werden:

$$x(t) = \sum_{n=0}^{\infty} A_n \cdot \cos(n\omega_0 t + \varphi_n). \quad (1.9)$$

Somit kann das Ausgangssignal $y(t) = \mathcal{T}[x(t)]$ eines linearen Systems, das durch die Übertragungsfunktion $H(j\omega)$ beschrieben ist, berechnet werden:

$$y(t) = \sum_{n=0}^{\infty} |H(jn\omega_0)| \cdot A_n \cdot \cos[n\omega_0 t + \varphi_n + \varphi_H(n\omega_0)]. \quad (1.10)$$

Eine elegantere Darstellung gelingt durch Einführung der komplexen Fourier-Koeffizienten

$$X(\pm jn\omega_0) = \frac{a_n \mp jb_n}{2}. \quad (1.11)$$

Somit kann das T_0 -periodische Signal in (1.7) bzw. (1.9) auch wie folgt geschrieben werden:

$$x(t) = \sum_{n=-\infty}^{+\infty} X(jn\omega_0) \cdot e^{jn\omega_0 t} \quad (1.12)$$

Die Fourier-Komponenten $X(jn\omega_0)$ können aus den Beziehungen (1.8) und (1.11) erhalten werden oder direkt aus

$$X(jn\omega_0) = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} x(t) \cdot e^{-jn\omega_0 t} dt \quad (1.13)$$

Für das Ausgangssignal $y(t)$ erhält man somit

$$y(t) = \sum_{n=-\infty}^{\infty} H(jn\omega_0) \cdot X(jn\omega_0) \cdot e^{jn\omega_0 t}. \quad (1.14)$$

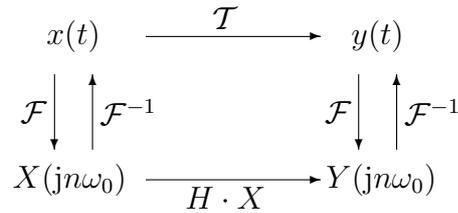


Bild 1.3: Kommutatives Diagramm für ein lineares System

Hieraus können sofort die Fourier-Komponenten $Y(jn\omega_0)$ des Ausgangssignals abgelesen werden:

$$Y(jn\omega_0) = H(jn\omega_0) \cdot X(jn\omega_0). \quad (1.15)$$

Die durch (1.13) definierte Abbildung

$$\mathcal{F} : x(t) \rightarrow X(jn\omega_0)$$

heißt *Fourier-Transformation* und die Menge aller Fourier-Komponenten $X(jn\omega_0)$ *Fourier-Transformierte* oder kurz *Spektrum* des Signals $x(t)$. Die *inverse Fourier-Transformation*

$$\mathcal{F}^{-1} : X(jn\omega_0) \rightarrow x(t)$$

ist durch (1.12) definiert.

Für Signale, die in der Praxis vorkommen, sind $x(t)$ und $\mathcal{F}^{-1} \circ \mathcal{F}[x(t)]$ für alle Zeitpunkte t identisch.¹⁾ Folglich ist das Diagramm in Abbildung 1.3 kommutativ. Hiernach kann das Ausgangssignal y des durch die Übertragungsfunktion H beschriebenen linearen Systems erhalten werden, indem H mit der Fourier-Transformierten des Eingangssignals x multipliziert und das Produkt dann rücktransformiert wird:

$$y(t) = \mathcal{F}^{-1} \left[H \cdot \mathcal{F}[x(t)] \right]. \quad (1.16)$$

Spektraldichte

Nichtperiodische Signale $x(t)$ können nicht durch eine unendliche Summe (1.12) dargestellt werden. An die Stelle des Spektrums $X(jn\omega_0)$ tritt hier die *Spektraldichte*

$$\boxed{
 \mathcal{F}[x(t)](\omega) \equiv X(j\omega) = \int_{-\infty}^{+\infty} x(t) \cdot e^{-j\omega t} dt \quad (1.17)
 }$$

Man erhält sie aus dem Grenzübergang

$$X(j\omega) = \lim_{T_0 \rightarrow \infty} T_0 \cdot X(jn2\pi/T_0). \quad (1.18)$$

¹⁾Aus mathematischer Sicht kann jedoch für einige Zeitpunkte das originale Signal vom hin- und rücktransformierten abweichen. Allerdings stimmen beide zumindest fast überall bez. des Lebesgue-Maßes überein. Das heißt, die Menge der Zeitpunkte, zu denen keine Übereinstimmung auftritt, hat auf der Zeitachse die Länge Null.

Nichtperiodische Signale werden also als „periodische Signale mit unendlich großer Periode T_0 “ aufgefaßt. Gleichung (1.17) definiert die Fourier-Transformation $\mathcal{F} : x(t) \rightarrow X(j\omega)$ für nichtperiodische Signale. Die zugehörige inverse Fourier-Transformation $\mathcal{F}^{-1} : X(j\omega) \rightarrow x(t)$ ist wie folgt definiert:

$$\mathcal{F}^{-1}[X(j\omega)](t) \equiv x(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} X(j\omega) \cdot e^{j\omega t} d\omega \quad (1.19)$$

Sie tritt an die Stelle der Reihendarstellung (1.12) periodischer Signale. Für die Existenz der Spektraldichte (1.18) eines Signals $x(t)$ sind die folgenden Bedingungen hinreichend:

1. Der Definitionsbereich von x kann in endlich viele Teilintervalle zerlegt werden, in denen $x(t)$ stetig und monoton ist
2. An jeder Unstetigkeitsstelle sind die Grenzwerte $x(t-0)$ und $x(t+0)$ definiert.
3. $x(t)$ ist absolut integrierbar, d.h., $\int_{-\infty}^{+\infty} |x(t)| dt$ existiert (und ist endlich).

Die ersten beiden Bedingungen sind nach DIRICHLET benannt. Die dritte Bedingung ist gleichbedeutend damit, dass $\int_{-\infty}^{+\infty} |x(t)|^2 dt$ existiert. An den Unstetigkeitsstellen liefert die Rücktransformierte $\mathcal{F}^{-1}[X(j\omega)](t) = [x(t-0) + x(t+0)]/2$.

In der Praxis wird die Existenz der Fourier-Transformierten in der Regel nicht in Frage gestellt, weil hier Signale zeitlich und in der Amplitude begrenzt anzusehen sind. Aber schon eine Sinusfunktion $x(t) = A \cdot \cos \omega_0 t$, mit $A, \omega \neq 0$, erfüllt nicht die Bedingung nach absoluter Integrierbarkeit. Wird die Transformation jedoch im Raum verallgemeinerter Funktionen (Distributionen) betrachtet, dann gilt $X(j\omega) = A\pi \cdot \delta(|\omega| - \omega_0)$, denn

$$\mathcal{F}^{-1}[X(j\omega)](t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} A\pi \cdot \delta(|\omega| - \omega_0) e^{j\omega t} d\omega = A \cdot \frac{e^{j\omega_0 t} + e^{-j\omega_0 t}}{2} = A \cdot \cos \omega_0 t = x(t) .$$

Die Spektraldichte (1.17) wird ebenso wie die Fourier-Komponenten (1.13) mit „ X “ bezeichnet. Tatsächlich handelt es sich hierbei jedoch um unterschiedliche Größen. Wird z.B. $x(t)$ in „Volt“ gemessen, so haben die Fourier-Komponenten dieselbe Einheit, wohingegen die Spektraldichte in „Volt/Hertz“ anzugeben ist. Zwischen beiden wird hier durch die Argumente „ $jn\omega_0$ “ (periodischer Fall) und „ $j\omega$ “ (nichtperiodischer Fall) unterschieden, oder der Kontext macht die Unterscheidung möglich. Eine einzelne harmonische Komponente eines nichtperiodischen Signals mit der Kreisfrequenz ω hat also die Amplitude $(2\pi)^{-1}|X(j\omega)| d\omega = 0$, falls $|X(j\omega)| < \infty$.

Für die Untersuchung des Übertragungsverhaltens linearer Systeme spielt die Spektraldichte jedoch genau die gleiche Rolle wie das Spektrum. Insbesondere ist auch im nichtperiodischen Fall das Diagramm in der Abbildung 3 kommutativ, wenn anstelle der Spektren $X(jn\omega_0)$ und $Y(jn\omega_0)$ die Spektraldichten $X(j\omega)$ und $Y(j\omega)$ treten.

Setzt man nun in (1.16) für $y(t)$ das Faltungsintegral aus (1.4) ein, so findet man

$$\mathcal{F}[y(t)] = \mathcal{F}[x(t) * h(t)] = \int_{-\infty}^{+\infty} \left[\int_{-\infty}^{+\infty} x(\tau) h(t - \tau) d\tau \right] e^{-j\omega t} dt$$

$$\begin{aligned}
 &= \int_{-\infty}^{+\infty} x(\tau) \int_{-\infty}^{+\infty} h(t - \tau) e^{-j\omega t} dt d\tau \\
 &= \int_{-\infty}^{+\infty} x(\tau) \int_{-\infty}^{+\infty} h(t) e^{-j\omega(t+\tau)} dt d\tau \\
 &= \int_{-\infty}^{+\infty} x(\tau) e^{-j\omega\tau} \int_{-\infty}^{+\infty} h(t) e^{-j\omega t} dt d\tau \\
 &= \int_{-\infty}^{+\infty} x(\tau) e^{-j\omega\tau} \mathcal{F}[h(t)] d\tau \\
 &= \mathcal{F}[x(t)] \cdot \mathcal{F}[h(t)] .
 \end{aligned} \tag{1.20}$$

Damit haben wir eine wichtige Beziehung hergeleitet: Die Fourier-Transformierte $Y(j\omega)$ des Ausgangssignals $y(t)$ eines linearen Systems ist gleich dem Produkt der Übertragungsfunktion $H(j\omega)$ mit der Fourier-Transformierten $X(j\omega)$ des Eingangssignals $x(t)$, $Y(j\omega) = X(j\omega) \cdot H(j\omega)$. Dabei ist die Übertragungsfunktion die Fourier-Transformierte der Gewichtsfunktion $h(t)$,

$$H(j\omega) = \mathcal{F}[h(t)] .$$

Einer Faltung im Zeitbereich entspricht also eine Multiplikation im Frequenzbereich mit anschließender Rücktransformation,

$$x(t) * h(t) = \mathcal{F}^{-1}[X(j\omega) \cdot H(j\omega)] \tag{1.21}$$

In der Praxis stellt sich häufig das Problem, ein unbekanntes System zu identifizieren, unter der Annahme, dass dieses System linear ist. Zur Identifikation muss dann die unbekannte Übertragungsfunktion bestimmt werden. Nach (1.21) kann man ein bekanntes Testsignal $x(t)$ auf das System geben und am Ausgang $y(t)$ messen. Damit sind dann auch die Transformierten $Y(j\omega)$ und $X(j\omega)$ bestimmbar und die Übertragungsfunktion erhält man aus

$$H(j\omega) = Y(j\omega)/X(j\omega) \quad \text{für alle } \omega \text{ mit } X(j\omega) \neq 0 .$$

Dies nennt man *Entfaltung*. Ein geeignetes Testsignal, das $X(j\omega) = \text{const.} \neq 0$ für alle Frequenzen ω liefert, stellt ein Rauschsignal dar.²⁾

Man findet auch, dass einer Multiplikation im Zeitbereich eine Faltung im Frequenzbereich entspricht. Denn analog zu (1.20) erhalten wir

$$\mathcal{F}^{-1}[X(j\omega) * H(j\omega)] = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \left[\int_{-\infty}^{+\infty} X(j\Omega) \cdot H(j(\omega - \Omega)) d\Omega \right] e^{j\omega t} d\omega$$

²⁾Tatsächlich kann hier ein beliebiges Signal genommen werden, welches keine linearen statistischen Abhängigkeiten hat. Später werden wir solche Signale unkorreliert nennen.

$$\begin{aligned}
&= \frac{1}{2\pi} \int_{-\infty}^{+\infty} X(j\Omega) \int_{-\infty}^{+\infty} H(j(\omega - \Omega)) e^{j\omega t} d\omega d\Omega \\
&= \frac{1}{2\pi} \int_{-\infty}^{+\infty} X(j\Omega) \int_{-\infty}^{+\infty} H(j\omega) e^{jt(\omega + \Omega)} d\omega d\Omega \\
&= 2\pi \cdot \frac{1}{2\pi} \int_{-\infty}^{+\infty} X(j\Omega) e^{jt\Omega} d\Omega \cdot \frac{1}{2\pi} \int_{-\infty}^{+\infty} H(j\omega) e^{jt\omega} d\omega \\
&= 2\pi \cdot x(t) \cdot h(t) .
\end{aligned} \tag{1.22}$$

Folglich gilt

$$x(t) \cdot h(t) = \frac{1}{2\pi} \cdot \mathcal{F}^{-1}[X(j\omega) * H(j\omega)] \tag{1.23}$$

1.4 Abtasttheorem

Eine digitale Signalverarbeitung setzt neben einer Quantisierung der Amplitude auch eine Diskretisierung der Zeit voraus. Hier wird angenommen, dass ein ursprünglich zeit-kontinuierliches Signal $x(t)$ in äquidistanten Zeitabständen T_s abgetastet wird (s steht für „sampling“).

Bei der praktischen Realisierung der Abtastung mit einem getakteten Schalter schwankt die Zeit zwischen jeweils zwei aufeinanderfolgenden Abtastungen infolge verschiedener Störungen. Dieses Schwanken um eine mittlere Abtastperiode wird *Zeit-Jitter* genannt.³⁾ Derartige Effekte werden hier jedoch nicht betrachtet, so dass wir im Folgenden von einer konstanten Abtastperiode T_s ausgehen.

Durch die Abtastung des Signals $x(t)$ wird eine Folge von Abtastwerten erzeugt,

$$\{x(t)\}_{t \in \mathbb{R}} \rightarrow \{x(t_0 + mT_s)\}_m \equiv \{x_m\}_{m = \dots, -2, -1, 0, 1, 2, \dots} \tag{1.24}$$

Im Folgenden wird ohne Beschränkung der Allgemeinheit $t_0 = 0$ gesetzt. Will man die zeitlichen Änderungen des abgetasteten Signals möglichst genau durch die Folge der Abtastwerte widerspiegeln, so ist schon intuitiv verständlich, dass hierzu die Abtastperiode T_s hinreichend klein gewählt werden muss. Andererseits sollte aber T_s nicht zu klein sein, würde doch dann unnötig viel Platz zur Speicherung der Abtastwerte eines bestimmten Zeitabschnittes des Ausgangssignals benötigt. Darüber hinaus ist eine größere *Abtastfrequenz* $f_s \equiv T_s^{-1}$ in der Regel mit einem höherem gerätetechnischen Aufwand verbunden. Somit steht die Frage nach einer „maximalen“ Abtastperiode bzw. „minimalen“ Abtastfrequenz, so dass aus den Abtastwerten noch eindeutig auf die Werte des Ausgangssignals zwischen den Abtastzeitpunkten geschlossen werden kann. Hierzu macht das Abtasttheorem eine Aussage.

Für die Formulierung des Abtasttheorems wird zunächst der Begriff der *Nyquist-Frequenz* eines Signals $x(t)$ eingeführt als die kleinste (Kreis-)Frequenz ω_N , ab welcher das Spektrum

³⁾Allgemein versteht man unter dem Begriff „Jitter“ unerwünschte, kleine, schnelle Schwankungen in der Zeit oder der Amplitude eines Signals.

$X(j\omega)$ verschwindet, d.h., $|X(j\omega)| = 0$ für $\omega \geq \omega_N$ und $|X(j\omega)| \neq 0$ für ein $\omega \in [\omega_N - \varepsilon, \omega_N)$, $\varepsilon > 0$, beliebig klein.⁴⁾

Ist die Nyquist-Frequenz endlich, so bedeutet dies im Falle periodischer Signale $x(t)$, dass es eine natürliche Zahl N gibt, mit der Eigenschaft, dass die Fourier-Komponenten die Bedingungen $X(j(N-1)\omega_0) \neq 0$ und $X(jn\omega_0) = 0$ für alle $n \geq N$ erfüllt, wobei $\omega_0/2\pi$ die triviale Periode von $x(t)$ ist. Die Nyquist-Kreis-Frequenz von $x(t)$ ist dann $N\omega_0$.

Für nichtperiodische Signale $x(t)$ ist f_N durch die Eigenschaft bestimmt, dass es für ein beliebiges $\delta f > 0$ im Frequenzintervall $[f_N - \delta f, f_N]$ eine Frequenz f gibt, so dass die Spektraldichte $X(j2\pi f) \neq 0$, wohingegen $X(j2\pi f) = 0$ für alle $f > f_N$. Es gilt dann das folgende

Abtasttheorem (Probensatz, Sampling Theorem): Ist $x(t)$ ein zeit-kontinuierliches Signal mit der endlichen Nyquist-Frequenz $f_N < \infty$, so kann aus den Abtastwerten $x(mT_s) \equiv x_m$ eindeutig $x(t)$ für alle Zeitpunkte t bestimmt werden, sofern die Abtastfrequenz $f_s \equiv T_s^{-1}$ die folgende Bedingung erfüllt:

$$f_s > 2f_N \quad (1.25)$$

Die Interpolationsformel lautet

$$x(t) = \sum_{m=-\infty}^{+\infty} x_m \cdot \frac{\sin[\pi(t/T_s - m)]}{\pi(t/T_s - m)} \quad (1.26)$$

Formel (1.26) stellt eine unendliche Reihe dar, deren direkte numerische Bestimmung im Allgemeinen recht aufwendig ist. Darüber hinaus müsste zur vollständigen Rekonstruktion die Reihe für einen jeden Zeitpunkt t des Kontinuums bestimmt werden, was offenbar praktisch nicht realisierbar ist. Eine Lösung dieses Problems ergibt sich aus der folgenden Überlegung.

Begründung des Abtasttheorems

Zunächst wird aus den Abtastwerten x_m zu einem beliebigen $\varepsilon \in [0, 1]$ das zeit-kontinuierliche Signal gebildet,

$$x_\varepsilon(t) = \begin{cases} \frac{x_m}{\varepsilon} & : m - \varepsilon/2 \leq t/T_s < m + \varepsilon/2 \\ 0 & : \text{sonst.} \end{cases} \quad (1.27)$$

In der Abbildung 1.4 ist diese Folge von Rechteckimpulsen der Pulsdauer εT_s schematisch dargestellt.

Wird nun diese Pulsdauer verkleinert und gleichzeitig die Impulshöhe $x_m/(\varepsilon T_s)$ derart vergrößert, dass die Fläche unter dem m -ten Impuls konstant $x_m T_s$ ist, so erhält man im Grenzfall $\varepsilon \rightarrow 0$ für $x_\varepsilon(t)$ eine Folge von Dirac-Impulsen,

$$x_0(t) \equiv \lim_{\varepsilon \rightarrow 0} x_\varepsilon(t) = \sum_{m=-\infty}^{+\infty} x_m \cdot T_s \cdot \delta(t - mT_s) \quad (1.28)$$

⁴⁾In der Literatur (z.B. [OSB04]) wird die Größe $2 \cdot \omega_N$ Nyquist-Rate genannt. Manche Autoren benutzen hierfür jedoch auch den Begriff Nyquist-Frequenz, so dass immer aus dem Kontext die richtige Interpretation gefunden werden muss.

Wird nun $x_0(t)$ nach Gleichung (1.17) fourier-transformiert, so folgt,

$$\begin{aligned}
 X_0(j\omega) &= \int_{-\infty}^{+\infty} x_0(t) \cdot e^{-j\omega t} dt \\
 &= T_s \sum_{m=-\infty}^{+\infty} \left[x_m \int_{-\infty}^{+\infty} \delta(t - mT_s) e^{-j\omega t} dt \right] \\
 &= T_s \sum_{m=-\infty}^{+\infty} x_m \cdot \exp\left(-j2\pi \frac{m\omega}{\omega_s}\right) .
 \end{aligned} \tag{1.29}$$

Offenbar gilt $X_0(j\omega) = X_0(j(\omega + n \cdot \omega_s))$ für ganzzahlige Werte von n , das heißt, die Spektraldichte von $x_0(t)$ ist eine periodische Funktion mit der Periode ω_s .

Um $X_0(j\omega)$ genauer zu bestimmen, wird zunächst $x_0(t)$ umgeformt. Unter Beachtung der Eigenschaften des Dirac-Impulses, kann (1.28) wie folgt geschrieben werden,

$$\begin{aligned}
 x_0(t) &= \sum_{m=-\infty}^{+\infty} x_m \cdot T_s \cdot \delta(t - mT_s) \\
 &= T_s \cdot \sum_{m=-\infty}^{+\infty} x(mT_s) \cdot \delta(t - mT_s) \\
 &= T_s \cdot x(t) \cdot \sum_{m=-\infty}^{+\infty} \delta(t - mT_s) \\
 &= T_s \cdot x(t) \cdot \delta_{T_s}(t) ,
 \end{aligned} \tag{1.30}$$

mit der T_s -periodischen Diracschen δ -Funktion (Distribution)

$$\delta_{T_s}(t) \equiv \sum_{m=-\infty}^{+\infty} \delta(t - mT_s) . \tag{1.31}$$

Die Fourier-Komponenten von $\delta_{T_s}(t)$ sind nach (1.13) alle gleich T_s^{-1} , denn mit $T_0 = T_s$ folgt

$$\mathcal{F}\left(\delta_{T_s}(t)\right) = T_s^{-1} \int_{-T_s/2}^{T_s/2} \delta_{T_s}(t) \cdot e^{-jn\omega_0 t} dt$$

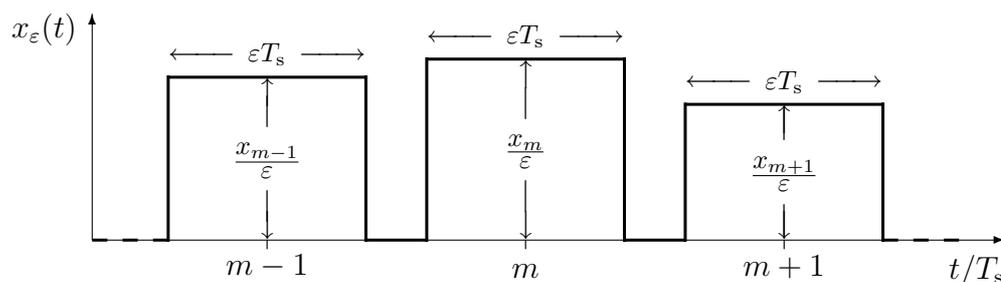


Bild 1.4: Signal der Abtastimpulse

$$\begin{aligned}
 &= T_s^{-1} \int_{-T_s/2}^{T_s/2} \delta(t) \cdot e^{-jn\omega_0 t} dt \\
 &= T_s^{-1}, \quad \text{für alle } n = 0, \pm 1, \pm 2 \dots
 \end{aligned}$$

Demzufolge hat δ_{T_s} die Fourier-Reihe-Darstellung (1.12)

$$\delta_{T_s}(t) = T_s^{-1} \sum_{n=-\infty}^{+\infty} e^{jn\omega_s t}. \quad (1.32)$$

Für $x_0(t)$ in (1.30) folgt somit die Darstellung

$$x_0(t) = x(t) \sum_{n=-\infty}^{+\infty} e^{jn\omega_s t}. \quad (1.33)$$

Wird $x_0(t)$ nun nach (1.17) fourier-transformiert, so erhält man die Spektraldichte

$$\begin{aligned}
 \mathcal{F}[x_0(t)] \equiv X_0(j\omega) &= \int_{-\infty}^{+\infty} x_0(t) e^{-j\omega t} dt \\
 &= \sum_{n=-\infty}^{+\infty} \int_{-\infty}^{+\infty} x(t) e^{j(n\omega_s - \omega)t} dt \\
 &= \sum_{n=-\infty}^{+\infty} X[j(n\omega_s - \omega)]. \quad (1.34)
 \end{aligned}$$

Die Spektraldichte der Abtastimpulsfolge $x_0(t)$ entsteht also, indem man die Spektraldichte $\mathcal{F}(x(t)) = X(j\omega)$ des abgetasteten Signals $x(t)$ nimmt, um alle ganzzahligen Vielfache der Abtast-Kreisfrequenz ω_s auf der Frequenzachse verschiebt und dann aufsummiert.

Für den Fall, dass die Nyquist-Frequenz $f_N \equiv \omega_N/2\pi$ von $x(t)$ die Abtastbedingung (1.25) erfüllt, sind die Beträge der Spektraldichten $X(j\omega)$ und $\mathcal{F}[x_0(t)] \equiv X_0(j\omega)$ in Abb. 1.5a) und b) schematisch dargestellt. Für die Rekonstruktion des ursprünglichen Signals $x(t)$ aus der Folge von Abtastwerten x_m ist es wesentlich, dass die Spektraldichten $X(j\omega)$ und $X_0(j\omega)$ im Intervall $[-\omega_N, \omega_N]$ übereinstimmen. Zur Rekonstruktion müßte dann das Signal $x_0(t)$ auf einen „idealen“ Tiefpaß gegeben werden, der durch die Übertragungsfunktion

$$H_{T,\omega_N}(j\omega) = \begin{cases} 1 & : -\omega_N \leq \omega \leq \omega_N \\ 0 & : \text{sonst.} \end{cases} \quad (1.35)$$

charakterisiert ist. Offenbar gilt dann

$$x(t) = \mathcal{F}^{-1}[H_{T,\omega_N}(j\omega) \cdot X_0(j\omega)] \quad (1.36)$$

Falls aber die Bedingung (1.25) infolge einer zu niedrig gewählten Abtastfrequenz verletzt wird, führt dies zu „Überlappungen“ der Frequenzbänder $X[j(n\omega_s - \omega)]$ in Gleichung (1.34), wie es in Abb. 1.5c) und d) schematisch dargestellt ist. Man nennt dies *aliasing* (gesprochen wie

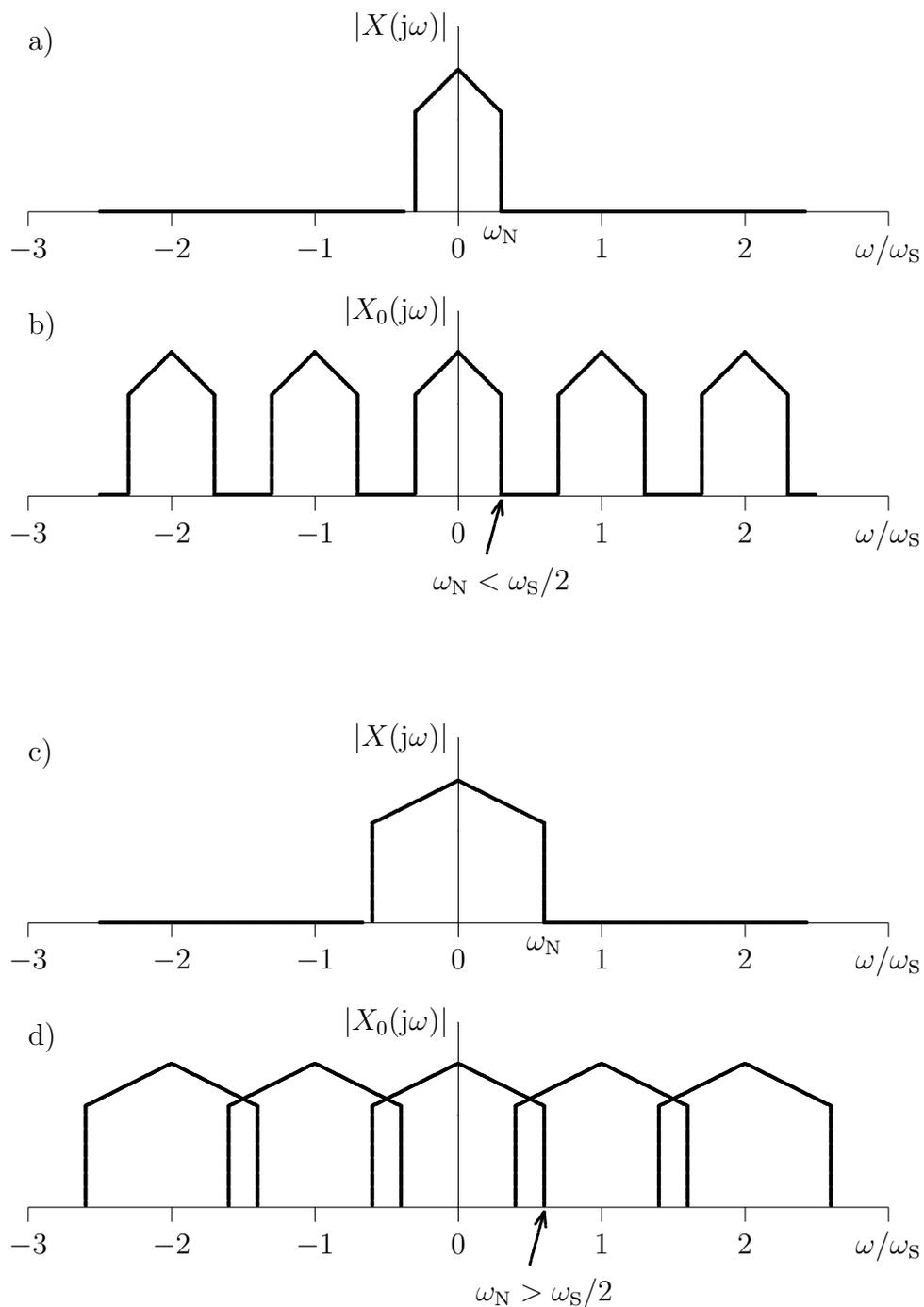


Bild 1.5: Schematische Darstellung der Spektraldichten der Signale $x(t)$ und $x_0(t)$ im Falle der Erfüllung der Abtastbedingung (a bzw. b) und ihrer Verletzung (c bzw. d). In d) sind die Beträge der Summanden der Reihe in Gl. (1.34) dargestellt. Im Überlappungsbereich $\omega = \omega_s - \omega_N \dots \omega_N$ kann aus dem Abtastsignal $x_0(t)$ das ursprüngliche Signal $x(t)$ nicht eindeutig rekonstruiert werden

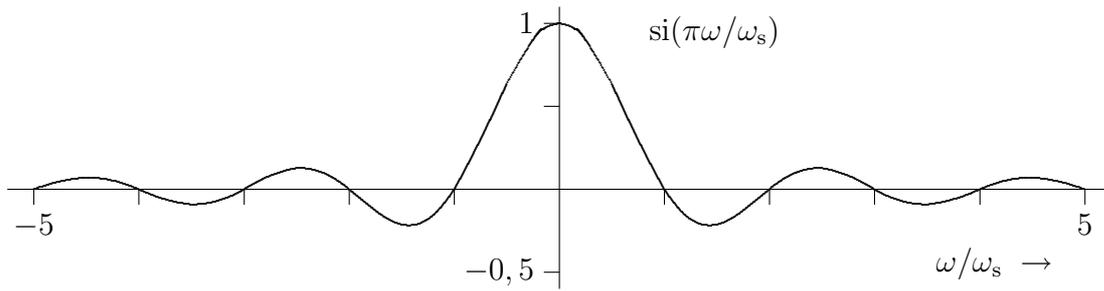


Bild 1.6: Graph der Spaltfunktion $\text{si}(\pi\omega/\omega_s) \equiv \frac{\sin(\pi\omega/\omega_s)}{\pi\omega/\omega_s}$

„eiliesing“). Hierbei weichen die Spektraldichten $X(j\omega)$ und $X_0(j\omega)$ im Intervall $[-\omega_N, \omega_N]$ voneinander ab, und eine eindeutige Rekonstruktion von $x(t)$ aus den Abtastwerten ist prinzipiell nicht mehr möglich.

Wird die Abtastbedingung (1.25) nur geringfügig verletzt, gilt z.B. $\omega_N > \omega_s/2$ mit $(\omega_N - \omega_s/2)/\omega_N \ll 1$, so kann es in der Praxis noch sinnvoll sein, aus den „unterabtasteten“ Werten x_m das ursprüngliche Signal $x(t)$ zu approximieren, indem für den Rekonstruktionstiefpaß in (1.36) anstelle der Grenzfrequenz $\omega_N/2\pi$ nun $(\omega_s - \omega_N)/2\pi$ gewählt wird. Das so erhaltene Signal $x^*(t) \equiv \mathcal{F}^{-1}[H_{T, \omega_s - \omega_N}(j\omega) \cdot X_0(j\omega)]$ unterscheidet sich dann von $x(t)$ nur in der Spektraldichte an den Stellen $|\omega| > \omega_s - \omega_N$.

Zusammenfassend lautet die Vorschrift zur Rückgewinnung des Analogsignals $x(t)$ aus den Abtastwerten x_m also wie folgt: Aus den Abtastwerten x_m wird das zeit-kontinuierliche Signal $x_0(t)$ von Dirac-Delta-Impulsen gebildet und auf einen „idealen“ Tiefpaß gegeben, mit der Nyquist-Frequenz von $x(t)$ als Grenzfrequenz des Tiefpasses. Das rekonstruierte Signal $x(t)$ kann am Ausgang des Tiefpasses abgegriffen werden.

Praktische Rekonstruktion

Bei der praktischen Rückgewinnung des originalen Signals aus Abtastwerten treten jedoch verschiedene Probleme auf. So kann man z.B. mit elektronischen Schaltungen keine Dirac-Impulse erzeugen. Ein reales Impulssignal gleicht eher der in der Abbildung 1.4 dargestellten Funktion $x_\varepsilon(t)$ mit $\varepsilon > 0$. Die Fourier-Transformierte hiervon lautet

$$X_\varepsilon(j\omega) = \frac{\sin(\pi\varepsilon\omega/\omega_s)}{\pi\varepsilon\omega/\omega_s} \cdot X_0(j\omega) . \quad (1.37)$$

Die Spektraldichten $X_{\varepsilon>0}(j\omega)$ und $X_0(j\omega)$ unterscheiden sich also nur in einer Gewichtsfunktion, welche die *Spaltfunktion*

$$\text{si}(\pi\varepsilon\omega/\omega_s) \equiv \frac{\sin(\pi\varepsilon\omega/\omega_s)}{\pi\varepsilon\omega/\omega_s} \quad (1.38)$$

darstellt (s. Abbildung 1.6).⁵⁾ Im Fall $\varepsilon = 1$ stellt $x_\varepsilon(t)$ eine Treppenfunktion dar, mit den Abtastwerten x_m als Stufenhöhe. Dies ist besonders unter praktischen Gesichtspunkten interessant, denn $x_1(t)$ kann relativ einfach schaltungstechnisch realisiert werden. Die zugehörige

⁵⁾ Statt *si* schreibt man zuweilen auch *sinc* für *sinus cardinalis*. Im Englischen heißt die Spaltfunktion auch *interpolation function*, *filtering function* oder *first spherical Bessel function*.

Spektraldichte lautet

$$X_1(j\omega) = \text{si}\left(\pi\frac{\omega}{\omega_s}\right) \cdot X_0(j\omega) \quad . \quad (1.39)$$

Die Gewichtsfunktion schwächt besonders höhere Frequenzen ab. Bei der halben Abtastfrequenz ist sie beispielsweise $\text{si}(\pi/2) = 2/\pi = 0,636\dots$

Der Einfluß der Gewichtsfunktion hinsichtlich der Abweichung zwischen den Spektraldichten $X_1(j\omega)$ und $X_0(j\omega)$ wird umso geringer, je höher das Signal $x(t)$ überabgetastet wird. Gilt z.B. $\omega_s = 10 \omega_N$, was mit Blick auf die Abtastbedingung (1.25) eine 5-fache Überabtastung bedeutet, so wird $\text{si}(\pi\omega_N/\omega_s) = \text{si}(\pi/10) = 0,983\dots$ Im relevanten Frequenzintervall $[-\omega_N, \omega_N]$ ist also der relative Fehler, den man macht, wenn anstelle der Dirac-Impulsfolge $x_0(t)$ die Treppenfunktion $x_1(t)$ auf den Rekonstruktionstiefpaß gegeben wird, kleiner als 2%. Bei vielen Anwendungen, z.B. in der Musikelektronik, ist dieser Fehler vernachlässigbar.

Sollte eine Überabtastung nicht möglich sein, oder ist eine besonders präzise Rekonstruktion des Analogsignals erforderlich, so kann die Abschwächung höherer Frequenzen ausgeglichen werden, indem anstelle eines einfachen Rekonstruktionstiefpasses sogenannte Formfilter (FF) verwendet werden. Diese haben ebenfalls eine Tiefpaß-Charakteristik, gleichen aber für niedrige Frequenzen ($< \omega_N$) den Einfluß der Gewichtsfunktion aus. Offenbar muß dazu die zugehörige Übertragungsfunktion $H_{FF}(j\omega)$ für Frequenzen $\omega < \omega_N$ die Übertragungsfunktion $|H_{FF}(j\omega)| = [\text{si}(\pi\omega/\omega_s)]^{-1}$ haben.

In der Praxis ist häufig die Sampling-Frequenz durch die Apparatur zur Datenerfassung fest vorgegeben oder nur in engen Grenzen wählbar. Soll nun ein Analogsignal $x(t)$ abgetastet werden, dessen Nyquist-Frequenz größer als die halbe Sampling-Frequenz ist, so ist dies zwar formal möglich, allerdings würde der Tiefpaß zur Rekonstruktion des Analogsignals anstelle von $x(t)$ ein davon verschiedenes Signal $x^*(t)$ liefern. Um zumindest im Frequenzband $[-\omega_s/2, \omega_s/2]$ eine umkehrbar eindeutige Entsprechung zwischen dem originalen Signal und den Abtastwerten zu haben, muß $x(t)$ vor der Abtastung mit einem Tiefpaß gefiltert werden, dessen Grenzfrequenz kleiner als $f_s/2$ ist. Entsprechend der Steilheit der Filtercharakteristik und gemäß den konkreten Anforderungen an die Unterdrückung von Frequenzen, die größer als $f_s/2$ sind, kann dabei die Grenzfrequenz mehr oder weniger dicht bei $f_s/2$ liegen. Diesen Tiefpaß nennt man *Anti-Aliasing-Filter*. Wird ein Signal zunächst unterabgetastet ($f_s < 2f_N$), so ist eine spätere Filterung mit dem Ziel, eine Übereinstimmung zwischen originalem Signal und den zeit-diskreten Werten im Frequenzband $[-\omega_s/2, \omega_s/2]$ zu erzielen, prinzipiell nicht mehr möglich!

Abbildung 1.7 fasst diese in einem allgemeinen Schema der Datenaufnahme und Rekonstruktion zusammen. Bei einer digitalen Weiterverarbeitung der Daten würde die Rekonstruktion durch Interpolationsalgorithmen realisiert.

1.5 Diskrete Fourier-Transformation

Hat man ein Signal $x(t)$ mit der Abtastperiode T_s unter Einhaltung der Abtastbedingung (1.25) zeitlich diskretisiert, so gibt es nach dem Abtasttheorem eine umkehrbar eindeutige Zuordnung zwischen Abtastwerten x_m und dem originalen zeit-kontinuierlichen Signal $x(t)$. Will man nun von $x(t)$ das Spektrum (1.13) bzw. nach (1.17) die Spektraldichte $X(j\omega)$ bestimmen, so stellt sich die Frage, wie man dies direkt aus den Abtastwerten erreichen kann. Eine Antwort gibt Formel (1.29). Wird hier durch T_s dividiert, so erhält man die *Zeit-Diskrete Fourier-Transformation*

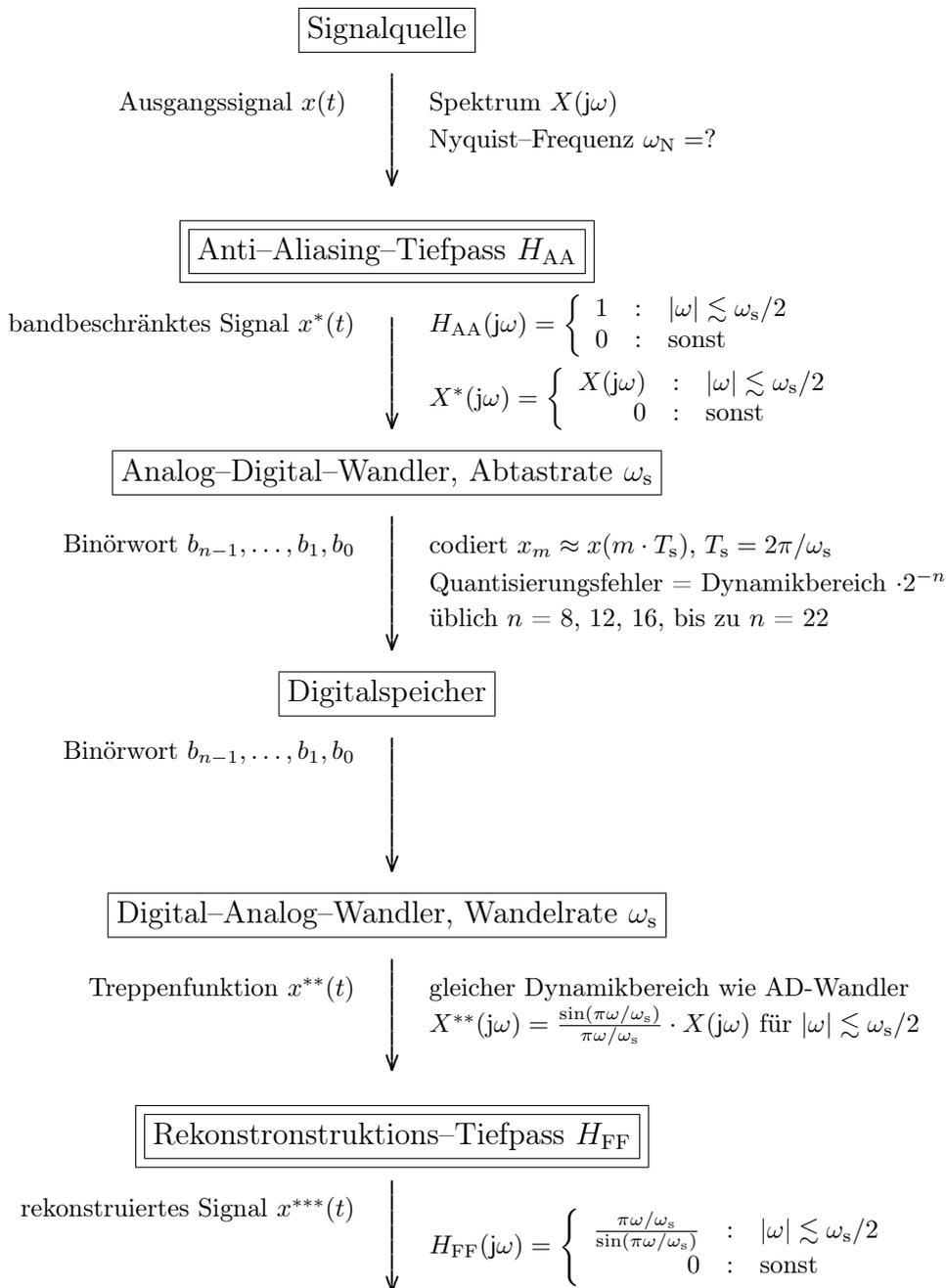


Bild 1.7: Allgemeine Anordnung zur Aufnahme und Rekonstruktion zeitkontinuierlicher Signale $x(t)$. Abweichungen des rekonstruierten Signals von $x^{***}(t)$ von $x(t)$ entstehen infolge von (i) Bandbeschränkungen nach Anti-Aliasing-Filterung und (ii) Quantisierungsrauschen. Ist letzteres vernachlässigbar, so stimmen $x^{***}(t)$ und $x(t)$ zumindest im Frequenzband $0 \dots \omega_s/2$ überein

der Abtastfolge $\{x_m\}$,⁶⁾

$$\mathcal{F}_D[x_m](\omega) \equiv \sum_{m=-\infty}^{+\infty} x_m \cdot \exp\left(-j2\pi \frac{m\omega}{\omega_s}\right) . \quad (1.40)$$

Wie schon oben bemerkt, ist $\mathcal{F}_D[x_m](\omega)$ periodisch in ω , mit der Periode ω_s , und im Intervall $[0, \omega_s]$ gilt $X(j\omega) = T_s \cdot \mathcal{F}_D[x_m](\omega)$. Man betrachtet deshalb die Frequenz ω nur im Intervall $[0, \omega_s]$. Wird darüber hinaus ω für ganzzahlige Vielfache von ω_s/N berechnet, also für

$$\omega = (n/N) \cdot \omega_s , \quad n = 0, 1, 2, \dots, N-1 ,$$

so folgt

$$T_s^{-1} \cdot X\left(j \frac{n}{N} \cdot \omega_s\right) = \mathcal{F}_D[x_m]\left(\frac{n}{N} \cdot \omega_s\right) = \sum_{m=-\infty}^{+\infty} x_m \cdot \exp\left(-j2\pi \frac{mn}{N}\right)$$

Diese Gleichung liefert noch immer keine praktikable Vorschrift zur numerischen Bestimmung des Spektrums eines Signals aus seinen Abtastwerten, können doch unendliche Reihen im Allgemeinen nicht bestimmt bzw. nur mit großen Aufwand geschätzt werden. Ist das Signal jedoch stationär oder zumindest quasistationär, so kann anstelle der unendlichen Reihe eine Summation über nur endlich viele ($N < \infty$) Glieder der Reihe betrachtet werden.

Dies führt schließlich auf die sogenannte *Diskrete Fourier-Transformation* (DFT),

$$X_n = \sum_{m=0}^{N-1} x_m \cdot \exp\left(-j2\pi \frac{mn}{N}\right) \quad (1.41)$$

Die Summation über der Zeit (Index m) läuft hier nur in einem endlichen Bereich, was dem Zeitbereich $0 \dots (N-1) \cdot T_s$ entspricht. Dies ist streng genommen nur dann zulässig, wenn $x_m = x(m \cdot T_s) = 0$ für alle $m < 0$ und $m \geq N$. Das ist aber das gleiche, als würden wir nicht das Signal $x(t)$ abtasten und transformieren, sondern das Produkt aus $x(t)$ mit der Fensterfunktion (Rechteckfenster) $h(t) = 1$ für $t \in [0, N \cdot T_s)$ und $h(t) = 0$ sonst,

$$\tilde{x}(t) = x(t) \cdot h(t) .$$

Nach (1.23) hat dieses Signal das Spektrum

$$\mathcal{F}[\tilde{x}(t)] = \frac{1}{2\pi} \cdot [X(j\omega) * H(j\omega)] .$$

Man berechnet also letztlich die Faltung der gesuchten Spektraldichte $X(j\omega)$ mit der Spektraldichte der Fensterfunktion (vgl. Rechnung auf S. 129),

$$H(j\omega) = \mathcal{F}[h(t)] = NT_s \cdot \frac{\sin(\omega NT_s/2)}{\omega NT_s/2} \cdot e^{-j\omega NT_s/2}$$

Es gilt dann

$$X_n = f_s \cdot \tilde{X}\left(j \frac{n}{N} \cdot \omega_s\right) .$$

⁶⁾Bemerkenswert ist die Dualität zu den kontinuierlichen periodischen Zeitsignalen, deren Spektren diskret sind, wohingegen hier die diskreten Zeitsignale kontinuierliche periodische Spektren haben.

Der Betrag der n -ten Fourier-Komponente $|X_n|$ ist also ein Maß für die Amplitude der harmonischen Komponente von $x(t)$ mit der Kreisfrequenz $(n/N)\omega_s$.

Die zugehörige *Inverse Diskrete Fourier-Transformation* (IDFT) lautet

$$x_m = \frac{1}{N} \sum_{n=0}^{N-1} X_n \cdot \exp\left(j2\pi \frac{mn}{N}\right) \quad (1.42)$$

Allgemeine Eigenschaften der DFT

1. Ist x_m für jedes $m = 0, 1, \dots, N - 1$ rein reell, so ist die DFT symmetrisch,

$$X_n = X_{N-n-1} \quad \text{für } n = 0, 1, \dots, N/2 - 1$$

2. Wird der Normierungsfaktor $1/N$ bei der Rücktransformation (1.42) symmetrisch auf die Hin- und Rücktransformierte aufgeteilt (jeweils Faktor $1/\sqrt{N}$) dann ist die DFT eine unitäre Transformation von $\{x_m\}_{m=0}^{N-1}$.

Hinweise zur Berechnung der DFT

Die DFT hat in der digitalen Signalverarbeitung eine außerordentlich große Bedeutung, insbesondere, weil für $N = 2^k$ mit $k = 3, 4, 5, \dots$ ein schneller numerischer Algorithmus zur Ausführung der DFT existiert, die sogenannte *Schnelle Fourier-Transformation* (*fast fourier transformation*: FFT).

Die direkte Berechnung nach

$$\begin{pmatrix} X_0 \\ X_1 \\ \vdots \\ X_{N-1} \end{pmatrix} = \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & \exp(-j2\pi \frac{1 \cdot 1}{N}) & \dots & \exp(-j2\pi \frac{(N-1) \cdot 1}{N}) \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \exp(-j2\pi \frac{1 \cdot (N-1)}{N}) & \dots & \exp(-j2\pi \frac{(N-1) \cdot (N-1)}{N}) \end{pmatrix} \times \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_{N-1} \end{pmatrix} \quad (1.43)$$

benötigt $2 \cdot N^2$ Multiplikationen mit reelwertigen Faktoren. Ist jedoch die Blocklänge N keine Primzahl, so kann der Rechenaufwand erheblich verringert werden, insbesondere dann, wenn N als ganzzahlige Potenz von 2 gewählt wird, beispielsweise $N = 256; 512; 1024; 2048$. Der schnelle Algorithmus nutzt den Umstand aus, dass dann gewisse Multiplikationen, die bei der direkten Implementierung mehrfach ausgeführt würden, nur einmal realisiert werden. Im Falle $N = 8$ ist dies aus der folgenden Darstellung der Transformationsmatrix \mathcal{U} leicht einzusehen ist (es werden nur die Produkte $mn \bmod N$, der Elemente $\exp(-j2\pi \frac{m \cdot n}{N})$ der Transformationsmatrix in (1.41) dargestellt)

$$\mathcal{U}_8 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 0 & 2 & 4 & 6 & 0 & 2 & 4 & 6 \\ 0 & 3 & 6 & 1 & 4 & 7 & 2 & 5 \\ 0 & 4 & 0 & 4 & 0 & 4 & 0 & 4 \\ 0 & 5 & 2 & 7 & 4 & 1 & 6 & 3 \\ 0 & 6 & 4 & 2 & 0 & 6 & 4 & 2 \\ 0 & 7 & 6 & 5 & 4 & 3 & 2 & 1 \end{pmatrix}$$

Zwei gleiche Zahlen in der Spalte bedeutet, das bei der direkten Implementierung eine Multiplikation unnötigerweise zweimal ausgeführt wird. Beispielsweise würde hier nach (1.43) der Abtastwert x_2 zweimal mit dem Faktor $\exp(-j2\pi\frac{2\cdot 1}{N}) = \exp(-j2\pi\frac{2\cdot 5}{N})$ multipliziert werden, um X_1 bzw. X_5 zu berechnen.

Man findet nun im Allgemeinen, dass die Anzahl der Multiplikationen für große Blocklängen N von der Größenordnung $N \log_2 N$ ist. Das bedeutet gegenüber N^2 bei der direkten Implementierung eine erhebliche Reduzierung der algorithmischen Komplexität. Beispielsweise reduziert sich der Rechenaufwand für $N = 2^{10}$ auf den Bruchteil $(N \log_2 N)/N^2 = \log_2 N/N = 10/1024$, also auf etwa 1% einer direkten Berechnung.

Ein PASCAL-Code zur FFT findet sich im Anhang (Seite 131).

Zwei-dimensionale DFT

Die 2-dimensionale DFT ist beispielsweise wichtig für Probleme der Bildverarbeitung und -codierung von großer Bedeutung.

Überdecken wir ein Bild mit $M \times N$ vielen gleich großen Boxen (Quadraten) B_{mn} und ordnen ihnen die Amplitude x_{mn} zu, welche eine Intensität oder Helligkeit darstellen könnte, oder einen mittleren Grauwert der Box B_{mn} . Der Datensatz

$$\{x_{mn}\}_{m,n=1}^{M,N}$$

kann auch *digitales $M \times N$ Bild* genannt werden. Die *Zwei-dimensionale diskrete Fourier-Transformation* (2D-DFT) eines $N \times M$ -Bildes ist dann definiert durch

$$X_{kl} \equiv \frac{1}{\sqrt{NM}} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} x_{mn} \exp \left[-j2\pi \left(\frac{ln}{N} + \frac{km}{M} \right) \right]$$

für $k = 0, 1, \dots, M-1, l = 0, 1, \dots, N-1$. Und die *zwei-dimensionale inverse diskrete Fourier-Transformation* (2D-IDFT) ist gegeben durch

$$x_{mn} = \frac{1}{\sqrt{NM}} \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} X_{kl} \exp \left[+j2\pi \left(\frac{ln}{N} + \frac{km}{M} \right) \right]$$

Eine geringfügige Umformung liefert

$$X_{kl} = \frac{1}{\sqrt{M}} \sum_{m=0}^{M-1} \exp \left(-j2\pi \frac{km}{M} \right) \left[\frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x_{mn} \exp \left(-j2\pi \frac{ln}{N} \right) \right].$$

Der in eckigen Klammern stehende Term kann als 1D-FFT einer Bildspalte bei festgehaltener Bildzeile m interpretiert werden. Die Summation über m ist letztlich eine weitere 1D-FFT. Zur effektiven Berechnung kann dazu in der Praxis jeweils der schnelle Algorithmus, also die 1D-FFT angewandt werden.

Kurzzeit–Fourieranalyse

Bei Signalen interessiert man sich häufig für Änderungen des Signalcharakters im Zeitablauf. Typische Beispiele hierfür stellen Sprachsignale dar. Die fortlaufende Änderung der Geometrie der Artikulationstraktes (Stellung der Unterkiefer, der Zunge usw.) sowie die Änderung der Erregungsbedingungen (Schwingung der Stimmritze im Kehlkopf, Turbulenzen an Engstellen im Artikulationstrakt u.v.a.m.) machen erst natürliche Sprache aus.

Um langsam veränderliche Parameter von Signalen zu erfassen, schneidet man aus dem Signal einen gewissen Abschnitt, ein sogenannten *Datenfenster* heraus und analysiert zunächst nur dieses. Dann wird ein weiteres Fenster ausgeschnitten und analysiert usw. Formal kann man dafür wie folgt schreiben: Zum Signal

$$\{x_m\}_{m=-\infty}^{+\infty}$$

erhalten wir das Fenster

$$\tilde{x}_n^{(m)} \equiv h_n x_{m+n}, \quad n = -\infty \dots +\infty .$$

$\{h_n\}_{n=-\infty}^{+\infty}$ nennt man Fensterfunktion. Einige häufig angewandte Fensterfunktionen sind:

Rechteck–Fenster der Länge N :

$$h_n = \begin{cases} 1 & : 0 \leq n < N \\ 0 & : \text{sonst.} \end{cases}$$

Hamming–Fenster der Länge N :

$$h_n = \begin{cases} 0,54 - 0,46 \cdot \cos \frac{2\pi n}{N-1} & : 0 \leq n < N \\ 0 & : \text{sonst.} \end{cases}$$

Trägt man nun den Betrag des Fourier–Spektrums grau codiert vertikal für die Daten–Fenster zu den verschiedenen Zeitpunkten auf, dann erhält man das sogenannte *Periodogramm* des Signals.

Periodogramme von Sprachsignalen

Abbildung 1.8 zeigt Periodogramme für verschiedene stimmhafte Sprachsignale, *Vokale*. Dies sind annähernd periodische Signale, deren Grundfrequenz auch *Pitchfrequenz* genannt wird. Sie spiegelt die Sprachmelodie (Grundfrequenz der Schwingung der Stimmbänder) wieder und liegt bei männlichen Sprechern etwa zwischen 50 und 150 Hz, bei Frauen bis über 250 Hz. Zusammen mit den Oberwellen erzeugt sie die horizontal verlaufende Streifenstruktur. Überlagert werden die Muster von breiteren dunklen Streifen, deren Mittenfrequenzen die sogenannten *Formantfrequenzen* bilden. Diese sind von Vokal zu Vokal unterschiedlich und spiegeln die unterschiedliche Geometrie von Rachen– und Mundraum bei der Spracherzeugung wider. Mittels der Formantfrequenzen kann deshalb eine Spracherkennung erfolgen.

Abbildung 1.9 das Periodogramm einer Probe fließender Sprachprobe. Hier ist eine grössere Variabilität der Sprachmelodie zu erkennen, beispielsweise steigt die Grundfrequenz bei $t \approx 27000T_s$ deutlich an (Stimme wird höher). Auch die Formantfrequenzen ändern sich zeitlich, was jedoch der Darstellung nicht mehr so deutlich zu entnehmen ist. In der On–line Sprachverarbeitung werden deshalb ausgefeiltere Methoden zur Spracherkennung verwendet.

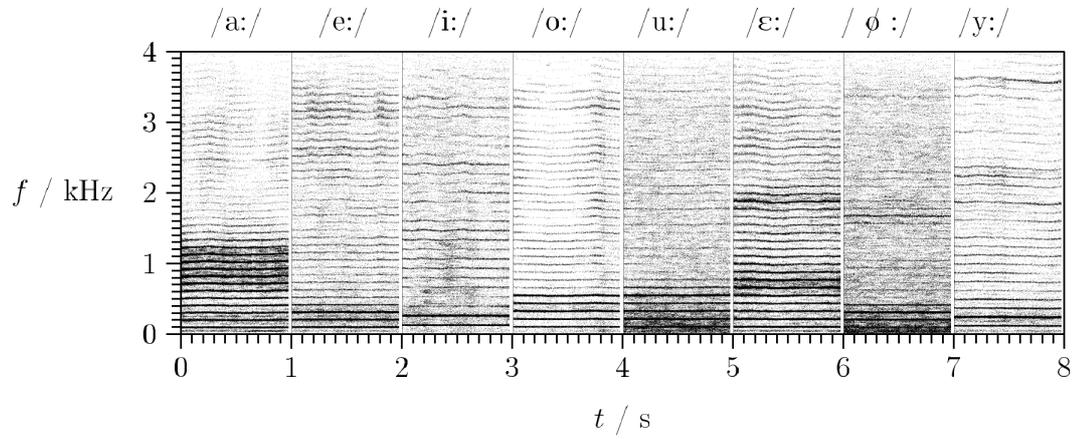


Bild 1.8: Spektrogramm von isoliert gesprochenen Vokalen (männlicher Sprecher im Brustregister)

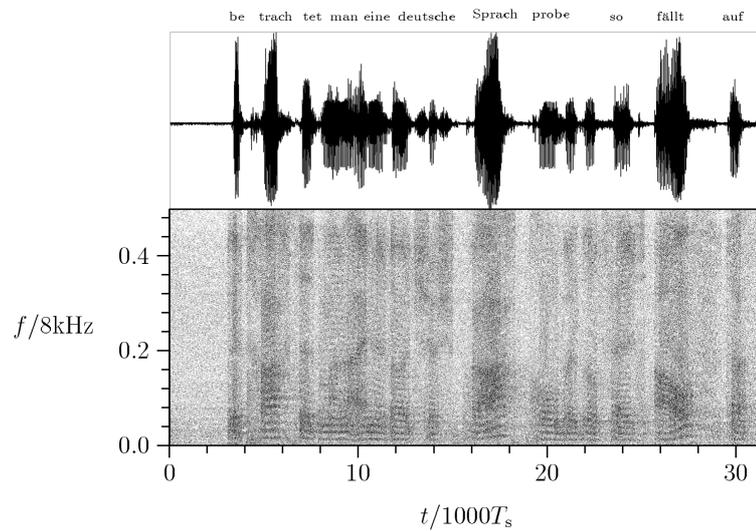


Bild 1.9: Oben: Sprachsignal eines männlichen Sprechers im modalen Register (Abtastperiode $T_s = 125\mu s$). Unten: zugehöriges Periodogramm

Kapitel 2

Interpolation (engl.)

2.1 Introduction

In this chapter digital filters are designed for interpolation and differentiation combined with interpolation of time–discrete signals. Upper bounds of the interpolation resp. differentiation errors are derived which depend on the Nyquist frequency of the signal and its power. The filters require low computational efforts.

In digital signal processing it is a standard task to interpolate and/or differentiate a signal which is discrete in time and amplitude. The sampling theorem says that interpolation is in principle uniquely possible if the sampling frequency f_s is at least twice the Nyquist frequency f_N of the signal, and it gives an interpolation formula which represents an infinite sum. However, in practice we would rather use any finite digital filter (e.g. FIR filter) which speeds up the procedure. But typically there remains an interpolation error which corrupts the interpolated signal.

In this paper we first (section 2) design digital filter for interpolation of a uniformly sampled bandlimited signal using Lagrange polynomials. Our point is to derive formulae representing upper bounds of the interpolation error as a function of the Nyquist frequency f_N of the original signal. The error can be arbitrarily diminished by increasing the "degree" of interpolation, provided f_N is related to the sampling frequency as $f_N < f_s/\pi$, i.e., our error estimation works only in the case of at least slight oversampling. However, from a practical point of view this is no essential restriction. Moreover, we find upper bounds for the interpolation errors which are caused by the quantization error and/or possible corruptions by noise of the sampled data.

In Sec. 3 we derive digital filters performing interpolation combined with differentiation simultaneously. Again, our point is to give precise error criterias in a similar way as for the interpolation filters of Sec.2.

The usefulness of our proposals is illustrated in the construction of Poincaré maps from a discrete time series and for higher–dimensional embeddings of a scalar time series.

2.2 Interpolation

Consider a band–limited time continuous signal $s^*(t)$ with a finite Nyquist frequency f_N . Let $f_s \equiv T_s^{-1}$ denote the sampling frequency, which should be constant, with $2f_N < f_s$. Then, according to the sampling theorem, from the discrete sequence

$$\{s^*(t_m) \equiv s(m)\}_{m=1}^M, \quad t_m = mT_s, \quad (2.1)$$

we can unambiguously reproduce the original continuous signal $s(t)$ by interpolation. Using Lagrangian polynomials of order n

$$L_i^*(t) = \prod_{l=0, l \neq i}^n \frac{t - t_{m+l}}{t_{m+i} - t_{m+l}} \quad (2.2)$$

we have

$$\hat{s}^*(t) = \sum_{i=0}^n L_i^*(t) s^*(t_{m+i}), \quad \text{for } t_m \leq t \leq t_{m+n} . \quad (2.3)$$

The interpolation error is given by

$$R^*(t, n) \equiv s^*(t) - \hat{s}^*(t) = \frac{d^{n+1}}{dt^{n+1}} s^*(t) \Big|_{\vartheta(t)} \frac{1}{(n+1)!} \prod_{i=0}^n (t - t_{m+i}) , \quad (2.4)$$

where $t_m \leq \vartheta(t) \leq t_{m+n}$.

Henceforth time t is measured in units of the sampling period T_s , thus (2.2) can be rewritten as

$$L_i^*(t) = \prod_{l=0, l \neq i}^n \frac{t - (m+l)}{(m+i) - (m+l)} = \prod_{l=0, l \neq i}^n \frac{t - m - l}{i - l} .$$

Writing $\tau \equiv t - m$ we get somewhat more conveniently:

$$L_i(\tau) \equiv L_i^*(m + \tau) = \frac{(-1)^{n-i}}{i! (n-i)!} \prod_{l=0, l \neq i}^n (\tau - l) , \quad (2.5)$$

for $0 \leq \tau \leq n$, $(-1)^0 \equiv 1$, and $0! \equiv 1$. We get now, instead of (2.3), the interpolation formula

$$\hat{s}(m + \tau) = \sum_{i=0}^n L_i(\tau) s(m + i), \quad \text{for } 0 \leq \tau \leq n , \quad (2.6)$$

with the interpolation error

$$R(\tau, n) \equiv R^*(m + \tau, n) = \frac{d^{n+1}}{dt^{n+1}} s(t) \Big|_{\vartheta(t)} \frac{1}{(n+1)!} \prod_{i=0}^n (\tau - i) , \quad (2.7)$$

where $m \leq \vartheta(t) \leq m + n$.

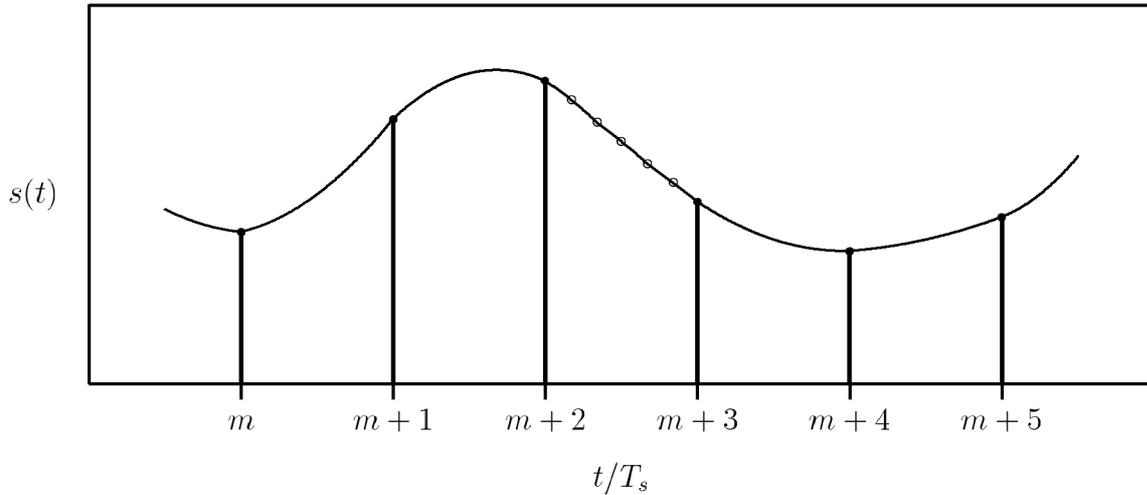
To get a more convenient expression, we consider in the following only odd values of n , $n = 3, 5, 7, \dots$, and we restrict the interpolation to equidistant values of τ in the interval $]\frac{n-1}{2}, \frac{n+1}{2}[$, i.e. $\tau \in \{\tau_j\}$,

$$\tau_j = \frac{n-1}{2} + \frac{j}{k}, \quad j = 1, 2, \dots, k-1 , \quad (2.8)$$

where $k \geq 2$. $k-1$ represents the number of points we want to obtain by interpolation from the samples $s(m), s(m+1), \dots, s(m+n)$. Figure 2.1 illustrates this situation.

The restriction to the middle interval has computational advantages, because the interpolation error is expected to be small there. Further, from (2.5) we can derive for each time instant τ_j the coefficients

$$L_{ij} \equiv L_i(\tau_j), \quad \text{where } i = 0, 1, \dots, n \quad \text{and} \quad j = 1, 2, \dots, k-1 . \quad (2.9)$$

Bild 2.1: Scheme of the interpolation of order $n = 5$, and $k - 1 = 5$ interpolated values

From (2.6) we see that they represent for each j a set of $n + 1$ coefficients of a (noncausal) digital filter (finite impuls response: FIR) from which we can obtain the interpolated values,

$$\hat{s}(m + \tau_j) = \sum_{i=0}^n L_{ij} s(m + i) . \quad (2.10)$$

Note that the coefficients L_{ij} do not depend on the time instant m . Hence they have to be calculated only once, and if we vary m from $1 \cdots M - n - 1$, we obtain from the discrete series (2.1) via the filter (2.10) for each $j = 1, 2, \dots, k - 1$ a series of interpolated values

$$\{\hat{s}(m + \tau_j)\}_{m=1}^{M-n-1} . \quad (2.11)$$

Together with the original series (2.1) they can be used to compose a times series

$$\left\{ \hat{s} \left(\frac{n-1}{2} + \frac{j}{k} \right) \right\}_{j=0}^{(M-n+1)k} ,$$

Neglecting the interpolation error, this can be considered as a sampled version of our signal $s(t)$, but now with a sampling frequency kf_s .

Bounds of the Interpolation Error

We intend now to find an upper bound of the magnitude of the interpolation error (2.7). We start with the product on the right hand side. For the considered case $\tau = \frac{n-1}{2} + \delta$, where $\delta \in [0, 1]$, $n = 3, 5, 7, \dots$, we obtain

$$\begin{aligned} \left| \prod_{i=0}^n (\tau - i) \right| &= \left| \prod_{i=0}^{(n-1)/2} (\delta + i)(1 - \delta + i) \right| \\ &\leq \prod_{i=0}^{(n-1)/2} \left(\frac{1}{2} + i \right)^2 \end{aligned} \quad (2.12)$$

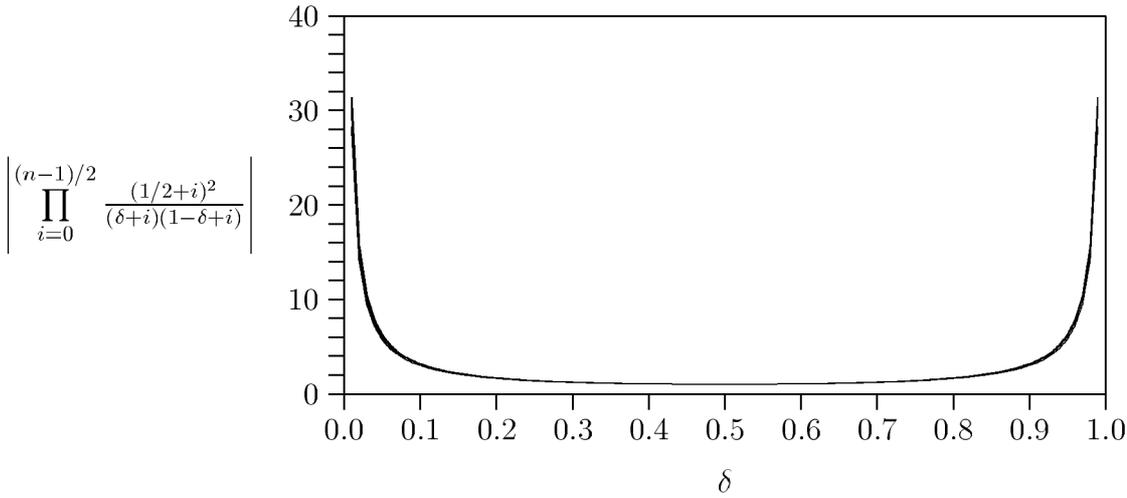


Bild 2.2: Ratio of overestimation according to Equ. (2.12) ($n = 3, 7, 11, \dots, 31$)

Note that this upper bound is independent on the instant δ . The equality holds for $\delta = 1/2$. The usefulness of this upper bound is illustrated in Figure 2.2.

Obviously only for $\delta \gtrsim 0$ and $\delta \lesssim 1$ it drastically overestimates the product on the left hand side of (2.12). Moreover, the ratio of overestimation is almost independent on the interpolation order n .

An upper bound of the derivative $|\mathrm{d}^{n+1}s(t)/\mathrm{d}t^{n+1}|$ in (2.7) can be obtained in dependence of the Nyquist frequency $f_N \equiv 2\pi/\omega_N$ of the signal $s(t)$ as follows: Consider the Fourier representation of the signal, which is presumed to be band-limited ($\omega_N < \infty$),

$$s(t) = \frac{1}{2\pi} \int_{-\omega_N}^{\omega_N} S(\omega) e^{j\omega t} \mathrm{d}\omega, \quad (2.13)$$

supposing that the integral exists. The spectral density $S(\omega)$ is obtained from

$$S(\omega) = \int_{-\infty}^{\infty} s(t) e^{-j\omega t} \mathrm{d}t, \quad (2.14)$$

with the imaginary unit $j \equiv \sqrt{-1}$. Now the $(n+1)$ th derivate of s can be written as

$$\left| \frac{\mathrm{d}^{n+1}s(t)}{\mathrm{d}t^{n+1}} \right| = \frac{1}{2\pi} \left| \int_{-\omega_N}^{\omega_N} (j\omega)^{n+1} S(\omega) e^{j\omega t} \mathrm{d}\omega \right|. \quad (2.15)$$

From the Schwarzian inequality

$$\left| \int f(\omega) \overline{g(\omega)} \mathrm{d}\omega \right|^2 \leq \int |f(\omega)|^2 \mathrm{d}\omega \cdot \int |g(\omega)|^2 \mathrm{d}\omega$$

we obtain

$$\begin{aligned} \left| \frac{\mathrm{d}^{n+1}s(t)}{\mathrm{d}t^{n+1}} \right| &\leq \frac{1}{2\pi} \sqrt{\int_{-\omega_N}^{\omega_N} \omega^{2(n+1)} \mathrm{d}\omega \cdot \int_{-\omega_N}^{\omega_N} |S(\omega)|^2 \mathrm{d}\omega} = \dots \\ &\dots = \frac{1}{2\pi} \sqrt{\frac{2\omega_N^{2n+3}}{2n+3}} \sqrt{\int_{-\omega_N}^{\omega_N} |S(\omega)|^2 \mathrm{d}\omega} \end{aligned}$$

Note that $|S(\omega)|^2$ represents the power spectral density of the signal $s(t)$, and hence

$$E_s \equiv \sqrt{\frac{1}{2\omega_N} \int_{-\omega_N}^{\omega_N} |S(\omega)|^2 d\omega} = \sqrt{\frac{1}{2f_N} \int_{-\infty}^{\infty} s^2(t) dt} \quad (2.16)$$

is the square root of the mean power spectral density. The equality (2.16) follows from the well-known fact that the power spectral density is the Fourier transform of the autocorrelation function (Wiener Khinchine theorem), and hence

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} |S(\omega)|^2 e^{j\omega\tau} d\omega \Big|_{\tau=0} = \int_{-\infty}^{\infty} s(t)s(t+\tau) dt \Big|_{\tau=0}.$$

Multiplication of both sides with $\frac{1}{2f_N}$ and taking the square root yields (2.16).

In the following E_s is assumed to be finite and not zero. With this abbreviation we can write

$$\left| \frac{d^{n+1}s(t)}{dt^{n+1}} \right| \leq \frac{\omega_N^{n+2}}{\pi\sqrt{2n+3}} E_s, \quad (2.17)$$

Summarizing (2.12) and (2.17) we obtain for the magnitude of the relative interpolation error (2.7)

$$\frac{|R(\tau, n)|}{E_s} \leq \frac{(2\pi f_N)^{n+2}}{\pi\sqrt{2n+3}} \frac{1}{(n+1)!} \prod_{i=0}^{(n-1)/2} \left(\frac{1}{2} + i\right)^2 \equiv r(f_N, n). \quad (2.18)$$

As already mentioned above, we measure time t in units of the sampling period T_s . Consequently the Nyquist frequency $f_N = \omega_N/(2\pi)$ is measured in units of the sampling frequency $f_s > 2f_N$, and the assumptions of the sampling theorem are satisfied if f_N , in Equ. (2.18), runs from $0 \cdots 1/2$. However, to guarantee that $r(f_N, n)$ diminishes with an increase of the interpolation order n , the relation $f_N < \pi^{-1}$ has to be satisfied. We conclude this from the formula

$$r(f_N, n+2m) = r(f_N, n) \sqrt{\frac{2n+3}{2n+4m+3}} \left(\prod_{l=1}^m \frac{1}{n+2l+1} \right) (\pi f_N)^{2m},$$

which is dominated by the exponential term $(\pi f_N)^{2m}$ if $m \rightarrow \infty$ ($n = 3, 5, 7, \dots$ and $m = 1, 2, 3, \dots$). For $\pi f_N > 1$ the bound $r(f_N, n+2m)$ diverges with increasing m . Moreover, from (2.18) we conclude that if $r(f_N, n)$ goes to zero with growing n then the magnitude of the interpolation error $|R(\tau, n)|$ goes to zero as well. The results are summarized in the following

Theorem:

Let $\{s(m)\}_{m=-\infty}^{\infty}$ be equally spaced samples of the signal $s(t)$, where the time is measured in units of the sampling period $T_s = f_s^{-1}$, and hence frequencies in units of f_s . The signal is presumed to be band-limited, i.e.,

$$\begin{aligned} S(\omega) &\equiv \int_{-\infty}^{\infty} s(t) e^{-j\omega t} dt \\ &= 0 \quad \text{for } |\omega| > \omega_N, \end{aligned}$$

with the Nyquist frequency

$$f_N \equiv \frac{\omega_N}{2\pi} < \frac{1}{\pi}. \quad (2.19)$$

Moreover, s is assumed to have finite energy,

$$E_s \equiv \sqrt{\frac{1}{2\omega_N} \int_{-\omega_N}^{\omega_N} |S(\omega)|^2 d\omega} < \infty. \quad (2.20)$$

Then we can reconstruct $s(t)$ from the samples according to the interpolation formula

$$s\left(m + \frac{n-1}{2} + \frac{j}{k}\right) = \sum_{i=0}^n L_{ij} s(m+i) + R(m+j/k, n) \quad (2.21)$$

where $n = 3, 5, 7, \dots$, $k = 2, 3, 4, \dots$, and $j = 1, 2, 3, k-1$. The coefficients are given by

$$\begin{aligned} L_{ij} &= \frac{(-1)^{n-i}}{i! (n-i)!} \prod_{l=0, l \neq i}^n \left(\frac{n-1}{2} + \frac{j}{k} - l \right) \\ &= \prod_{l=0, l \neq i}^n \left(\frac{n-1}{2} + \frac{j}{k} - l \right) \frac{1}{i-l}. \end{aligned} \quad (2.22)$$

The magnitude of the interpolation error is bounded from above,

$$|R(m+j/k, n)| \leq r(f_N, n) E_s \quad (2.23)$$

where

$$r(f_N, n) \equiv \frac{(2\pi f_N)^{n+2}}{\pi \sqrt{2n+3}} \frac{1}{(n+1)!} \prod_{i=0}^{(n-1)/2} \left(\frac{1}{2} + i \right)^2 \quad (2.24)$$

The interpolation error vanishes if the interpolation order n is increased,

$$\lim_{n \rightarrow \infty} |R(m+j/k, n)| = 0.$$

Remarks:

The condition (2.19) says that we need a slight oversampling of the signal to guarantee that the interpolation error can be diminished by increasing the interpolation order n . According to the sampling theorem $f_N < 1/2$ would be sufficient, in general, but not in our approach. However, $1/\pi$ is close to $1/2$, and hence (2.19) is, from a practical point of view, no essential restriction.

Further it should be noted that the error bound in (2.23) approaches the faster to zero the smaller f_N as it is illustrated in Figure 2.3.

The curves reflect the exponential decay of the interpolation error when the interpolation order n increases. Consider, e.g., the curve for $\pi f_N = 0.7$ together with the horizontal line at $r(f_N, n) = 10^{-3}$. They tell us that if we have sampled the signal with the frequency $f_s \geq \frac{\pi f_N}{0.7}$ then the magnitude $r(f_N, n)$ of the relative interpolation error is less than 10^{-3} if we choose the interpolation order $n \geq 17$.

Example 1:

Consider the *sinus cardinalis*

$$s(t) = A \cdot \frac{\sin \omega_N t}{\omega_N t},$$

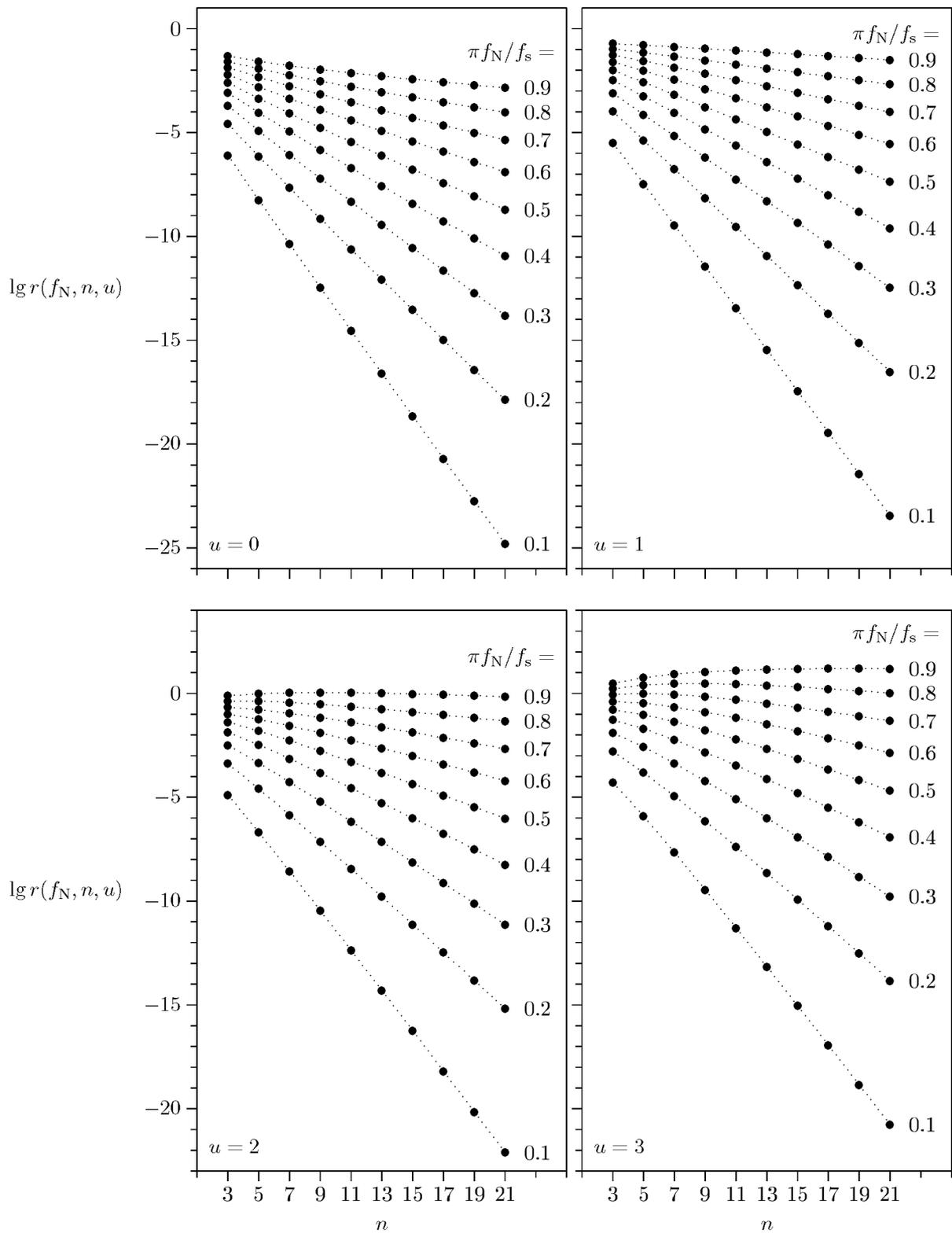


Bild 2.3: Decay of the magnitude $r(f_N, n)$ of the relative interpolation error with increasing interpolation order n . Parameter: Relative Nyquist frequency $\pi f_N / f_s$ and differentiation order u

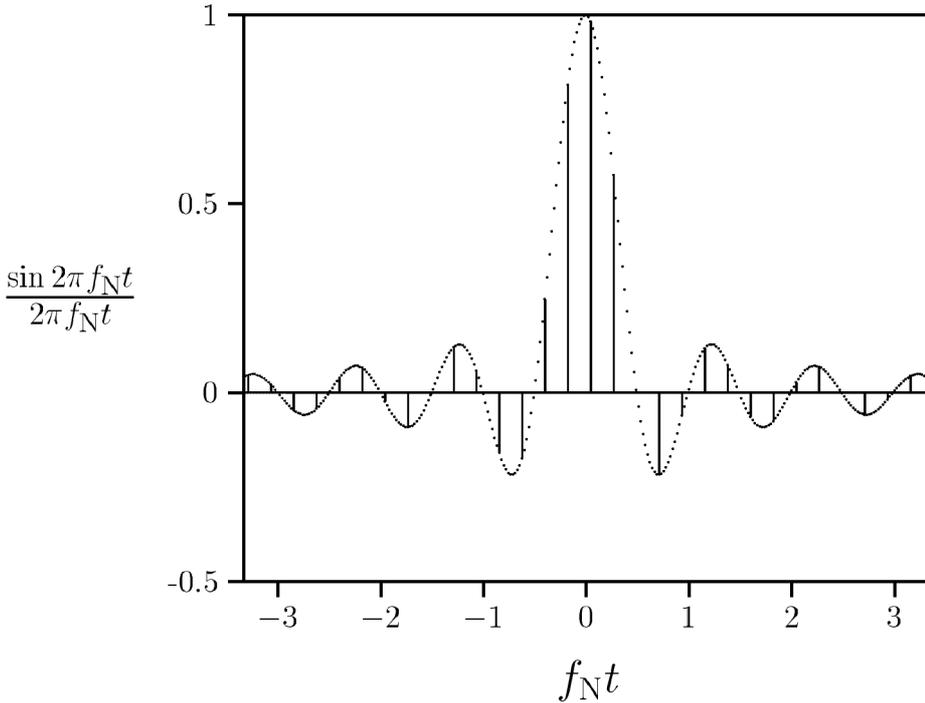


Bild 2.4: Interpolated data of the signal of example 1 ($k = 10$, $n = 13$, $f_s = \frac{\pi}{0.7} f_N$). The vertical lines indicate the original samples. The magnitude of the interpolation error is less than 10^{-3} .

which is sampled at the rate $f_s = \frac{\omega_N}{1.4} = \frac{\pi}{0.7} f_N \approx 4.49 \cdot f_N$. Take, e.g., $\{s(\pi^{-1} + m/f_s)\}_m$. We want to interpolate with an error magnitude less than $0.001 \cdot A$. Then we can proceed according to our theorem as follows:

The power spectral density of s is given by

$$|S(\omega)| = \begin{cases} A\pi/\omega_N & : |\omega| < \omega_N \\ 0 & : \text{otherwise.} \end{cases}$$

Hence, from (2.20) we obtain $E_s = \frac{A}{2f_N} < \infty$, and thus, from (2.23), we conclude that $r(f_N, n) \cdot \frac{A}{2f_N} < 0.001 \cdot A$, or equivalently $\log_{10} r(\frac{0.7}{\pi}, n) < -3 - \log_{10} \frac{\pi}{1.4} \approx -3.35$, is sufficient to have the desired interpolation accuracy. From Figure 2.3 (see curve with $\pi f_N/f_s = 0.7$) we see that this can be performed with an interpolation order $n \geq 13$. If the signal was sampled with a larger frequency, say $f_s = \frac{\pi f_N}{0.4} \approx 7.85 \cdot f_N$, then $n \geq 5$ is sufficient. Figure 2.4 illustrates the result working with the interpolation order $n = 13$. The corresponding actual interpolation error and the upper bounds are illustrated in Figure 2.5.

Interpolation in the presence of noise

In a typical practical situation the samples (2.1) are corrupted by noise. It has its origin in the analog-to-digital conversion of the data ("quantization error") or by additional random sources. Consequently we have to consider, instead of (2.1), the actually attainable sequence

$$\{s^*(m)\}_{m=1}^M, \tag{2.25}$$

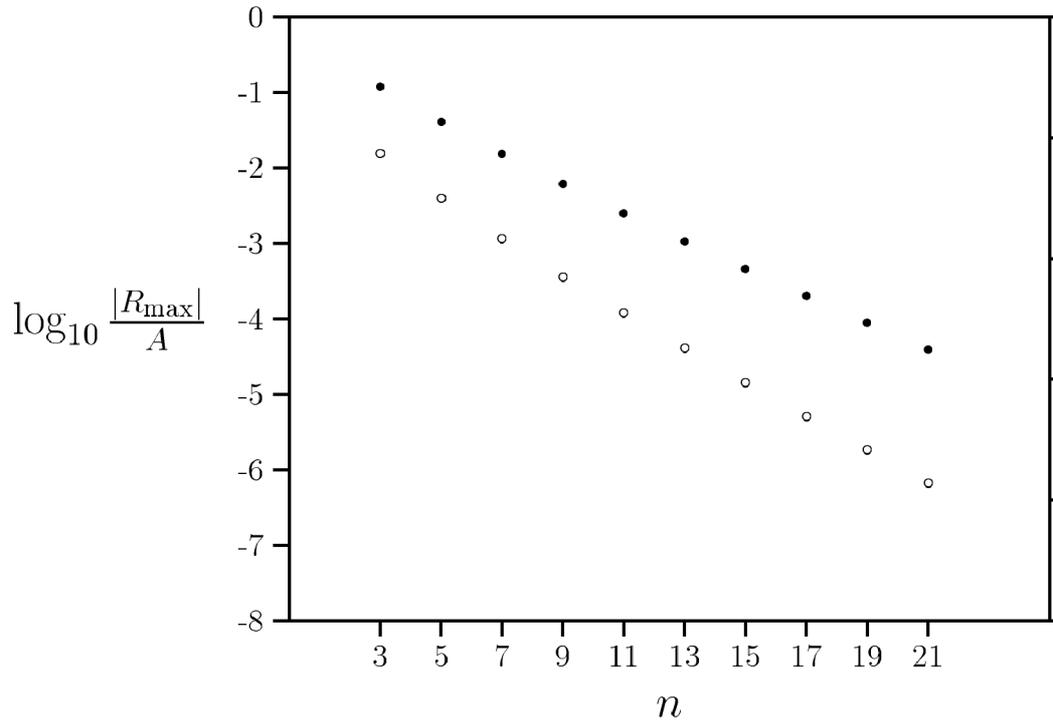


Bild 2.5: Maximum R_{\max} of the actual interpolation error (open circles) and its estimated upper bound (filled circles) for the function $s(t) = A \frac{\sin \omega_N t}{\omega_N t}$ of example 1

where $s^*(m)$ differs from $s(m)$ by the error

$$d(m) \equiv s^*(m) - s(m). \quad (2.26)$$

(Again, time m is measured in units of the sampling period T_s which is assumed to be constant.) Consequently, in formula (2.10) we have to work with $s^*(m)$ instead of $s(m)$, leading to

$$\begin{aligned} \hat{s}^*(m + \tau_j) &= \sum_{i=0}^n L_{ij} s^*(m + i) \\ &= \sum_{i=0}^n L_{ij} [s(m + i) + d(m + i)] \\ &= \hat{s}(m + \tau_j) + e(m + \tau_j). \end{aligned} \quad (2.27)$$

$\hat{s}(m + \tau_j)$ is the interpolated value (2.10) in the noiseless case and

$$e(m + \tau_j) \equiv \sum_{i=0}^n L_{ij} d(m + i) \quad (2.28)$$

denotes the interpolation error of $\hat{s}^*(m + \tau_j)$ due to noise. (τ_j is defined in (2.8).)

Suppose now that the noise amplitudes $\{|d(m)|\}$ are limited by d_{\max} , i.e.

$$d_{\max} \equiv \max_{m=1, \dots, M} \{|d(m)|\}. \quad (2.29)$$

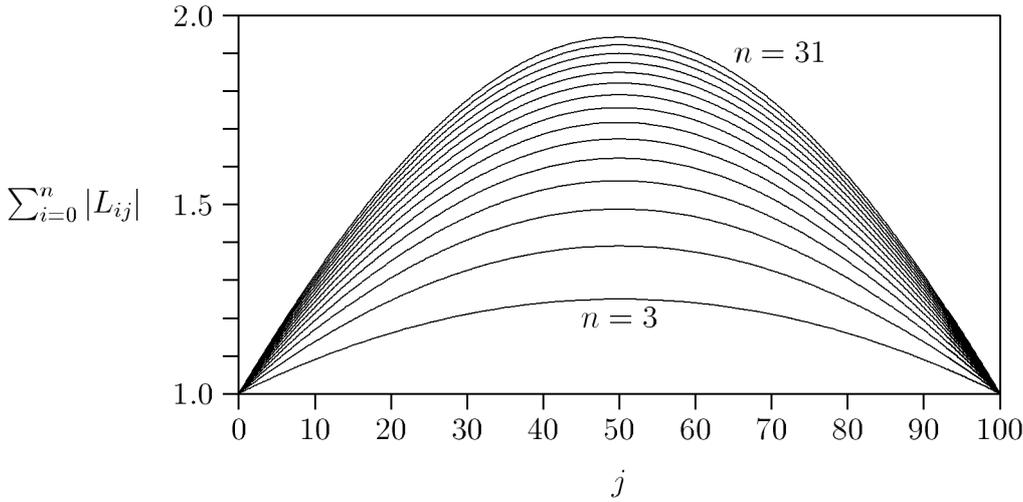


Bild 2.6: Variation of the sum of the magnitude of the interpolation coefficients for $\tau_j = \frac{n-1}{2} + \frac{j}{k}$, $k = 100$. Parameter : interpolation order n

Obviously this is a useful assumption if d contains only quantization noise. In this case d_{\max} is given by the smallest quantization step. Then the magnitude of the interpolation error (2.28) is bounded by

$$|e(m + \tau_j)| \leq d_{\max} \sum_{i=0}^n |L_{ij}|. \quad (2.30)$$

The sum over the magnitude of the interpolation coefficients L_{ij} are limited as Figure 2.6 illustrates. For each interpolation order n it reaches its maximum value for $j/k = 1/2$. Notably, for $n \leq 31$ it does not exceed the value 2, which means that the interpolation in the presence of noise satisfies

$$|e(m + \tau_j)| \leq 2 d_{\max}, \quad \text{for } n \leq 31. \quad (2.31)$$

The practical consequence of this consideration is as follows: The proposed interpolation algorithm yields a quantity $\hat{s}^*(t)$ that differs from the actual signal value $s(t)$ by two kinds of errors — that which is inherent in the algorithm, and that which follows from corruption by noise. Bounds for the first are given by (2.23), and for the latter by (2.30) resp. (2.31). Thus we can estimate the magnitude of the relative total interpolation error by

$$\left| \frac{\hat{s}^*(t) - s(t)}{E_s} \right| \leq r(f_N, n) + \frac{2 d_{\max}}{E_s}, \quad \text{for } n \leq 31. \quad (2.32)$$

As we have shown above, $r(f_N, n)$ could be arbitrarily diminished if we increase the interpolation order n (see Figure 2.3). However, an increase of n can essentially improve the performance of the algorithm only until $r(f_N, n) \lesssim 2d_{\max}/E_s$, otherwise in (2.32) the term $2d_{\max}/E_s$ would dominate, which is, for large n , almost independent on n (see Figure 2.6). Thus we would, in practice, choose n such that

$$r(f_N, n) \approx 10 \cdot \frac{2 d_{\max}}{E_s}, \quad (2.33)$$

where the somewhat arbitrarily chosen factor 10 on the right hand side guarantees that, from a practical point of view, the performance cannot be improved by further increase of the interpolation order n . Otherwise we would only increase the complexity of our algorithm.

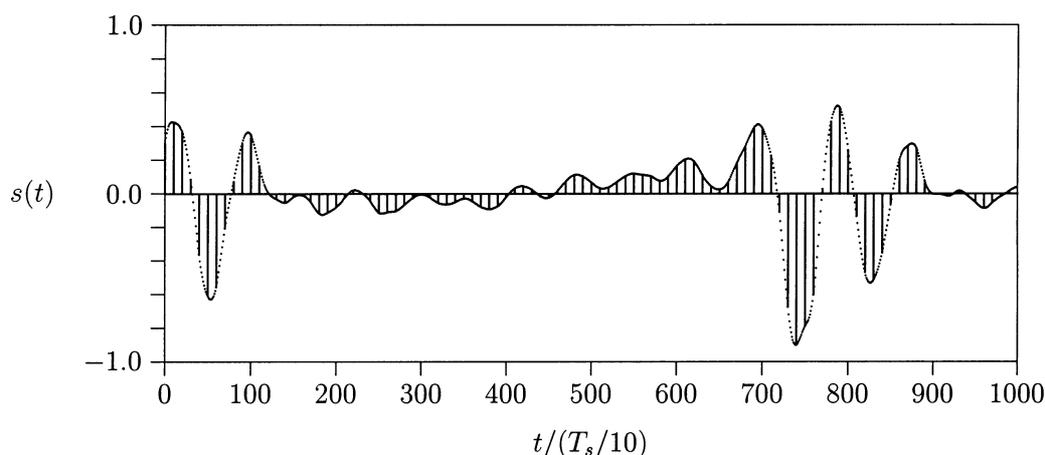


Bild 2.7: Interpolated speech signal (interpolation factor $k = 10$, interpolation order $n = 13$)

Example 2: Speech Signal

As an example Figure 2.7 shows the results for a speech signal which is an isolated spoken allophone [a:] of a male speaker in a modal register. The original samples are marked by vertical lines. They were obtained from an 8-bit analog/digital converter at a sampling period of $T_s = 125 \mu s$. The interpolated values are represented by points.

2.3 Differentiation

We use now the interpolation technique of the former section to derive formulas for interpolation combined with differentiation of the signal $s(t)$, provided only the samples (2.1) are given. From (2.4) and (2.6) we obtain for the u -th derivative

$$\frac{d^u s(m + \tau)}{d\tau^u} = \sum_{i=0}^n \frac{d^u L_i(\tau)}{d\tau^u} s(m + i) + \frac{d^u R(\tau, n)}{d\tau^u}, \quad \text{for } 0 \leq \tau \leq n. \quad (2.34)$$

The first derivative of the interpolation coefficients (2.5) are obtained as

$$\begin{aligned} L_i^{(1)}(\tau) &\equiv \frac{d}{d\tau} L_i(\tau) = \frac{d}{d\tau} \exp(\ln L_i(\tau)) \\ &= L_i(\tau) \cdot \frac{d}{d\tau} \ln L_i(\tau) \\ &= L_i(\tau) \cdot \frac{d}{d\tau} \left[\ln \left(\frac{(-1)^{n-i}}{i! (n-i)!} \right) + \sum_{l=0, l \neq i}^n \ln(\tau - l) \right] \\ &= L_i(\tau) \cdot \sum_{l=0, l \neq i}^n (\tau - l)^{-1}. \end{aligned} \quad (2.35)$$

The u -th derivatives of the interpolation coefficients (2.5) are obtained for any $u > 0$ applying LEIBNIZ formula¹⁾

$$\begin{aligned}
 L_i^{(u)}(\tau) &\equiv \frac{d^u L_i(\tau)}{d\tau^u} = \frac{d^{u-1}}{d\tau^{u-1}} L_i^{(1)}(\tau) \\
 &= \frac{d^{u-1}}{d\tau^{u-1}} \left(L_i(\tau) \cdot \sum_{l=0, l \neq i}^n (\tau - l)^{-1} \right) \\
 &= \sum_{m=0}^{u-1} \binom{u-1}{m} L_i^{(m)}(\tau) \sum_{l=0, l \neq i}^n \frac{d^{u-1-m}}{d\tau^{u-1-m}} \frac{1}{\tau - l} \\
 &= \sum_{m=0}^{u-1} \left[\frac{(u-1)!}{(u-1-m)! m!} L_i^{(m)}(\tau) \sum_{l=0, l \neq i}^n (u-1-m)! (-1)^{u-1-m} (\tau - l)^{m-u} \right]
 \end{aligned}$$

Herein we have set $L_i^{(0)}(\tau) \equiv L_i(\tau)$. Thus, we get the u -th differentiate of the interpolation coefficients,

$$L_i^{(u)}(\tau) = \sum_{m=0}^{u-1} \left[L_i^{(m)}(\tau) \frac{(u-1)! (-1)^{u-1-m}}{m!} \sum_{l=0, l \neq i}^n (\tau - l)^{m-u} \right] \quad (2.36)$$

The second derivatives, e.g., are obtained from

$$L_i^{(2)}(\tau) = -L_i(\tau) \sum_{l=0, l \neq i}^n \frac{1}{(\tau - l)^2} + L_i^{(1)}(\tau) \sum_{l=0, l \neq i}^n \frac{1}{\tau - l}, \quad (2.37)$$

and the third derivatives are given by

$$L_i^{(3)}(\tau) = 2 L_i(\tau) \sum_{l=0, l \neq i}^n \frac{1}{(\tau - l)^3} - 2 L_i^{(1)}(\tau) \sum_{l=0, l \neq i}^n \frac{1}{(\tau - l)^2} + L_i^{(2)}(\tau) \sum_{l=0, l \neq i}^n \frac{1}{\tau - l}. \quad (2.38)$$

For computational reasons we rewrite (2.36) as follows,

$$L_i^{(u)}(\tau) = \sum_{m=0}^{u-1} a_{m,i}^{(u-1)}(\tau) L_i^{(m)}(\tau), \quad \text{where} \quad a_{m,i}^{(u)}(\tau) \equiv \frac{u! (-1)^{u-m}}{m!} \sum_{l=0, l \neq i}^n (\tau - l)^{m-u-1}$$

Thus we can recursively calculate the filter coefficients as (omitting indices i and τ for convenience),

$$\begin{aligned}
 L^{(1)} &= a_0^{(0)} L^{(0)} &= & a_0^{(0)} \cdot L^{(0)} \\
 L^{(2)} &= a_0^{(1)} L^{(0)} + a_1^{(1)} L^{(1)} &= & \left(a_0^{(1)} + a_1^{(1)} a_0^{(0)} \right) \cdot L^{(0)} \\
 L^{(3)} &= a_0^{(2)} L^{(0)} + a_1^{(2)} L^{(1)} + a_2^{(2)} L^{(2)} &= & \left[a_0^{(2)} + a_1^{(2)} a_0^{(0)} + a_2^{(2)} \left(a_0^{(1)} + a_1^{(1)} a_0^{(0)} \right) \right] \cdot L^{(0)}
 \end{aligned}$$

¹⁾For any functions $x(t)$ and $y(t)$ we have $\frac{d^n(xy)}{dt^n} = \sum_{m=0}^n \binom{n}{m} \frac{d^m x}{dt^m} \cdot \frac{d^{n-m} y}{dt^{n-m}}$.

With the abbreviations

$$\begin{aligned}
 b^{(0)} &\equiv 1 \\
 b^{(1)} &\equiv a_0^{(0)} = a_0^{(0)} b_0 \\
 b^{(2)} &\equiv a_0^{(1)} + a_1^{(1)} a_0^{(0)} = a_0^{(1)} b^{(0)} + a_1^{(1)} b^{(1)} \\
 b^{(3)} &\equiv a_0^{(2)} + a_1^{(2)} a_0^{(0)} + a_2^{(2)} \left(a_0^{(1)} + a_1^{(1)} a_0^{(0)} \right) = a_0^{(2)} b^{(0)} + a_1^{(2)} b^{(1)} + a_2^{(2)} b^{(2)} \\
 &\dots \\
 b_i^{(u)}(\tau) &\equiv \sum_{m=0}^{u-1} a_{m,i}^{(u-1)}(\tau) b_i^{(m)}(\tau)
 \end{aligned}$$

we get for each $u \geq 0$

$$L_i^{(u)}(\tau) = L_i^{(0)}(\tau) b_i^{(u)}(\tau) .$$

The differentiated signal (2.34) is obtained in terms of the original time units of $t = \tau T_s$ by

$$\frac{d^u s((m + \tau)T_s)}{dt^u} = \frac{d^u s(m + \tau)}{d\tau^u} \cdot \frac{1}{(2\pi f_N / f_s)^u} .$$

Moreover, for time instants τ_j , defined in (2.8), we write

$$L_{ij}^{(u)} \equiv \frac{d^u L_i(\tau_j)}{d\tau^u} . \quad (2.39)$$

The magnitude of the relative interpolation error can be estimated, using the relations (2.7) and (2.12),

$$\left| \frac{d^u R(\tau, n)}{d\tau^u} \right| \leq (n + 1)^u |R(\tau, n)| .$$

Thus, from (2.18) we finally get for the magnitude of the relative differentiation–interpolation error

$$\frac{1}{E_s} \left| \frac{d^u R(\tau, n)}{d\tau^u} \right| \leq (n + 1)^u r(f_N, n) . \quad (2.40)$$

Again, the right hand term of (2.40) diverges for $n \rightarrow \infty$ if the NYQUIST–frequency, which is still measured in units of the sampling frequency f_s , satisfies (2.19).

A source code (PASCAL) for the interpolation procedure is given in the Appendix C (p. 133).

Differentiation in the presence of noise

We consider again the situation where the samples are corrupted by noise (see formulas (2.25) to (2.31)). However, instead of the interpolation coefficients L_{ij} we now have to take the differentiation–interpolation coefficients (2.35). Then, similarly to (2.28), the magnitude of the differentiation–interpolation error

$$e^{(u)}(m + \tau_j) \equiv \sum_{i=0}^n L_{ij}^{(u)} d(m + i) \quad (2.41)$$

can be estimated by

$$e^{(u)}(m + \tau_j) \leq d_{\max} \sum_{i=0}^n |L_{ij}^{(u)}| \quad (2.42)$$

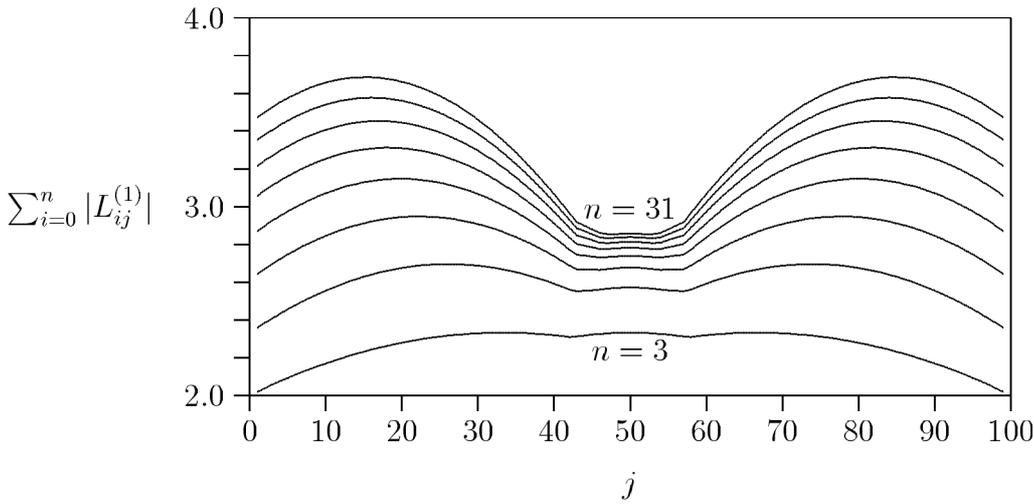


Bild 2.8: Variation of the sum of the magnitude of the interpolation/differentiation coefficients for $\tau_j = \frac{n-1}{2} + \frac{j}{k}$, $k = 100$. Parameter : interpolation order $n = 3, 7, 11, \dots, 31$

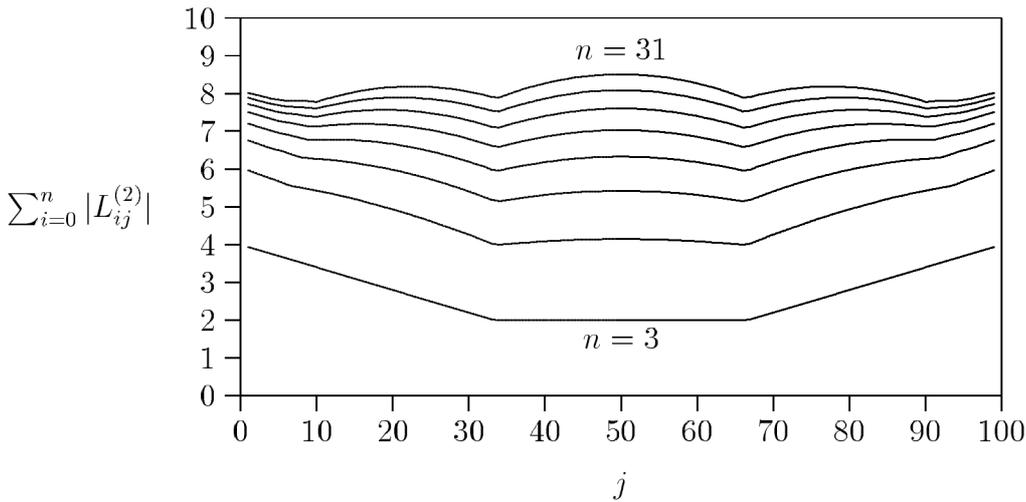


Bild 2.9: Variation of the sum of the magnitude of the interpolation/differentiation coefficients for $\tau_j = \frac{n-1}{2} + \frac{j}{k}$, $k = 100$. Parameter : interpolation order $n = 3, 7, 11, \dots, 31$

The sum over the magnitude of the interpolation coefficients $L_{ij}^{(u)}$ is shown in Figures 2.8 and 2.9 for $u = 1$ and 2 , respectively.

Noteably, for $n \leq 31$ it does not exceed the value 4 for $u = 1$ and 9 for $u = 2$.

Example 3: Speech Signal

As an example Figure 2.10 shows the first derivative of the speech signal of Example 2. The vertical lines indicate the original samples.

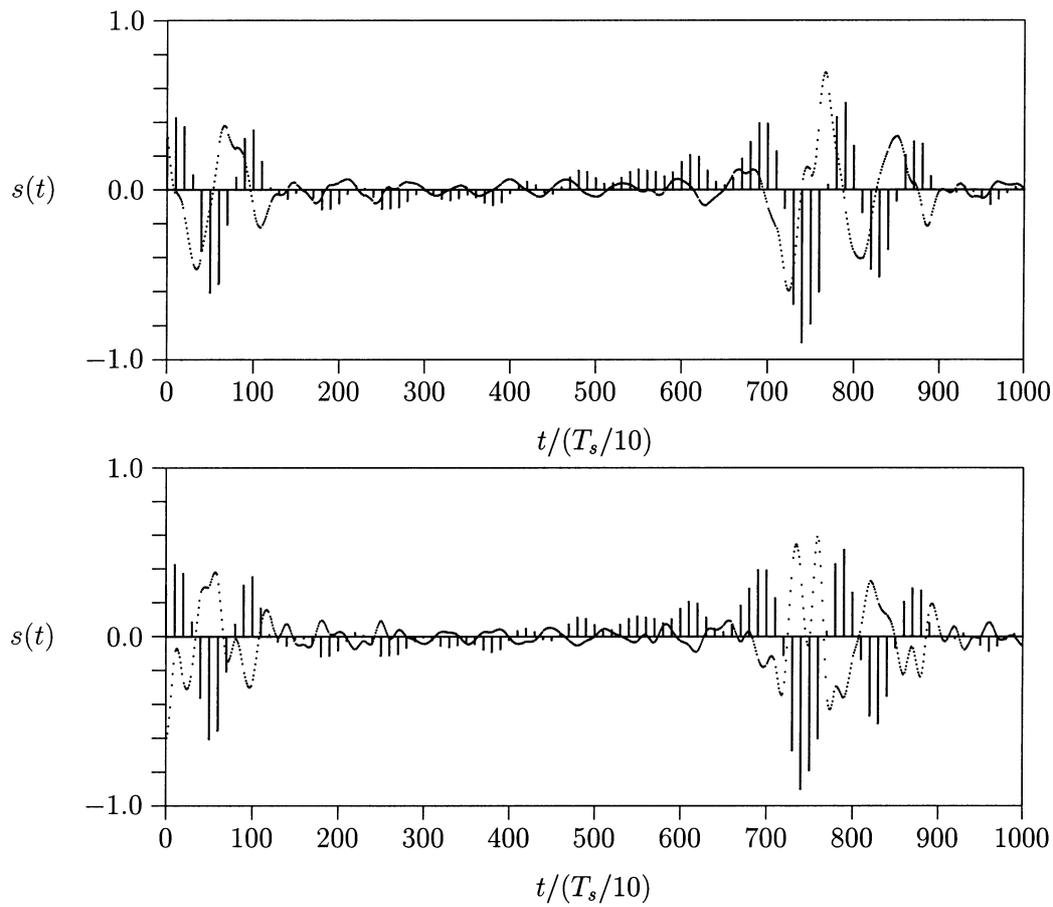


Bild 2.10: Interpolated and one (top) and two (bottom) times differentiated speech signal (interpolation factor $k = 10$, interpolation order $n = 13$)

Kapitel 3

Korrelation

3.1 Definition und graphische Interpretation

Der Korrelationskoeffizient zweier Zeitreihen

$$\{x_t\}_{t=1}^T \quad \text{und} \quad \{y_t\}_{t=1}^T \quad (3.1)$$

ist ein Maß für ihren *linearen* statistischen Zusammenhang. Wir verfolgen hier zunächst eine anschauliche geometrische Interpretation des Korrelationskoeffizienten. Dazu stellen wir die Punkte $\{(x_t, y_t)\}_{t=1}^T$ in der Ebene dar.¹⁾ Ein Beispiel zeigt Bild 3.1 für den Fall, dass $\{y_t\}$ aus einer zeitlichen Verschiebung von $\{x_t\}$ um τ entsteht, $y_t = x_{t+\tau}$.

Wir wollen nun durch die Punktwolke eine Regressionsgerade

$$y_t^*(a, b) = a + bx_t \quad , \quad (3.2)$$

legen, so dass

$$f(a, b) \equiv \frac{1}{T} \sum_{t=1}^T (y_t^*(a, b) - y_t)^2 = \min ! \quad (3.3)$$

Aus dieser Bedingung der kleinsten mittleren quadratischen Abweichung können die unbekannt-ten Parameter a und b bestimmt werden. Im Englischen spricht man vom *least mean square fit* (LMS fit), was wir im Folgenden genauer ausführen:

Wir setzen (3.2) in (3.3) ein und erhalten

$$f(a, b) = \frac{1}{T} \sum_{t=1}^T (a + bx_t - y_t)^2$$

Um das Minimum von f zu finden, setzen wir die partiellen Ableitungen Null,

$$\frac{\partial f}{\partial a} = \frac{2}{T} \sum_{t=1}^T a + bx_t - y_t = 0 \quad (3.4)$$

$$\frac{\partial f}{\partial b} = \frac{2}{T} \sum_{t=1}^T x_t(a + bx_t - y_t) = 0 \quad (3.5)$$

¹⁾Diese Punktwolke heisst im Englischen *scatter plot* oder auch *delay plot*.

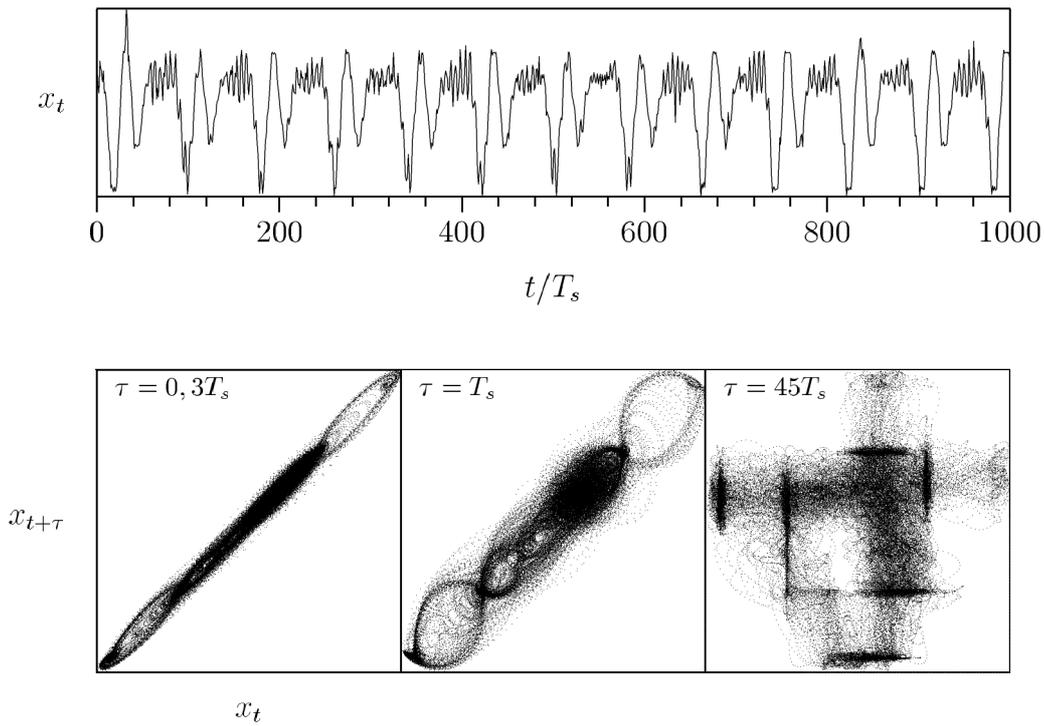


Bild 3.1: Oben: Zeitreihe zum isoliert gesprochenen Laut /dsch:/ wie in „joule“ (Abtastperiode: $T_s = 125\mu s$, Amplitudenauflösung: 8 bit, linear interpoliert). Unten: Verzögerungsdarstellungen zur obigen Zeitreihe (Gesamtzeit 1s, 10-fach hochgetastet durch Lagrange-Interpolation)

Mit den Abkürzungen für die Mittelwerte,

$$\overline{x_t} \equiv \bar{x} \equiv \frac{1}{T} \sum_{t=1}^T x_t \quad \text{und} \quad \overline{y_t} \equiv \bar{y} \equiv \frac{1}{T} \sum_{t=1}^T y_t$$

erhalten wir aus (3.4),

$$a + b\bar{x} = \bar{y} . \quad (3.6)$$

Weiterhin kann Gleichung (3.5) mit den Abkürzungen

$$\overline{x_t^2} \equiv \overline{x^2} \equiv \frac{1}{T} \sum_{t=1}^T x_t^2 \quad \text{und} \quad \overline{x_t y_t} \equiv \overline{xy} \equiv \frac{1}{T} \sum_{t=1}^T x_t y_t$$

wie folgt geschrieben werden,

$$a\bar{x} + b\overline{x^2} = \overline{xy} . \quad (3.7)$$

Aus (3.6) und (3.7) folgt für die gesuchten Parameter

$$a = \bar{y} - b\bar{x} \quad \text{und} \quad b = \frac{\overline{xy} - \bar{x}\bar{y}}{\overline{x^2} - \bar{x}^2} . \quad (3.8)$$

Um festzustellen, ob an diesen Stellen $f(a, b)$ ein Minimum hat, bilden wir die zweiten Ableitungen,

$$D \equiv \begin{pmatrix} \frac{\partial^2 f}{\partial a^2} & \frac{\partial^2 f}{\partial b \partial a} \\ \frac{\partial^2 f}{\partial a \partial b} & \frac{\partial^2 f}{\partial b^2} \end{pmatrix} = 2 \begin{pmatrix} 1 & \bar{x} \\ \bar{x} & \overline{x^2} \end{pmatrix}$$

Ist die Zeitreihe nicht entartet, d.h., sind nicht alle $x_t = \text{const.}$, so gilt

$$\det D/2 = \overline{x^2} - \bar{x}^2 = \overline{(x - \bar{x})^2} > 0 .$$

Damit hat f ein relatives Extremum an der durch (3.8) gegebenen Stelle. Aus $\frac{\partial^2 f}{\partial a^2} = 2 > 0$ folgt, dass dies ein relatives Minimum ist.

Um zu Ausdrücken zu kommen, die besser interpretierbar sind, führen wir die *Varianz (Streuung)* einer Zeitreihe ein,

$$\sigma_x^2 \equiv \frac{1}{T} \sum_{t=1}^T (x_t - \bar{x})^2 = \overline{(x - \bar{x})^2} .$$

Die Wurzel σ_x aus der Varianz σ_x^2 heisst *Standardabweichung*. Wegen

$$\sigma_x^2 = \overline{(x - \bar{x})^2} = \overline{x^2 - 2x\bar{x} + \bar{x}^2} = \overline{x^2} - \bar{x}^2$$

und

$$\overline{(x - \bar{x})(y - \bar{y})} = \overline{xy} - \bar{x}\bar{y}$$

können wir b in (3.8) auch wie folgt schreiben,

$$b = \frac{\overline{(x - \bar{x})(y - \bar{y})}}{\sigma_x^2} . \quad (3.9)$$

Setzen wir nun a aus (3.8) in (3.2) ein, so erhalten wir

$$y_t^* = \bar{y} + b(x_t - \bar{x}) .$$

Dies kann schließlich unter Beachtung von (3.9) wie folgt geschrieben werden:

$$\frac{y_t^* - \bar{y}}{\sigma_y} = \frac{\overline{(x - \bar{x})(y - \bar{y})}}{\sigma_x \sigma_y} \frac{x_t - \bar{x}}{\sigma_x} \quad (3.10)$$

Der Anstieg

$$\rho_{x,y} \equiv \frac{\overline{(x - \bar{x})(y - \bar{y})}}{\sigma_x \sigma_y} \quad (3.11)$$

heißt *Korrelationskoeffizient* der Zeitreihen (3.1). Eine einfache Rechnung ergibt

$$\rho_{x,y} = \frac{\overline{xy} - \bar{x}\bar{y}}{\sigma_x \sigma_y} . \quad (3.12)$$

Häufig normiert man die Signale auf Mittelwert 0 und die Varianz 1, indem man von x_t zu

$$X_t \equiv \frac{x_t - \bar{x}}{\sigma_x} \quad \text{bzw.} \quad Y_t \equiv \frac{y_t - \bar{y}}{\sigma_y} \quad (3.13)$$

übergeht. Dann vereinfacht sich (3.11) zu

$$\rho_{x,y} = \overline{XY} . \quad (3.14)$$

Gleichung (3.10) geht dann über in

$$Y_t^* = \rho_{x,y} X_t . \quad (3.15)$$

Somit erkennen wir, dass der Korrelationskoeffizient $\rho_{x,y}$ der Anstieg der Regressionsgeraden (3.15) durch die Punktwolke $\{(X_t, Y_t)\}_t$ ist, sofern man also die normierten Größen Y_t über X_t darstellt (vgl. Bild 3.1).

Der Zähler in (3.11), also

$$\text{cov}_{x,y} \equiv \overline{(x - \bar{x})(y - \bar{y})} , \quad (3.16)$$

heißt *Kovarianz*. Der Korrelationskoeffizient kann somit als normierte Kovarianz aufgefasst werden.

3.2 Eigenschaften

Wir führen nun einige wichtige Eigenschaften des Korrelationskoeffizienten auf:

Symmetrie: Unmittelbar aus der Definition (3.11) folgt

$$\rho_{x,y} = \rho_{y,x} .$$

Beschränktheit: Es gilt immer

$$-1 \leq \varrho_{x,y} \leq 1 . \quad (3.17)$$

Dies folgt aus der Schwarzschen Ungleichung

$$\left(\sum_t x_t y_t \right)^2 \leq \sum_t x_t^2 \cdot \sum_t y_t^2 ,$$

die für beliebige Vektoren $(x_1, \dots, x_t, \dots, x_T) \in \mathbb{R}^T$ und $(y_1, \dots, y_t, \dots, y_T) \in \mathbb{R}^T$ gilt. Ersetzt man hier x_t durch $(x_t - \bar{x})/T$ sowie y_t durch $(y_t - \bar{y})/T$, und zieht auf beiden Seiten die Wurzel, so folgt

$$\left| \overline{(x - \bar{x})(y - \bar{y})} \right| \leq \sigma_x \sigma_y .$$

Unter Beachtung von (3.11) folgt schließlich (3.17). Wegen $\varrho_{x,y} = 1$ für $y = x$ und $\varrho_{x,-x} = -1$ für $y = -x$ kann die Ungleichung nicht verschärft werden.

Man nennt die Signale im Fall $\varrho_{x,y} = 0$ *unkorreliert* oder auch *linear unabhängig*. Für $\varrho_{x,y} > 0$ bzw. $\varrho_{x,y} < 0$ heißen die Signale *positiv* bzw. *negativ korreliert*.

Determiniertheit: Für $\varrho_{x,y} \pm 1$ sind x und y gegenseitig *determiniert*. Das bedeutet, es gilt $\varrho_{x,y} = \pm 1$ dann und nur dann, wenn y eine Funktion von x ist, und zwar

$$y_t = a + b x_t , \quad (3.18)$$

mit reellen Konstanten a und $b \neq 0$. y_t ist hier also durch x_t eindeutig bestimmt (determiniert) bzw. umgekehrt auch x_t durch $(y_t - a)/b$.

Zum Beweis nehmen wir zunächst an, dass (3.18) gilt. Dann erhalten wir $|b|\sigma_x = \sigma_y$ und

$$\varrho_{x,y} = \frac{\overline{(x - \bar{x})(a + b x - a + b x)}}{|b|\sigma_x^2} = \frac{\overline{b(x - \bar{x})^2}}{|b|\sigma_x^2} = \operatorname{sgn} b = \pm 1 .$$

Gelte nun $\varrho_{x,y} = 1$, dann folgt wegen (3.14)

$$\overline{XY} = 1 .$$

Somit gilt auch

$$\overline{(X - Y)^2} = \overline{X^2 - 2XY + Y^2} = 0 .$$

Daraus folgt, dass mit Sicherheit (also Wahrscheinlichkeit 1) $X_t - Y_t = 0$ gilt, und folglich (vgl. (3.13))

$$y_t = \bar{y} + \sigma_y \frac{x_t - \bar{x}}{\sigma_x} .$$

Im Falle $\varrho_{x,y} = \overline{XY} = -1$ zeigt man aus

$$\overline{(X + Y)^2} = \overline{X^2 + 2XY + Y^2} = 1 - 2 + 1 = 0 ,$$

dass dann mit Sicherheit $X_t + Y_t = 0$, also

$$y_t = \bar{y} - \sigma_y \frac{x_t - \bar{x}}{\sigma_x}$$

gilt.

Unabhängigkeit: Statistisch unabhängige Signale sind unkorreliert, denn dann gilt für lange Signale ($T \rightarrow \infty$) mit Sicherheit $\overline{xy} = \bar{x} \bar{y}$. Aus (3.12) folgt dann $\rho_{x,y} = 0$.

Die Umkehrung gilt nicht. Zuweilen trifft man auf die falsche Vorstellung, dass auch umgekehrt aus der Unkorreliertheit die Unabhängigkeit folge. Dies stammt daher, dass der Umkehrschluss in Spezialfällen gemacht werden kann, etwa dann, wenn (x, y) (zweidimensional) normalverteilt ist.

3.3 Korrelationsfunktion

Fasst man den Korrelationskoeffizient (3.11) zweier Zeitreihen $\{x_t\}$ und $\{y_{t+\tau}\}$ als Funktion vom Zeitversatz τ auf, so erhält man die *Korrelationsfunktion*,

$$\rho_{x,y}(\tau) \equiv \frac{\overline{(x_t - \bar{x})(y_{t+\tau} - \bar{y})}}{\sigma_x \sigma_y} \quad (3.19)$$

Im Falle $x_t = y_t$ nennt man sie auch *Autokorrelationsfunktion* (AKF) und andernfalls *Kreuzkorrelationsfunktion* (KKF). Zur Vereinfachung schreiben wir im Falle der AKF auch $\rho(\tau)$ für $\rho_{x,x}(\tau)$.

Zur Illustration betrachten wir die Zeitreihen in Bild 3.2. Dabei handelt es sich um folgende Signale:

a) **Sinus Signal:**

$$x_t = \sin(2\pi \cdot 0,023 \cdot (t - 1)) \quad , \quad t = 1, 2, \dots, 1000 \quad .$$

b) — e) **Sprachsignale:** Isoliert (lang) gesprochene Allophone /a:/, /m:/, /dsch:/ und /f:/ im Frequenzbereich von 30 — 3600 Hz, bei einer Abtastperiode $T_s = 125\mu\text{s}$ ($1/T_s = 8\text{kHz}$) und mit einer Amplitudenauflösung von 8 bit.

f) **Chaotisches Pendel:** Winkelgeschwindigkeit eines parametrisch erregten Pendels im chaotischen Bewegungsregime.

g) **Chaotische 1D–Abbildung:** Logistische Abbildung,

$$x_{t+1} = 4x_t(1 - x_t) \quad , \quad x_1 = 1/\pi \quad , \quad t = 1, 2, \dots, 1000 \quad . \quad (3.20)$$

h) **Chaotische 2D–Abbildung:** Hénon–Abbildung,

$$x_{t+1} = (1 - ax_t^2 + bx_{t-1}) \quad , \quad x_1 = 1 \quad , \quad x_0 = 0 \quad , \quad t = 1, 2, \dots, 1000 \quad .$$

i) **Wechselkurs–Änderung:** $\{y_t\}$ sei die Zeitreihe der Wechselkurse US\$/Can.\$ der Jahre 1971 bis 1974, dann ist unsere Zeitreihe

$$x_t = \ln \frac{y_t}{y_{t-1}} \quad , \quad t = 1, 2, \dots, 1000 \quad .$$

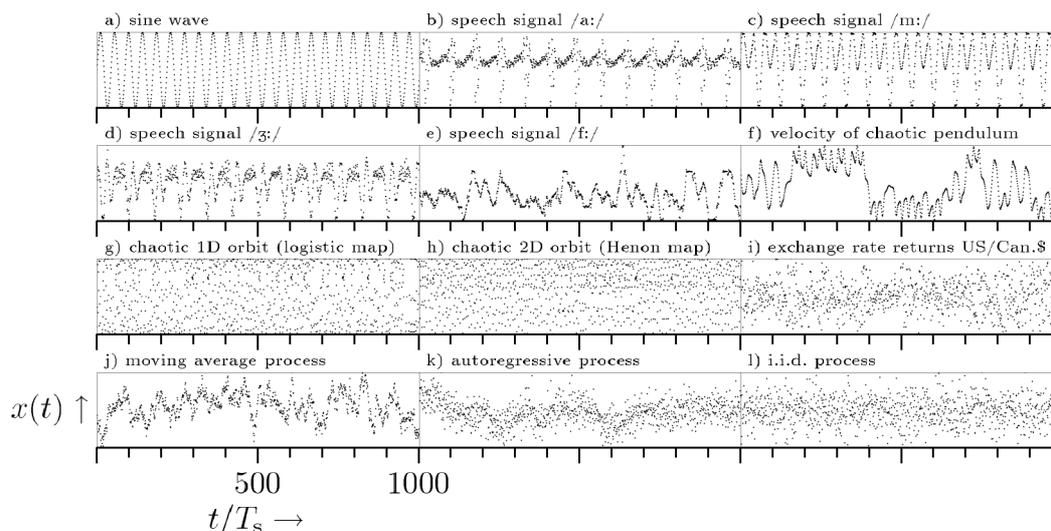


Bild 3.2: Zeitreihen, für welche im Bild 3.3 die Korrelationsfunktionen dargestellt sind

j) Gleitendes Mittel:

$$x_t = \frac{1}{\sqrt{20}} \sum_{k=1}^{20} n(t-k) , \quad t = 1, 2, \dots, 1000 . \quad (3.21)$$

Dabei stammt $\{n_t\}$ aus einem gaußverteilter i.i.d. Prozess (s. S. 61).

k) Autoregressiver Prozess:

$$x_t = 0,19 \cdot x_{t-6} + 0,2 \cdot \sum_{k=7}^{10} x(t-k) + 0,01 \cdot n(t) , \quad t = 1, 2, \dots, 1000 .$$

Dabei stammt $\{n_t\}$ aus einem gaußverteilter i.i.d. Prozess.

l) Rauschen: Vollkommen statistisch unabhängiger gaußverteilter Prozess (i.i.d. Prozess).

Die zugehörigen AKF sind im Bild 3.3 zu sehen. Es fällt zunächst auf, dass die AKF an der Stelle $\tau = 0$ gleich eins ist. Dies muss ganz allgemein gelten, haben wir doch oben schon gezeigt, dass $y = x$ (vgl. Gl. (3.18) für $a = 0$ und $b = 1$) dann und nur dann gilt, wenn $\varrho_{x,y} = 1$ gilt. Es sei bemerkt, dass die KKF an der Stelle $\tau = 0$ nicht notwendig eins ist.

Weiterhin fällt im Bild 3.3 auf, dass die AKF spiegelsymmetrisch zur Stelle $\tau = 0$ ist, also

$$\rho(\tau) = \rho(-\tau) .$$

Dies gilt ebenfalls ganz allgemein für die AKF, sofern die Zeitreihe $\{x_t\}$ stationär ist.²⁾ Wir haben dann

$$\rho(\tau) = \varrho_{x_t, x_{t+\tau}} = \frac{\overline{x_t x_{t+\tau}} - \bar{x}^2}{\sigma_x^2} = \frac{\overline{x_{t-\tau} x_t} - \bar{x}^2}{\sigma_x^2} = \frac{\overline{x_t x_{t-\tau}} - \bar{x}^2}{\sigma_x^2} = \rho(-\tau) .$$

²⁾ Stationarität einer Zeitreihe bedeutet, dass sich die statistischen Eigenschaften einer Zeitreihe nicht mit der Zeit t ändern. Das bedeutet insbesondere, dass für eine nahezu beliebige Funktion $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ der Mittelwert $\overline{g(x_t, y_{t+\tau})}$ nur vom Zeitversatz τ , nicht aber vom Zeitpunkt t abhängt. Später sagen wir noch genauer, was wir unter Stationarität verstehen.

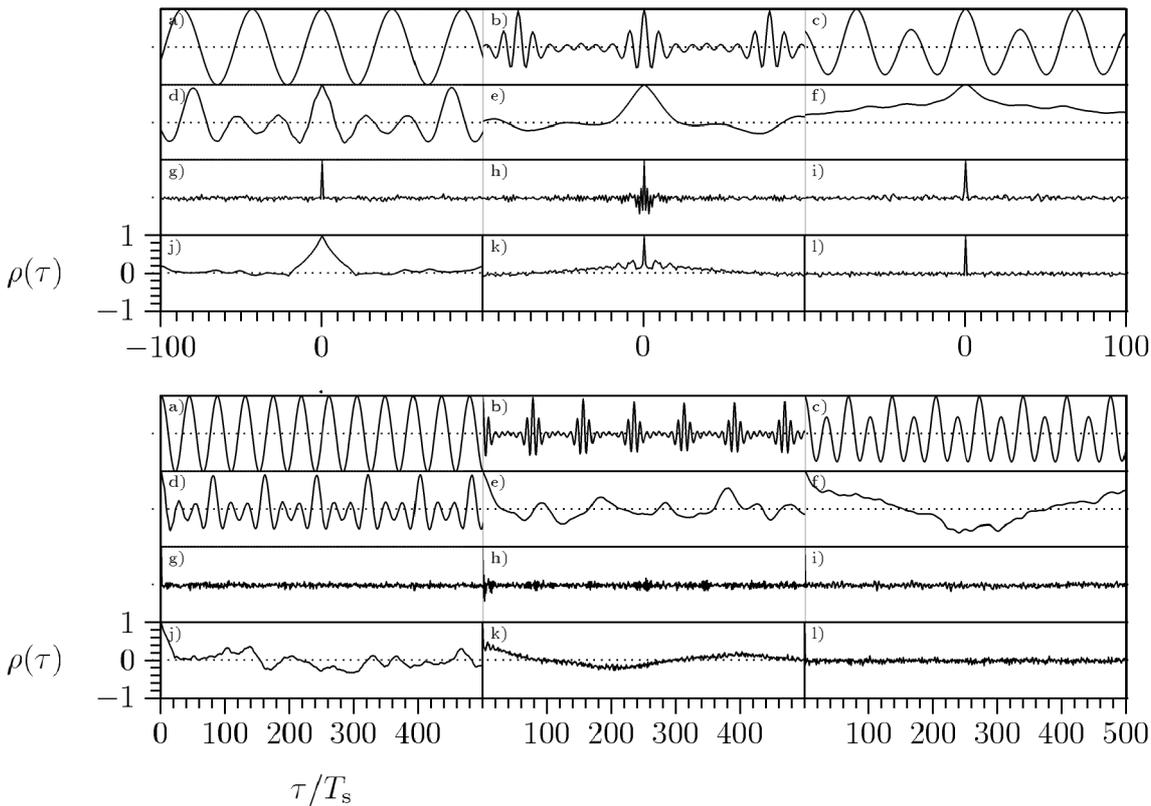


Bild 3.3: Autokorrelationsfunktionen der Zeitreihen im Bild 3.2

Wegen dieser Symmetrie stellt man die AKF üblicherweise nur für $\tau \geq 0$ dar, wie im Bild 3.3 unten. Die KKF ist allerdings im Allgemeinen asymmetrisch, $\varrho_{x,y}(\tau) \neq \varrho_{x,y}(-\tau)$.

Das Signal a) in Bild 3.2 ist periodisch, und die stimmhaften Sprachsignale b) — d) sind nahezu periodisch. Die entsprechenden AKF sind dann auch (nahezu) periodisch (vgl. Bild 3.3 unten). Dies ist eine allgemein gültige Eigenschaft der AKF: Ist $\{x_t\}$ periodisch, also $x_t = x_{t+T_0}$, so ist dies auch die zugehörige AKF, und zwar mit derselben Periode T_0 , $\text{AKF}(\tau) = \text{AKF}(\tau+T_0)$. Dies folgt aus

$$\rho(\tau) = \frac{\overline{x_t x_{t+\tau}} - \bar{x}^2}{\sigma_x^2} = \frac{\overline{x_t x_{t+\tau+T_0}} - \bar{x}^2}{\sigma_x^2} = \rho(\tau + T_0) .$$

3.4 Korrelationsfunktion für zeitkontinuierliche Signale

Wir haben bislang angenommen, dass eine Zeitreihe (ein Signal) zeitlich diskret vorliegt. Für praktische Belange der Zeitreihenanalyse ist dies durchaus ausreichend. Für die theoretische Beschreibung von Signalen empfiehlt es sich jedoch, das Kalkül auch für zeitkontinuierliche Signale auszubauen.

Die Kreuzkorrelationsfunktion wird für zeitkontinuierliche Signale $x(t)$ und $y(t)$, $t \in \mathbb{R}$, wie

folgt definiert,

$$p_{x,y}(\tau) \equiv \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t)y(t+\tau) dt \quad (3.22)$$

Diese Definition wird vor allem in der Signaltheorie und Elektrotechnik verwendet. Ein Vergleich mit (3.12) und (3.19) ergibt

$$\rho_{x,y}(\tau) = \frac{p_{x,y}(\tau) - \bar{x}\bar{y}}{\sigma_x\sigma_y} . \quad (3.23)$$

Darin ist der Mittelwert eines zeitkontinuierlichen Signals $x(t)$ durch

$$\bar{x} \equiv \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t) dt$$

definiert, und seine Varianz durch

$$\sigma_x^2 \equiv \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} [x(t) - \bar{x}]^2 dt .$$

Sind die Signale normiert, so dass $\bar{x} = 0 = \bar{y}$ und $\sigma_x = 1 = \sigma_y$ gelten, dann folgt offenbar $\rho_{x,y}(\tau) = p_{x,y}(\tau)$. Im Falle der AKF, also für $x = y$, schreiben wir auch hier wieder vereinfacht $\rho(\tau)$ statt $\rho_{x,x}(\tau)$.

3.4.1 Kreuzkorrelationsfunktion für periodische Signale

Sind die Signale zeitkontinuierlich und periodisch mit der Periode $T_0 \equiv 2\pi/\omega_0$, dann vereinfacht sich der Ausdruck (3.22) zu

$$p_{x,y}(\tau) = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} x(t)y(t+\tau) dt . \quad (3.24)$$

Unter Verwendung der Fourier-Reihenentwicklungen (1.12) und (1.13) für $x(t)$ und $y(t)$ erhalten wir

$$\begin{aligned} p_{x,y}(\tau) &= \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} x(t) \left[\sum_{n=-\infty}^{+\infty} Y(jn\omega_0) e^{jn\omega_0(t+\tau)} \right] dt \\ &= \sum_{n=-\infty}^{+\infty} \left[Y(jn\omega_0) e^{jn\omega_0\tau} \cdot \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} x(t) e^{jn\omega_0 t} dt \right] \\ &= \sum_{n=-\infty}^{+\infty} X^*(jn\omega_0) Y(jn\omega_0) e^{jn\omega_0\tau} . \end{aligned}$$

Mit dem sogenannten *Kreuzleistungsspektrum*,

$$P_{x,y}(jn\omega_0) \equiv X^*(jn\omega_0)Y(jn\omega_0) , \quad (3.25)$$

folgt schließlich

$$p_{x,y}(\tau) = \sum_{n=-\infty}^{+\infty} P_{x,y}(jn\omega_0) e^{jn\omega_0\tau} . \quad (3.26)$$

Um die Bezeichnung ‘‘Kreuzleistungsspektrum’’ zu rechtfertigen, bestimmen wir nun die mittlere Leistung des Signals $x(t) + y(t + \tau)$,

$$\begin{aligned} l_{x+y}(\tau) &= \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} [x(t) + y(t + \tau)]^2 dt \\ &= \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} x^2(t) dt + \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} y^2(t) dt + \frac{2}{T_0} \int_{-T_0/2}^{T_0/2} x(t)y(t + \tau) dt \\ &= l_x + l_y + 2 p_{x,y}(\tau) . \end{aligned}$$

Somit kann $p_{x,y}(\tau)$ im Wesentlichen (bis auf den Faktor 2) als *mittlere Kreuzleistung* der Signale $x(t)$ und $y(t + \tau)$ aufgefasst werden. Fur $\tau = 0$ erhalten wir

$$p_{x,y}(0) = \sum_{n=-\infty}^{+\infty} P_{x,y}(jn\omega_0) .$$

Aus dem Vergleich des Ausdrucks (3.26) mit (1.12) folgern wir, dass die Kreuzkorrelationsfunktion $p_{x,y}(\tau)$ ebenfalls periodisch ist, mit der Periode T_0 . Aus (1.13) finden wir

$$P_{x,y}(jn\omega_0) = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} p_{x,y}(\tau) e^{-jn\omega_0\tau} d\tau .$$

Wir haben somit gezeigt, dass das Kreuzleistungsspektrum $P_{x,y}(jn\omega_0)$ die Fourier–Transformierte der Kreuzkorrelationsfunktion $p_{x,y}(\tau)$ ist. Umgekehrt ist nach Gl. (3.26) die Kreuzkorrelationsfunktion die inverse Fourier–Transformierte des Kreuzleistungsspektrums (3.26). Die Kreuzkorrelationsfunktion enthalt also genau dieselbe Information wie das Kreuzleistungsspektrum.

3.4.2 Autokorrelationsfunktion fur periodische Signale

Fur $x(t) = y(t)$ vereinfacht sich das Kreuzleistungsspektrum (3.25) zu

$$P(jn\omega_0) \equiv P_{x,x}(jn\omega_0) = |X(jn\omega_0)|^2 .$$

Damit folgt fur die KKF (3.26), die nun zur AKF wird,

$$p(\tau) = \sum_{n=-\infty}^{+\infty} |X(jn\omega_0)|^2 e^{jn\omega_0\tau} . \quad (3.27)$$

Das Leistungsspektrum (power spectrum) $P(jn\omega_0)$ ist die Fourier–Transformierte der Autokorrelationsfunktion $p(\tau)$. Umgekehrt ist die Autokorrelationsfunktion die inverse Fourier–Transformierte des Leistungsspektrums.

Als Beispiel betrachten wir die harmonische Funktion

$$x(t) = A \sin(\omega_0 t + \varphi) .$$

Sie hat das Power-Spektrum

$$|X(jn\omega_0)| = \begin{cases} A/2 & : n = \pm 1 \\ 0 & : \text{sonst.} \end{cases}$$

Setzen wir dies in (3.27) ein, so erhalten wir die Autokorrelationsfunktion eines harmonischen Signals,

$$p(\tau) = \frac{A^2}{4} (e^{j\omega_0\tau} + e^{-j\omega_0\tau}) = \frac{A^2}{2} \cdot \cos \omega_0\tau . \quad (3.28)$$

Es gilt hier $\bar{x} = 0$ sowie $p(0) = A^2/2 = \sigma_x^2$. Für den Korrelationskoeffizient zwischen $\{x(t)\}_t$ und $\{x(t + \tau)\}_t$ folgt dann nach (3.23),

$$\varrho(\tau) = \frac{p(\tau)}{p(0)} = \cos \omega_0\tau .$$

Dies ist in Übereinstimmung mit Bild 3.3a).

Unter Verwendung bekannter trigonometrischer Additionstheoreme³⁾ kann (3.28) auch leicht wie folgt erhalten werden,

$$\begin{aligned} p(\tau) &= \frac{1}{T_0} \int_0^{T_0} A \sin(\omega_0 t + \varphi) \cdot A \sin(\omega_0(t + \tau) + \varphi) dt \\ &= \frac{A^2}{2} \cdot \frac{1}{T_0} \int_0^{T_0} \cos \omega_0\tau dt \\ &= \frac{A^2}{2} \cdot \cos \omega_0\tau . \end{aligned}$$

³⁾Es gilt $\cos(x - y) = \cos x \cos y + \sin x \sin y$.

Kapitel 4

Stochastischer Prozess

4.1 Grundlegenden Begriffe

Messen wir in der Praxis ein Signal mehrere Male unter vermeintlich gleichen Bedingungen, so werden wir niemals vollkommen gleiche Signale erhalten. Wir sagen dazu, dass reale Signale immer mehr oder weniger *verrauscht* sind. Diese Zufälligkeiten werden im Allgemeinen umso deutlicher, wenn wir genauer messen. Sie finden mit dem Begriff vom *stochastischen Prozess* eine mathematische Beschreibung.

Ein *stochastischer* oder auch *zufälliger Prozess* $\{X_t\}_t$ ist eine Folge von Zufallsgrößen. Eine *Zufallsgröße* X_t ist wieder eine Abbildung von der Menge der betrachteten zufälligen Ereignisse in die reellen Zahlen \mathbb{R} . Die zufälligen Ereignisse sind bei uns die möglichen Messwerte. Gibt es nur abzählbar viele Ereignisse, dann heißt der Prozess *diskret* und andernfalls *kontinuierlich*. Ein Signal $\{x(t)\}$ wird als Realisierung von einem $\{X(t)\}_{t \in I}$ aufgefasst. Beim sogenannten *zeitdiskreten Prozess* durchläuft die Zeit t alle ganzen Zahlen, d. h., für die Indexmenge gilt $I = \mathbb{I}$. Hingegen läuft die Zeit beim *kontinuierlichen Prozess* in den reellen Zahlen, $I = \mathbb{R}$.

Aus einem Prozess $\{X_t\}$ können wir den D -dimensionalen Zufallsvektor

$$\mathbf{X}_\tau(t) \equiv (X_{t+\tau_1}, X_{t+\tau_2}, \dots, X_{t+\tau_d}, \dots, X_{t+\tau_D}) \quad (4.1)$$

konstruieren, mit den Zeitversätzen

$$\boldsymbol{\tau} \equiv (\tau_1, \tau_2, \dots, \tau_d, \dots, \tau_D) \quad , \quad \tau_d \in I \quad . \quad (4.2)$$

Der Prozess wird *stationär*, genannt, wenn für beliebige $\boldsymbol{\tau}$ und zu beliebigen Zeitpunkten t_1 und t_2 die Vektoren $\mathbf{X}_\tau(t_1)$ und $\mathbf{X}_\tau(t_2)$ die gleiche Wahrscheinlichkeitsverteilung haben. In diesem Sinne hängen bei einem stationären Prozess $\{X_t\}$ die statistischen Eigenschaften nicht vom Zeitpunkt t ab.¹⁾

¹⁾Nehme der Zufallsvektor $\mathbf{X} = (X_1, \dots, X_D)$ die Werte $\mathbf{x} = (x_1, \dots, x_D) \in \mathbb{R}^D$ an. Dann ist die Wahrscheinlichkeitsverteilung (kurz: Verteilung) S von \mathbf{X} definiert als die Wahrscheinlichkeit, mit der $X_d \leq x_d$, $d = 1, \dots, D$, gilt. S ist also die Funktion $S : \mathbb{R}^D \rightarrow [0, 1]$, so dass

$$P(x_1, \dots, x_D) = \text{prob}(X_1 \leq x_1, \dots, X_D \leq x_D) = \int_{x_d^* : x_d^* \leq x_d, d=1, \dots, D} dP(\mathbf{x}^*) \quad .$$

Wenn es eine Funktion $p : \mathbb{R}^D \rightarrow \mathbb{R}$, $p(\mathbf{x}) \geq 0$, gibt, aus der die Verteilung $P(\mathbf{x})$ gemäß

$$P(x_1, \dots, x_D) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_D} p(\mathbf{x}^*) d\mathbf{x}^*$$

Häufig interessieren wir uns für den *Mittelwert einer Funktion* $f(x_1(t), \dots, x_D(t)) : \mathbb{R}^D \rightarrow \mathbb{R}$. Dabei unterscheiden wir vier Fälle:

Zeitmittel zeitdiskreter Signale:

$$\bar{f}^{\text{Zeit}} \equiv \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{t=T} f(x_1(t), \dots, x_D(t)) \quad , \quad (4.3)$$

sofern dieser Grenzwert existiert, was wir hier immer voraussetzen.

Zeitmittel zeitkontinuierlicher Signale:

$$\bar{f}^{\text{Zeit}} \equiv \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T f(x_1(t), \dots, x_D(t)) dt \quad , \quad (4.4)$$

sofern das Integral und der Grenzwert existieren.

Scharmittel amplitudendiskreter Signale: Nehme X_d die Werte x_{d,m_d} , $m_d = 1, \dots, M_d$, mit den Wahrscheinlichkeiten p_{d,m_d} an, und (X_1, \dots, X_D) die Werte $(x_{1,m_1}, \dots, x_{D,m_D})$ mit den Verbundwahrscheinlichkeiten $s_{m_1 \dots m_D}$. Dann definieren wir

$$\bar{f}^{\text{Schar}} \equiv \sum_{m_1, \dots, m_D=1}^{M_1, \dots, M_D} f(x_{1,m_1}, \dots, x_{D,m_D}) s_{m_1 \dots m_D} \quad . \quad (4.5)$$

Scharmittel amplitudenkontinuierlicher Signale: Nehme X_d die Werte $x_d \in \mathbb{R}$ mit der Wahrscheinlichkeitsdichte $p_d(x_d)$ an, und (X_1, \dots, X_D) die Werte $\mathbf{x} \equiv (x_1, \dots, x_D) \in \mathbb{R}^D$ mit der Verbundwahrscheinlichkeitsdichte $p(\mathbf{x})$. Dann definieren wir

$$\bar{f}^{\text{Schar}} \equiv \int_{\mathbb{R}^D} f(\mathbf{x}) p(\mathbf{x}) d\mathbf{x} \quad , \quad (4.6)$$

sofern dieser Grenzwert existiert.

berechnet werden kann, dann heisst p *Verteilungsdichte*. Das entsprechende *Wahrscheinlichkeitsmaß* $\mu : B \rightarrow [0, 1]$, das den Borelmengen $B \subseteq \mathbb{R}^D$ eine reelle Zahl (Wahrscheinlichkeit)

$$\mu(B) = \int_B p(\mathbf{x}^*) d\mathbf{x}^*$$

zuordnet, ist dann absolut stetig bezüglich des Lebesgue-Maßes λ auf den Borelmengen aus \mathbb{R}^D . μ heisst *absolut stetig* bezüglich λ , wenn aus $\lambda(B) = 0$ auch $\mu(B) = 0$ folgt. Jede Nullmenge bezüglich λ ist hier also auch Nullmenge bezüglich μ .

Eine *Borelmenge* ist eine Teilmenge aus \mathbb{R}^D . Beispielsweise ist jeder offener, abgeschlossener und halboffener D -dimensionaler Quader eine Borelmenge. Darüber hinaus ist jede abzählbare (möglicherweise unendliche) Vereinigung oder auch jeder abzählbare Durchschnitt von Borelmengen wieder eine Borelmenge. Die Borelmengen bilden somit eine sogenannte σ -Algebra von Teilmengen des \mathbb{R}^D , auf der ein Wahrscheinlichkeits-Maß in natürlicher Weise definiert werden kann.

Die d -te eindimensionale *Randverteilung* von P ist die Funktion $P_d : \mathbb{R} \rightarrow \mathbb{R}$,

$$P_d(x_d) \equiv P(+\infty, \dots, +\infty, x_d, +\infty, \dots, +\infty) \quad .$$

Sie entsteht also im Falle absolut stetiger Zufallsgrößen durch Integration der Dichte s über alle Variablen, außer x_d .

Solche Wahrscheinlichkeitsdichten existieren nicht immer. Beispielsweise hat $(X(t), X(t + \tau))$ keine Dichte, wenn die Realisierung von $\{X(t)\}$ eines der chaotischen Signale f — h) in Bild 3.2 ist. In der Praxis der Signalverarbeitung kann man jedoch häufig in guter Näherung von der Existenz einer solchen Dichte ausgehen, denn reale Zeitreihen sind immer verrauscht.

Mit der Verteilung $P(\mathbf{x})$ von \mathbf{X} können die beiden Mittelwerte (4.5) und (4.6) auch einheitlich wie folgt geschrieben werden,

$$\bar{f}^{\text{Schar}} \equiv \int_{\mathbb{R}^D} f(\mathbf{x}) dP(\mathbf{x}) .$$

Das Zeitmittel kann man auch für Mischformen aus zeit-diskreten und -kontinuierlichen Zufallsgrößen definieren und das Scharmittel aus Mischformen amplituden-diskreter und -kontinuierlicher Zufallsgrößen.

In der Praxis der Signalverarbeitung setzt man häufig voraus, dass für nahezu beliebige Funktionen f Zeit- und Scharmittel gleich sind, was nicht selbstverständlich ist. Gilt die Gleichheit

$$\bar{f}^{\text{Zeit}} = \bar{f}^{\text{Schar}} ,$$

so nennt man die Signale *ergodisch*. Man lässt dann die Beschriftungen „Zeit“ bzw. „Schar“ weg.

Fassen wir die Signale $\{x(t)\}$ und $\{y(t)\}$ als Realisierungen zweier ergodischer stochastischer Prozesse $\{X(t)\}$ bzw. $\{Y(t)\}$ auf, dann können wir den Korrelationsfunktion (3.19) als Mittelwert über die Zufallsgröße

$$f(X(t), Y(t + \tau)) = \frac{[X(t) - \bar{X}] \cdot [Y(t + \tau) - \bar{Y}]}{\sigma_X \cdot \sigma_Y} \quad (4.7)$$

auffassen, wobei über die Verbundverteilung $P_\tau(x, y)$ von $(X(t), Y(t + \tau))$ zu mitteln ist, also

$$\rho_{X,Y}(\tau) = \overline{f(X(t), Y(t + \tau))} = \int_{\mathbb{R}^2} \frac{(x - \bar{X}) \cdot (y - \bar{Y})}{\sigma_X \cdot \sigma_Y} dP_\tau(x, y) . \quad (4.8)$$

Im Fall, dass es eine Verbunddichte $p(x, y)$ gibt, gilt dann $dP(x, y) = p(x, y) dx dy$.

Für $X = Y$ geht (4.8) über in die Autokorrelationsfunktion,

$$\begin{aligned} \rho(\tau) &= \int_{\mathbb{R}^2} \frac{(x_1 - \bar{X})(x_2 - \bar{X})}{\sigma_X^2} dP_\tau(x_1, x_2) \\ &= \frac{1}{\sigma_X^2} \left(\int_{\mathbb{R}^2} x_1 x_2 dP_\tau(x_1, x_2) - \bar{X}^2 \right) . \end{aligned} \quad (4.9)$$

4.2 Zusammenhang von Korrelationsfunktion und Spektrodichte

Haben die Signale eine endliche Energie, gelten also $\int_{-\infty}^{+\infty} |x(t)| dt < \infty$ und $\int_{-\infty}^{+\infty} |y(t)| dt < \infty$, so definiert man in der Signal- bzw. Nachrichtentheorie die Funktion

$$e_{x,y}(\tau) \equiv \lim_{T \rightarrow \infty} \int_{-T/2}^{T/2} x(t)y(t + \tau) dt \quad (4.10)$$

und nennt sie ebenso Kreuzkorrelationsfunktion, bzw. für $x = y$ Autokorrelationsfunktion,

$$e(\tau) \equiv \lim_{T \rightarrow \infty} \int_{-T/2}^{T/2} x(t)x(t + \tau) dt \quad (4.11)$$

Im Unterschied zu (3.22) fehlt hier der Faktor $1/T$. Somit ist $e_{x,y}$ im Wesentlichen die *Kreuzenergie*, wohingegen $p_{x,y}$ eine Kreuzleistung ist.

Unter Beachtung von (4.4) können wir auch wie folgt schreiben,

$$e_{x,y}(\tau) = \overline{T \cdot X(t) Y(t + \tau)}^{\text{Zeit}} .$$

Zwischen der Kreuzkorrelationsfunktion (4.10) und der Fourierpektraldichte (vgl. Gln.1.17 und 1.19) gibt es einen Zusammenhang. Es gilt:

$$\begin{aligned} e_{x,y}(\tau) &= \int_{-\infty}^{+\infty} x(t)y(t + \tau) dt \\ &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} x(t) \int_{-\infty}^{+\infty} Y(j\omega)e^{j\omega(t+\tau)} d\omega dt \\ &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} Y(j\omega)e^{j\omega\tau} \int_{-\infty}^{+\infty} x(t)e^{j\omega t} dt d\omega \\ &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} Y(j\omega)X^*(j\omega)e^{j\omega\tau} d\omega \end{aligned}$$

Der Ausdruck $E_{x,y}(j\omega) \equiv Y(j\omega)X^*(j\omega)$ heißt *Kreuzenergiespektraldichte*. Wir schreiben damit

$$\begin{aligned} e_{x,y}(\tau) &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} E_{x,y}(j\omega)e^{j\omega\tau} d\omega \\ E_{x,y}(j\omega) &= \int_{-\infty}^{+\infty} e_{x,y}(\tau)e^{-j\omega\tau} d\tau \end{aligned} \quad (4.12)$$

Die Kreuzkorrelationsfunktion $e_{x,y}(\tau)$ ist also die inverse Fourier-Transformierte der Kreuzenergiespektraldichte $E_{x,y}(j\omega)$, bzw. ist umgekehrt $E_{x,y}$ gleich der Fourier-Transformierten von $e_{x,y}$. Dieser Zusammenhang (4.12) ist das sogenannte *Wiener-Chintchin-Theorem*.

Für zwei Zufallsgrößen X und Y wird die Kreuzspektraldichte $E_{X,Y}(j\omega)$ definiert als die Fouriertransformierte der Kreuzkorrelationsfunktion (4.10).

Gilt $x(t) = y(t)$, so ist $E(j\omega) \equiv E_{x,x}(j\omega)$ die sogenannte *Energiespektraldichte* des Signals x (im Engl., *Power Spektrum*),

$$E(j\omega) = X(j\omega)X^*(j\omega) = |X(j\omega)|^2 .$$

Entsprechend sind dann die AKF (4.11) und das Power Spektrum über die Fourier-Transformation miteinander verknüpft. Power Spektrum und AKF sind demzufolge äquivalente Möglichkeiten, lineare statistische Zusammenhänge in einem Signal zu beschreiben.

4.3 Beispiele

4.3.1 Vollkommen unabhängiger Prozess

Ein stationärer Prozess $\{X(t)\}$ heisst *vollkommen unabhängig*,²⁾ wenn die Verbundverteilung des Zufallsvektors (4.1) zu jedem Vektor von Zeitversätzen (4.2) als Produkt der eindimensionalen Randverteilungen darstellbar ist,

$$P_{\tau}(\mathbf{x}) = \prod_{d=1}^D P_d(x_d) .$$

Bezeichne nun $P_{\tau}(x_1, x_2)$ die Verteilung von $(X(t), X(t+\tau))$. Für die Autokorrelationsfunktion (4.9) erhalten wir somit für $\tau \neq 0$,

$$\begin{aligned} \sigma^2 \cdot \rho(\tau) + \bar{X}^2 &= \int_{\mathbb{R}^2} x_1 \cdot x_2 \, dP_{\tau}(x_1, x_2) \\ &= \int_{\mathbb{R}^2} x_1 \cdot x_2 \, dP_1(x_1) \, dP_2(x_2) \\ &= \int_{\mathbb{R}} x_1 \, dP_1(x_1) \cdot \int_{\mathbb{R}} x_2 \, dP_2(x_2) = \bar{X}^2 . \end{aligned}$$

Daraus folgt, falls $\sigma^2 \neq 0$,

$$\rho(\tau) = 0 .$$

Aus früheren Überlegungen wissen wir schon, dass $\rho(0) = 1$. Somit gilt insgesamt

$$\rho(\tau) = \begin{cases} 1 & : \tau = 0 \\ 0 & : \text{sonst.} \end{cases}$$

Man sagt dazu, der i.i.d. Prozess ist δ -korreliert. Ein Beispiel ist das Signal 1) in den Bildern 3.2 sowie 3.3.

Die Bezeichnung „ δ -korreliert“ rührt daher, dass im Falle ergodischer Signale die folgende Gleichheit von Schar- und Zeit-Mittel gilt,

$$\int_{\mathbb{R}^2} x_1 \cdot x_2 \, dP_{\tau}(x_1, x_2) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t) \cdot x(t+\tau) \, dt .$$

Für die Korrelationsfunktion (4.11) bedeutet dies, dass sie an der Stelle $\tau = 0$ divergiert. Sie kann dann aber für mittelwertfreie Signale, $\bar{X}^2 = 0$, durch eine δ -Funktion beschrieben werden,

$$e(\tau) = \sigma^2 \cdot \delta(\tau) .$$

Setzen wir dies in (4.12) ein, so erhalten wir

$$E_{x,y}(j\omega) = \int_{-\infty}^{+\infty} \sigma^2 \cdot \delta(\tau) \cdot e^{-j\omega\tau} \, d\tau = \sigma^2 ,$$

²⁾Im Englischen nennt man solch einen Prozess *independent identically distributed* (i.i.d.). Dabei bedeutet „identically“, dass alle Zufallsgrößen $X(t)$ dieselbe Verteilung haben, nicht aber, dass dies notwendig die Gleichverteilung ist.

für $\omega > 0$. Die Energiespektraldichte eines i.i.d. Prozesses ist also für alle Werte der Kreisfrequenz ω gleich der Varianz σ^2 des Prozesses. In Analogie zum Lichtspektrum sagt man hierfür auch: „das Spektrum eines i.i.d. Signals ist weiß“.

Die Gesamtenergie eines i.i.d. Signals ist unendlich, sofern $\sigma^2 > 0$. Man gibt daher die Energie von Rauschsignalen besser bezogen auf ein Frequenzband an, etwa in Einheiten von Ws/Hz.

Generell heisst ein Prozess mit konstanter Spektraldichte *weiß*. Dies ist nach dem Wiener–Chintchin–Theorem gleichbedeutend damit, dass der Prozess δ –korreliert ist. Ist die Spektraldichte nicht konstant, so nennt man den Prozess auch *farbig*. Farbige Prozesse sind immer auch korreliert.

Wir haben oben gezeigt, dass ein jeder i.i.d. Prozess δ –korreliert (weiß) ist. Die Umkehrung gilt jedoch im Allgemeinen nicht.

4.3.2 Korrelation normalverteilter Zufallsgrößen

Wir betrachten einen Zufallsvektor (X, Y) , der zwei–dimensional normalverteilt ist. Beide Zufallsgrößen seien normiert (Varianz 1, Mittelwert 0). Die Wahrscheinlichkeitsdichte ist dann (Bild4.1)

$$p(x, y) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left[-\frac{x^2 - 2\rho xy + y^2}{2(1-\rho^2)}\right]. \quad (4.13)$$

Daraus berechnen sich wie folgt die Randverteilungsdichten, unter Beachtung von $x^2 - 2\rho xy + y^2 = x^2(1-\rho^2) + (y-\rho x)^2$:

$$\begin{aligned} s_1(x) &= \int_{\mathbb{R}} s(x, y) dy \\ &= \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left[-\frac{x^2}{2}\right] \int_{\mathbb{R}} \exp\left[-\frac{1}{2}\left(\frac{y-\rho x}{\sqrt{1-\rho^2}}\right)^2\right] dy. \end{aligned}$$

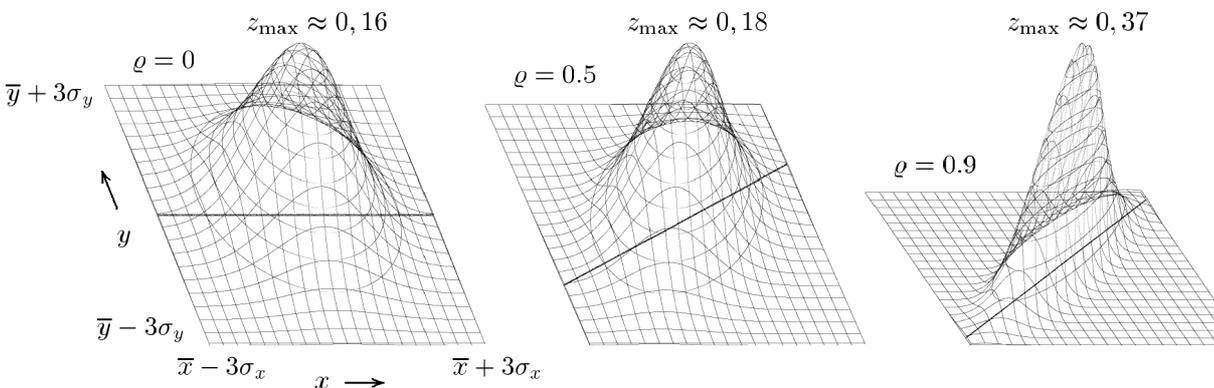


Bild 4.1: Zweidimensionale Normalverteilung bei verschiedenen Korrelationskoeffizienten ρ . Die vertikale Skaleneinteilung ist auf $z_{\max} = \left(2\pi\sqrt{1-\rho^2}\right)^{-1}$ skaliert. Der dick gezeichnete Strich im Definitionsbereich ist die Regressionsgerade, deren Anstieg ist ρ

Wegen³⁾

$$\int_{\mathbb{R}} \exp \left[-\frac{1}{2} \left(\frac{y - \varrho x}{\sqrt{1 - \varrho^2}} \right)^2 \right] dy = \sqrt{1 - \varrho^2} \sqrt{2\pi}$$

folgt schließlich

$$s_1(x) = \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{x^2}{2} \right] .$$

Analog erhalten wir die andere Randdichte

$$s_2(y) = \int_{\mathbb{R}} s(x, y) dx = \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{y^2}{2} \right] .$$

Den Korrelationskoeffizient zwischen X und Y erhalten wir aus⁴⁾

$$\begin{aligned} p_{X,Y} &= \int_{\mathbb{R}^2} xy s(x, y) dx dy \\ &= \frac{1}{2\pi\sqrt{1 - \varrho^2}} \int_{\mathbb{R}} x \exp \left[-\frac{x^2}{2} \right] \int_{\mathbb{R}} y \exp \left[-\frac{1}{2} \left(\frac{y - \varrho x}{\sqrt{1 - \varrho^2}} \right)^2 \right] dy dx \\ &= \frac{1}{2\pi\sqrt{1 - \varrho^2}} \int_{\mathbb{R}} x \exp \left[-\frac{x^2}{2} \right] \cdot \varrho x \sqrt{1 - \varrho^2} \sqrt{2\pi} dx \\ &= \frac{1}{2\pi\sqrt{1 - \varrho^2}} \sqrt{2\pi} \cdot \varrho \sqrt{1 - \varrho^2} \sqrt{2\pi} dx \\ &= \varrho . \end{aligned}$$

Offenbar folgt hier aus $\varrho = 0$ die statistische Unabhängigkeit,

$$s(x, y) = s_1(x) \cdot s_2(y) .$$

Umgekehrt hatten wir schon allgemein gezeigt, dass aus der Unabhängigkeit die Unkorreliertheit folgt. **Somit sind zwei normalverteilte Zufallsgrößen dann und nur dann statistisch unabhängig, wenn sie unkorreliert sind.**

Damit zeichnet sich die Familie der Normalverteilungen gegenüber anderen Familien von Verteilungen aus. Sie ist aber mit dieser Eigenschaft durchaus nicht die einzige Familie. Gleiches trifft beispielsweise für die folgende Familie von Wahrscheinlichkeitsdichten $\{s_\delta(x, y)\}_{\delta \in [0,1]}$ zu,

$$s_\delta(x, y) \equiv \begin{cases} 1 + \delta & : (x, y) \in (B_0 \times B_0) \cup (B_1 \times B_1) \\ 1 - \delta & : (x, y) \in (B_0 \times B_1) \cup (B_1 \times B_0) \\ 0 & : \text{sonst,} \end{cases}$$

mit den Intervallen $B_0 \equiv [0, 1/2[$ und $B_1 \equiv [1/2, 1[$. Hier erhält man nach einfacher Rechnung $\varrho = 3/4\delta$, wobei wie im Falle der Normalverteilung aus $\varrho = 0 = \delta$ die Unabhängigkeit folgt.

³⁾Wir benutzen hier das *Poisson-Integral* $\int_{\mathbb{R}} e^{-(y-a)^2/(2b^2)} dy = b\sqrt{2\pi}$.

⁴⁾Wir verwenden hier $\int_{\mathbb{R}} ye^{-\frac{(y-a)^2}{2b^2}} dy = ab\sqrt{2\pi}$ und $\int_{\mathbb{R}} x^2 e^{-\frac{x^2}{2}} dx = \sqrt{2\pi}$.

4.3.3 Vollkommen abhängiger δ -korrelierter Prozess

Wir betrachten den vorwärts-determinierten nichtlinearen Prozess

$$X(t+1) = f(X(t)) = 4X(t)(1-X(t)) . \quad (4.14)$$

Eine typische Realisierung ist Signal g in Bild 3.2. Man erhält sie, indem man (per Zufall) eine reelle Zahl $x(1)$ aus dem Intervall $[0, 1]$ wählt. Dies ist dann mit Wahrscheinlichkeit eins eine irrationale Zahl, die nach der Iterationsvorschrift, der sogenannten *logistischen Abbildung* (3.20), eine typischen chaotischen Orbit $\{x(t)\}_{t=1,2,\dots}$ generiert. Typisch bedeutet, dass seine Werte $x(t)$ nach folgender Wahrscheinlichkeitsdichte im Intervall $[0, 1]$ verteilt sind

$$p(x) = \begin{cases} \frac{1}{\pi\sqrt{x(1-x)}} & : 0 < x < 1 \\ 0 & : \text{sonst.} \end{cases} \quad (4.15)$$

Mit diesem Wahrscheinlichkeitsmaß ist der Prozess ergodisch.

Im Bild 4.2 sind einige Iterierte der logistischen Abbildung f dargestellt, sowie die zugehörige f -invariante Wahrscheinlichkeitsdichte.⁵⁾

Solch ein Orbit (Signal) ist δ -korreliert, was wir aus folgender Rechnung erhalten:
Der Mittelwert ist hier $\bar{X} = \int_0^1 x \cdot p(x) dx = 1/2$. Unter Beachtung von $x_\tau = f^\tau(x_0)$ folgt

$$\begin{aligned} \sigma_X^2 \cdot \rho(\tau) &= \int_0^1 \int_0^1 (x_0 - 0,5) \cdot (x_\tau - 0,5) dP(x_0, x_\tau) \\ &= \int_0^1 (x - 0,5) \cdot (f^\tau(x) - 0,5) \cdot p(x) dx . \end{aligned} \quad (4.16)$$

Der Integrand $h(x) \equiv (x - 0,5)(f^\tau(x) - 0,5)p(x)$ ist schiefsymmetrisch bez. $x = 0,5$, denn f ist symmetrisch, $f(x) = f(1-x)$. Folglich ist auch die τ -te Iterierte $f^\tau \equiv f \circ f^{\tau-1}$ symmetrisch, also $f^\tau(x) = f^\tau(1-x)$ für $\tau \geq 1$. (Hierbei setzen wir $f^0(x) = x$.) Für die Dichte p gilt das Gleiche, $p(x) = p(1-x)$. Damit haben wir

$$h(1-x) = ((1-x) - 0,5)(f^\tau(1-x) - 0,5)p(1-x) = -h(x) .$$

Somit verschwindet das Integral in (4.16), es gilt also $\rho(\tau) = 0$ für $\tau \neq 0$. Ein typischer Orbit der logistischen Abbildung (4.14) ist somit δ -korreliert. **Es gibt also Prozesse, die determiniert und gleichzeitig unkorreliert sind.** Die Abhängigkeiten zwischen $X(t)$ und $X(t+\tau)$ sind hier rein nichtlinear. Sie werden mit der Korrelationsfunktion nicht erfasst, da diese nur lineare Zusammenhänge beschreibt. Dies macht die bedingte Interpretationsmöglichkeit der Korrelationsfunktion als Maß für statistische Abhängigkeiten besonders deutlich.

4.3.4 Gleitmittel-Prozess

Sei $\{X(t)\}_t$ ein i.i.d. Prozess mit Mittelwert Null und Varianz σ_X^2 . Daraus konstruieren wir einen neuen Prozess $\{Y(t)\}_t$, indem wir

$$Y(t) \equiv \sum_{d=0}^D a_d \cdot X(t-d) \quad (4.17)$$

⁵⁾Ist B eine Teilmenge (Borelmenge) aus dem Intervall $[0, 1]$ und μ das zur Dichte (4.15) gehörige Wahrscheinlichkeitsmaß, so gilt $\mu(B) = \mu(f^{-1}(B))$. Man sagt dafür, dass μ ein *invariantes Maß* bezüglich der Abbildung f ist. Dabei ist $f^{-1}(B)$ die Menge aller Punkte aus $[0, 1]$, die durch f auf B abgebildet wird. Ist ein Maß invariant bezüglich f , so ist es offenbar auch invariant bezüglich jeder höheren Iterierten f^τ .

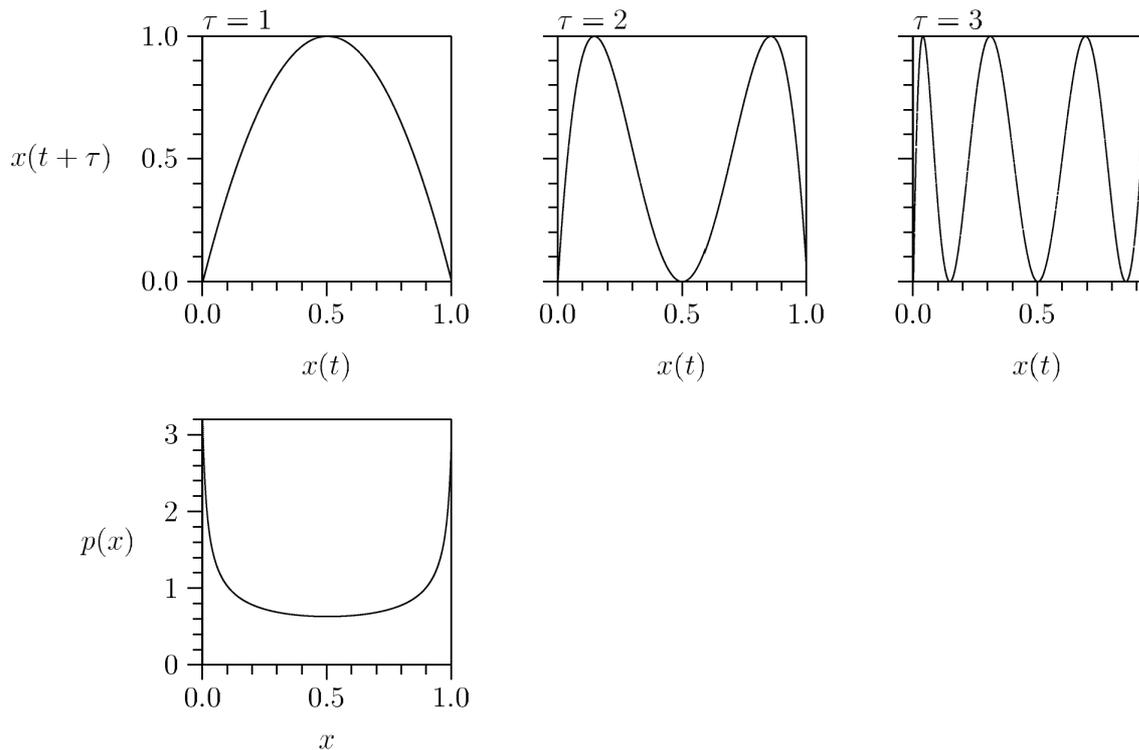


Bild 4.2: Oben:Logistische Abbildung $x(t) \rightarrow x(t + \tau) = f^\tau(x(t))$ für verschiedene Iterierte $\tau = 1; 2; 3$ Unten: Zugehörige Wahrscheinlichkeitsdichte (4.15)

setzen. Dieser Prozess heisst dann *Gleitmittel-Prozess der Ordnung D* und wird mit $MA(D)$ abgekürzt. Hierbei setzen wir $a_D \neq 0$ voraus, $a_d \in \mathbb{R}$ sind Konstanten. Man skaliert $X(t)$ gewöhnlich so, dass $a_0 = 1$ gilt. ⁶⁾

Für den Mittelwert erhalten wir

$$\bar{Y} = 0 .$$

Die Varianz ist

$$\sigma_Y^2 = \sigma_X^2 \cdot \sum_{d=0}^D a_d^2 ,$$

was durch einfache Rechnung aus der Unabhängigkeit der $X(t)$ folgt. Die Autokorrelationsfunktion des Prozesses (4.17) ist

$$\rho(\tau) = \begin{cases} 1 & : \tau = 0 \\ \left(\frac{\sigma_X}{\sigma_Y}\right)^2 \cdot \sum_{d=0}^{D-\tau} a_d \cdot a_{d+\tau} & : \tau = 1, \dots, D \\ 0 & : \tau > D \end{cases} \quad (4.18)$$

Unser Beispiel (3.21) für ein MA-Signal (Bild 3.2 j)) hat die Ordnung 20, folglich sollte die AKF im Bild 3.3 j) für $\tau > 20$ verschwinden. Dass dies nicht genau gilt, hängt damit

⁶⁾Im Englischen heisst der Gleitmittel-Prozess *moving average* (MA). Für $a_d = 1/(D+1)$, $d = 0; 1; \dots; D$, liefert dieser Prozess die Bildung des arithmetischen Mittelwertes über Fenster $x(t), x(t-1), \dots, x(t-D)$ der Zeitreihe $\{x(t)\}$. Diese Fenster gleiten bei laufendem t über die Zeitreihe, woher die Bezeichnung „Gleitmittel-Prozess“ stammt. Heutzutage wird dieser Begriff aber für beliebige Koeffizientensätze $\{a_d\}$ verwendet.

zusammen, dass wir die AKF zu einem konkreten endlichen Signal immer nur schätzen können. Die Abweichungen von der theoretischen AKF (4.18) sind die Schätzfehler.

4.3.5 Autoregressiver Prozess

Sei $\{X(t)\}_t$ ein i.i.d. Prozess mit Mittelwert Null und Varianz σ_X^2 . Daraus konstruieren wir einen neuen Prozess $\{Y(t)\}_t$, indem wir

$$Y(t) \equiv X(t) + \sum_{d=1}^D a_d \cdot Y(t-d) \quad (4.19)$$

setzen. Falls $a_D \neq 0$ gilt, heisst dieser Prozess *autoregressiver Prozess der Ordnung D*. Er wird mit AR(D) abgekürzt.

Autoregressiver Prozess 1. Ordnung, AR(1): Für $D = 1$ ist (4.19) ein autoregressiver Prozess 1. Ordnung,

$$Y(t) \equiv a \cdot Y(t-1) + X(t) \quad , \quad (4.20)$$

mit dem Parameter $a \equiv a_1$. Durch sukzessive Substitutionen können wir hierfür auch wie folgt schreiben,

$$\begin{aligned} Y(t) &= a[aY(t-2) + X(t-1)] + X(t) \\ &= a^2[aY(t-3) + X(t-2)] + aX(t-1) + X(t) \\ &= \dots \\ &= \sum_{d=0}^{+\infty} a^d X(t-d) \quad . \end{aligned} \quad (4.21)$$

Der Vergleich mit (4.17) ergibt, dass ein autoregressiver Prozess 1. Ordnung auch als Gleitmittelprozess der Ordnung unendlich, MA(+∞), aufgefasst werden kann.

Zuweilen ist es nützlich, (4.20) unter Verwendung des *Verschiebeoperators* V^d umzuschreiben,

$$V^d Y(t) \equiv Y(t-d) \quad , \quad d = 0; 1; 2; \dots \quad (4.22)$$

Dann geht (4.20) über in

$$(1 - aV)Y(t) = X(t) \quad . \quad (4.23)$$

Der Vergleich mit (4.21) liefert

$$Y(t) = \frac{1}{1 - aV} X(t) = \left(\sum_{d=0}^{+\infty} a^d V^d \right) X(t) \quad . \quad (4.24)$$

Durch einfache Rechnung erhalten wir den Mittelwert von AR(1),

$$\bar{Y} = 0 \quad ,$$

und die Varianz

$$\sigma_Y^2 = \sigma_X^2 \sum_{d=0}^{+\infty} a^{2d} \quad . \quad (4.25)$$

Die Varianz ist endlich, wenn die Reihe konvergiert, also für $a^2 < 1$ bzw. $|a| < 1$. Wir erhalten dann

$$\sigma_Y^2 = \frac{\sigma_X^2}{1 - a^2} .$$

Eine endliche Varianz, hier also die Bedingung $|a| < 1$, ist wiederum eine notwendige Bedingung dafür, dass der Prozess stationär ist. Konstruieren wir aus dem Term auf der linken Seite in (4.23) ein Polynom in V , $1 - aV = 0$, so können wir für die Stabilität des Prozesses auch fordern, dass die Wurzel $V^* = 1/a$ dieses Polynoms außerhalb des Einheitskreises ($-$ intervalls) liegt, also $V^* > 1$.

Aus (4.18) und (4.21) finden wir unter Beachtung von (4.25) die AKF von AR(1),

$$\begin{aligned} \rho(\tau) &= \left(\frac{\sigma_X}{\sigma_Y} \right)^2 \cdot \sum_{d=0}^{+\infty} a^d a^{d+\tau} \\ &= (1 - a^2) a^\tau \sum_{d=0}^{+\infty} a^{2d} \\ &= a^\tau . \end{aligned}$$

Für $|a| < 1$ fällt also der Betrag der AKF exponentiell mit τ ab. Ist a negativ, dann alterniert die AKF zwischen positiven (τ gerade) und negativen (τ ungerade) Werten.

Setzen wir in (4.20) den Parameter $a = 1$, erhalten wir eine sogenannte *Irrfahrt*. Dies ist ein nichtstationärer Prozess, mit dem Mittelwert $\overline{Y(t)} = t\overline{X}$ und der Varianz $\sigma_{Y(t)}^2 = t\sigma_X^2$. Als Irrfahrt werden zum Beispiel Aktienkurse modelliert.

Autoregressiver Prozess beliebiger Ordnung: Der autoregressive Prozess (4.19) kann auch für höhere Ordnungen, $D > 1$, auf einen Gleitmittelprozess MA(+ ∞) zurückgeführt werden, denn durch sukzessive Substitutionen können wir wie folgt schreiben,

$$\begin{aligned} Y(t) &= \sum_{d_1=1}^D a_{d_1} Y(t - d_1) + X(t) \\ &= \sum_{d_1=1}^D a_{d_1} \left[\sum_{d_2=1}^D a_{d_2} Y(t - d_1 - d_2) + X(t - d_1) \right] + X(t) \\ &= \sum_{d_1, d_2=1}^D a_{d_1} a_{d_2} Y(t - d_1 - d_2) + \sum_{d_1=1}^D a_{d_1} X(t - d_1) + X(t) \\ &= \sum_{d_1, d_2=1}^D a_{d_1} a_{d_2} \left[\sum_{d_3=1}^D a_{d_3} Y(t - d_1 - d_2 - d_3) + X(t - d_1 - d_2) \right] + \sum_{d_1=1}^D a_{d_1} X(t - d_1) + X(t) \\ &= \sum_{d_1, d_2, d_3=1}^D a_{d_1} a_{d_2} a_{d_3} Y(t - d_1 - d_2 - d_3) \\ &\quad + \sum_{d_1, d_2=1}^D a_{d_1} a_{d_2} X(t - d_1 - d_2) + \sum_{d_1=1}^D a_{d_1} X(t - d_1) + X(t) \\ &= \dots \end{aligned}$$

Für

$$\lim_{K \rightarrow \infty} \sum_{d_1, \dots, d_K=1}^D a_{d_1} \dots a_{d_K} Y(t - d_1 - \dots - d_K) = 0$$

erhalten wir also die Darstellung

$$Y(t) = X(t) + \sum_{K=1}^{+\infty} \sum_{d_1, \dots, d_K=1}^D a_{d_1} \dots a_{d_K} X(t - d_1 - \dots - d_K) . \quad (4.26)$$

Somit können wir $Y(t)$ als einen Gleitmittelprozess der Ordnung unendlich ansehen,

$$Y(t) = \sum_{d=0}^{+\infty} b_d X(t - d) . \quad (4.27)$$

Durch Koeffizientenvergleich mit (4.26) erhalten wir die Beziehungen zwischen den a_d und b_d ,

$$\begin{aligned} b_0 &= 1 \\ b_1 &= a_1 \\ b_2 &= a_1^2 + a_2 \\ b_3 &= a_1^3 + 2a_1 a_2 + a_3 \end{aligned} \quad (4.28)$$

$$\dots \quad (4.29)$$

Eine hinreichende Bedingung für die Konvergenz und somit für die Stabilität des Prozesses (4.27) ist, dass die Reihe $\sum_d |b_d|$ konvergiert.

Aus der vorausgesetzten Mittelwertfreiheit $\overline{X(t)} = 0$ folgt, dass auch der Mittelwert von AR(D) verschwindet, $\overline{Y(t)} = 0$. Für die Varianz von AR(D) erhalten wir

$$\sigma_Y^2 = \sigma_X^2 \sum_{d=1}^{+\infty} b_d^2 .$$

Die AKF lautet für $\tau \geq 0$ (vgl. (4.18)),

$$\rho(\tau) = \frac{\sigma_X^2}{\sigma_Y^2} \sum_{d=0}^{+\infty} b_d b_{d+\tau} = \frac{\sum_{d=0}^{+\infty} b_d b_{d+\tau}}{\sum_{d=1}^{+\infty} b_d^2} . \quad (4.30)$$

Um die AKF aus den Koeffizienten a_d zu berechnen, müssen wir den Zusammenhang der Koeffizienten b_d des Prozesses (4.27) mit den Regressions-Koeffizienten a_d auffinden. Dies ist im Allgemeinen algebraisch umständlich. Nimmt man jedoch an, dass der Prozess stationär ist, dann vereinfacht sich die Berechnung. Dazu multiplizieren wir beide Seiten in (4.19) mit $Y(t - \tau)/\sigma_Y^2$ und bilden dann den Mittelwert,

$$\frac{\overline{Y(t)Y(t - \tau)}}{\sigma_Y^2} \equiv \sum_{d=1}^D a_d \frac{\overline{Y(t - d)Y(t - \tau)}}{\sigma_Y^2} + \frac{\overline{X(t)Y(t - \tau)}}{\sigma_Y^2} .$$

Für $\tau > 0$ verschwindet der letzte Summand $\overline{X(t)Y(t - \tau)}/\sigma_Y^2$, denn $X(t)$ und $Y(t - \tau)$ sind unabhängig. Für den Mittelwert gilt nach Voraussetzung $\overline{X(t)} = 0$. Aus der vorausgesetzten

Stationarität und unter Beachtung der Symmetrie $\rho(\tau) = \rho(-\tau)$ folgt dann für die AKF des Prozesses (4.19) das Gleichungssystem

$$\rho(\tau) \equiv \sum_{d=1}^D a_d \rho(\tau - d) \quad , \quad \text{für alle } \tau > 0 \quad . \quad (4.31)$$

Dieses lineare System von Differenzgleichungen ist unter dem Namen *Yule-Walker-Gleichungen* bekannt. Aus ihm kann, bei gegebenen Koeffizienten a_d , die Korrelationsfunktion $\rho(\tau)$ berechnet werden.

Wir betrachten zunächst den Fall $D = 2$. Dann geht (4.31) über in

$$\begin{aligned} \rho(1) &= a_2 \rho(-1) + a_1 \rho(0) \\ \rho(2) &= a_2 \rho(0) + a_1 \rho(1) \\ \rho(3) &= a_2 \rho(1) + a_1 \rho(2) \\ \rho(4) &= a_2 \rho(2) + a_1 \rho(3) \\ &\dots \end{aligned}$$

Unter Beachtung von $\rho(-1) = \rho(1)$ und $\rho(0) = 1$ finden wir aus der ersten Gleichung

$$\rho(1) = \frac{a_1}{1 - a_2} \quad .$$

Damit gehen wir in die zweite Gleichung ein und berechnen

$$\rho(2) = a_2 + \frac{a_1^2}{1 - a_2} \quad .$$

Iterativ können wir somit auch für $\tau = 3; 4, \dots$ die AKF vom stationären Prozess AR(2) berechnen,

$$\rho(\tau) = a_1 \rho(\tau - 1) + a_2 \rho(\tau - 2) \quad .$$

Unter Verwendung des Verschiebeoperators (4.22) können wir (4.19) umschreiben

$$\begin{aligned} X(t) &= Y(t) - \sum_{d=1}^D a_d Y(t - d) \\ &= \left(1 - \sum_{d=1}^D a_d V^d \right) \cdot Y(t) \quad . \end{aligned}$$

Analog zum Fall erster Ordnung (4.24) schreiben wir auch hier wie folgt,

$$Y(t) = \frac{1}{1 - \sum_{d=1}^D a_d V^d} X(t) = \left(\sum_{d=0}^{+\infty} b_d V^d \right) X(t) \quad . \quad (4.32)$$

Das Polynom in V vom Grad D ,

$$1 - \sum_{d=1}^D a_d V^d = 0 \quad , \quad (4.33)$$

heißt *charakteristische Gleichung* des Prozesses AR(D). Man kann zeigen, dass die (im Allgemeinen komplexwertigen) Wurzeln des charakteristischen Polynoms außerhalb des Einheitskreises liegen müssen, damit der Prozess stationär ist (s. z. B. [BoJe76], S. 54).

Kapitel 5

Autoregressives Prädiktor–Modell

5.1 Lineares Modell

Gegeben sei eine Zeitreihe

$$\{x_t\}_{t=1}^T$$

zu der wir ein Prädiktor–Modell

$$\begin{aligned} \hat{x}_t &\equiv f_D(x_{t-1}, x_{t-2}, \dots, x_{t-D}) \\ f_D : \quad \mathbb{R}^D &\longrightarrow \mathbb{R} \end{aligned}$$

entwickeln wollen. Dabei soll die mittlere Energie des Prädiktions–Fehler–Signals minimal werden, also

$$\overline{(x_t - \hat{x}_t)^2} = \text{Minimum!} \quad (5.1)$$

Darin bedeutet

$$\overline{(x_t - \hat{x}_t)^2} \equiv \lim_{T \rightarrow \infty} \frac{1}{T - D} \sum_{t=1+D}^T (x_t - \hat{x}_t)^2$$

D heißt *Prädiktorordnung*.

Wir machen zunächst den *linearen* Prädiktoransatz

$$\hat{x}_t = \sum_{d=1}^D a_d x_{t-d} .$$

Setzen wir diesen in die linke Seite von (5.1) ein, so erhalten wir

$$g(a_1, \dots, a_{d^*}, \dots, a_D) \equiv \overline{\left(x_t - \sum_{d=1}^D a_d x_{t-d} \right)^2} . \quad (5.2)$$

Zur Bestimmung des Minimums der Fehler–Funktion g bilden wir alle partiellen Ableitungen nach den noch unbekanntem Prädiktorkoeffizienten a_{d^*} und setzen diese null. Wir erhalten dann das Gleichungssystem

$$\begin{aligned} 0 &= \frac{1}{2} \frac{\partial g}{\partial a_{d^*}} = \overline{\left(x_t - \sum_{d=1}^D a_d x_{t-d} \right) x_{t-d^*}} \\ &= \overline{x_t x_{t-d^*}} - \sum_{d=1}^D \overline{a_d x_{t-d} x_{t-d^*}} . \end{aligned}$$

Darin läuft d^* von 1 bis D .

Nehmen wir nun an, daß die Zeitreihe stationär und o.B.d.A. auch mittelwertfrei ist, mit der Varianz 1. Dann ist

$$\overline{x_t x_{t-d^*}} = \varrho(d^*) \quad \text{bzw.} \quad \overline{x_{t-d} x_{t-d^*}} = \varrho(|d^* - d|)$$

die Autokorrelationsfunktion an der Stelle d^* bzw. $|d^* - d|$. Damit geht das obige Gleichungssystem über in

$$\boxed{\varrho(d^*) = \sum_{d=1}^D a_d \varrho(|d^* - d|)} \quad d^* = 1, 2, \dots, D .$$

Das ist ein lineares inhomogenes Gleichungssystem für die Prädiktorkoeffizienten a_d . Ausführlich lautet es

$$\begin{pmatrix} \varrho(1) \\ \varrho(2) \\ \vdots \\ \varrho(D) \end{pmatrix} = \begin{pmatrix} 1 & \varrho(1) & \dots & \varrho(D-1) \\ \varrho(1) & 1 & \dots & \varrho(D-2) \\ & & \ddots & \\ \varrho(D-1) & \varrho(D-2) & \dots & 1 \end{pmatrix} \times \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_D \end{pmatrix} \quad (5.3)$$

Kürzer schreibt man auch

$$\boldsymbol{\varrho}^T = \mathbf{C} \mathbf{a}^T .$$

Die Autokorrelationsmatrix \mathbf{C} ist symmetrisch mit Einsen auf der Hauptdiagonalen. Diese sogenannte Toeplitz-Struktur ist für eine numerische Lösung des Gleichungssystems vorteilhaft.

Beispiel: unkorreliertes Signal

Für unkorrelierte mittelwertfreie Signale ist \mathbf{C} gleich der Einheitsmatrix und folglich $\mathbf{a} = \mathbf{o}$. Das bedeutet, im Rahmen einer linearen Vorhersage minimiert $\hat{x}_t = 0$ den mittleren quadratischen Vorhersagefehler. Wäre das Signal nicht auf den Mittelwert \bar{x} Null normiert worden, so hätten wir $\hat{x}_t = \bar{x}$ erhalten. Wird also bei unkorrelierten Signalen der Mittelwert als Vorhersagewert genommen, so minimiert dieser im Rahmen einer linearen Modellbildung den mittleren quadratischen Vorhersagefehler.

5.2 Nichtlineares Modell

Ein unkorreliertes Signal kann in bestimmtem Maße prädiktierbar sein, nämlich dann, wenn *nichtlineare* Abhängigkeiten bestehen. Diese werden mit nichtlinearen Prädiktoransätzen erfaßt, so zum Beispiel mit:

$$\hat{x}_t = \sum_{d=1}^D a_d x_{t-d} + \sum_{d_1, d_2=1}^D b_{d_1 d_2} x_{t-d_1} x_{t-d_2} .$$

Dieser *nichtlineare* Prädiktoransatz führt wiederum auf ein *lineares* Gleichungssystem für die $D + D^2$ unbekanntenen Koeffizienten a_d und $b_{d_1 d_2}$. Die Koeffizientenmatrix besteht dann aber neben den Korrelationskoeffizienten $\overline{x_{t-d} x_{t-d^*}} \equiv \varrho(|d^* - d|)$ auch aus Korrelationen höherer Ordnung, zum Beispiel

$$\overline{x_{t-d_1} x_{t-d_2} x_{t-d_1^*} x_{t-d_2^*}} .$$

5.3 The General Prediction Problem

For processes

$$\{\xi_0(t)\}_t \quad (5.4)$$

and

$$\{\boldsymbol{\xi}(t)\}_t, \quad \boldsymbol{\xi}(t) \equiv (\xi_D(t), \dots, \xi_d(t), \dots, \xi_1(t)), \quad (5.5)$$

a predictor f_τ for $\xi_0(\tau)$ is considered as a function of the observable variables (observables) $\boldsymbol{\xi}$. We suppose stationary processes, hence the predictor f_τ should depend only on the prediction step τ but not on the instant t . The prediction error

$$\delta(\tau) \equiv \xi_0(\tau) - f_\tau(\boldsymbol{\xi}) \quad (5.6)$$

causes costs $c(\delta)$. Depending on the field of application, there is a variety of possible costs functions. It is useful to give c some general properties:

1. c is not negative, $c(\delta) \geq 0$, and the exact prediction causes no costs, $c(0) = 0$.
2. A greater magnitude of the prediction error should not lead to lower costs, $c(\delta) \leq c(\delta^*)$ for $|\delta| < |\delta^*|$.

From a mathematical point of view, the prediction problem is to specify f_τ , belonging to a certain set of allowed functions, such that the costs of the prediction error (5.6) are minimized in the mean,

$$\langle c(\delta(\tau)) \rangle = \min_{f_\tau} \quad (5.7)$$

Herein the mean $\langle \cdot \rangle$ has to be taken over the (joint) probability measure $\mu_\tau(\mathbf{x}, x_0)$ of the $(D+1)$ -dimensional random vector

$$(\boldsymbol{\xi}, \xi_0(\tau)) \equiv (\xi_D, \dots, \xi_1, \xi_0(\tau)). \quad (5.8)$$

Thus the left hand side of (5.7) can be rewritten as

$$\langle c(\delta(\tau)) \rangle = \int_{\mathbb{R}^{1+D}} c(x_0 - f_\tau(\mathbf{x})) \, d\mu_\tau(\mathbf{x}, x_0). \quad (5.9)$$

If μ_τ is absolutely continuous we could also write

$$d\mu_\tau(\mathbf{x}, x_0) = s_\tau(\mathbf{x}, x_0) \, d\mathbf{x} \, dx_0,$$

where s_τ denotes the corresponding $(D+1)$ -dimensional probability density. In the following we mostly operate with such densities though most of the theory can be expressed also for singular probability measures. This is the situation, e.g., in most chaotic dynamical systems, where the measures are often fractal. However, in practice of data analysis absolutely continuous measures are a more appropriate model due to superimposed noise of any real data.

5.4 Best Least Mean Square Predictor

In (5.7), we search for a specific f_τ among all allowed predictors, for which the integral on the right hand side of (5.9) exists. Of course, before starting our search we have to specify the cost function c . Widely used is the square function,

$$c(\delta) = \delta^2 \quad . \quad (5.10)$$

With it we look in (5.7) for the so-called *least mean square* (LMS) of the prediction error. Now we can derive an explicit expression for the *best predictor* as follows: We introduce the D -dimensional marginal density

$$p(\mathbf{x}) \equiv \int_{\mathbb{R}} s_\tau(\mathbf{x}, x_0) dx_0 \quad (5.11)$$

and the conditional probability density

$$s_{\tau, \mathbf{x}}(x_0) \equiv \frac{s_\tau(\mathbf{x}, x_0)}{p(\mathbf{x})} \quad , \quad (5.12)$$

which is, for a fixed $\mathbf{x} \in \mathbb{R}^D$, a function of $x_0 \in \mathbb{R}$ alone.

We derive now a formula for the best LMS predictor. Using (5.10) we can rewrite (5.9) as follows,

$$\langle c(\delta(\tau)) \rangle = \int_{\mathbb{R}^D} g(f_\tau(\mathbf{x})) p(\mathbf{x}) \, d\mathbf{x} \quad , \quad (5.13)$$

where we have set

$$g(f_\tau(\mathbf{x})) \equiv \int_{\mathbb{R}} [x_0 - f_\tau(\mathbf{x})]^2 s_{\tau, \mathbf{x}} \, dx_0 \quad .$$

This can be rewritten as

$$g(f_\tau) = f_\tau^2 - 2f_\tau \int_{\mathbb{R}} x_0 s_{\tau, \mathbf{x}} \, dx_0 + \int_{\mathbb{R}} x_0^2 s_{\tau, \mathbf{x}} \, dx_0 \quad .$$

From $p(\mathbf{x}) \geq 0$ and $g(f_\tau(\mathbf{x})) \geq 0$ for all $\mathbf{x} \in \mathbb{R}^D$ we conclude that the minimum in (5.13) is attained if for (almost) each possible vector \mathbf{x} the term $g(f_\tau)$ attains its minimum. Then, from $dg/df_\tau|_{f_\tau=f_\tau^{\text{best}}} = 0$, and $d^2g/df_\tau^2 = 2 > 0$ we get an explicit expression for the best LMS predictor,

$$f_\tau^{\text{best}}(\mathbf{x}) = \int_{\mathbb{R}} x_0 s_{\tau, \mathbf{x}}(x_0) \, dx_0 \quad (5.14)$$

Thus we see that, in the LMS sense, the best prediction from the set of observations \mathbf{x} is the expected value of the conditional density (5.12). In general, f_τ^{best} is a non-linear function of \mathbf{x} .

A similar result can be derived if the random variables (5.4) and (5.5) are discrete in amplitude. In order to include also the case of a vanishing prediction error, which is the deterministic case, $s_{\tau, \mathbf{x}} dx_0$ in (5.14) has to be replaced by a suitable singular conditional measure $d\mu_{\tau, \mathbf{x}}$.

5.5 Nearest Neighbor Predictor

In practice we are confronted with the difficulty of computing (estimating) the required conditional probabilities and finally the corresponding expectations (5.14). Several methods have been proposed for estimating such expectations directly from the time series. (see e.g. [FAR87, CAS89, CAO99]). Motivated by chaos theory, we first construct a $(D+1)$ -dimensional phase space. Given only one time series

$$\{x_0(t)\}_{t=1}^T, \quad (5.15)$$

this can be done, e.g., by some delay-coordinates ([PAC80, TAK80, SAU91])

$$x_d(t) = x_0(t + (d-1)\vartheta), \quad d = 1, \dots, D, \quad (5.16)$$

for a time lag $\vartheta < 0$, leading to a so-called D -dimensional embedding

$$\mathbf{x}(t) = (x_0(t + (D-1)\vartheta), \dots, x_0(t + \vartheta), x_0(t)). \quad (5.17)$$

Then the prediction scheme is based on a nearest neighbors search from which the expectations are estimated. More precisely, our forecast for $\xi_0(t + \tau)$ is now

$$f_\tau(\mathbf{x}(t)) = \frac{1}{\#\theta_{\mathbf{x}(t)}} \sum_{t^* \in \theta_{\mathbf{x}(t)}} x_0(t^* + \tau). \quad (5.18)$$

$\theta_{\mathbf{x}(t)}$ denotes the set of all instants t^* such that the past $\mathbf{x}(t^*)$ is similar to the most recent past $\mathbf{x}(t)$ within uncertainty $\pm\varepsilon/2$,

$$\theta_{\mathbf{x}(t)} \equiv \{t^* : \|\mathbf{x}(t) - \mathbf{x}(t^*)\| < \varepsilon/2, t^* = (1 - (D-1)\vartheta), \dots, t - \tau\}.$$

Herein, $\|\cdot\|$ is some suitable norm, e.g. the maximum norm, and $\#\theta_{\mathbf{x}(t)}$ denotes the number of elements in the set $\theta_{\mathbf{x}(t)}$. The parameter $\varepsilon > 0$ should be on the one hand small enough to be specific to the given $\mathbf{x}(t)$, and on the other hand, ε should be large enough to avoid an empty $\theta_{\mathbf{x}(t)}$. Here, the problem is to find “good” embeddings. This means, we need “good” delay steps ϑ and a sufficiently high but not too high embedding dimension D . A method to find such an optimal delay step ϑ based on the mutual information function of the time series was proposed in [FRA86]. However, though widely used, this method does not always work well. Later in this chapter we offer an alternative when we search for delay-embeddings that are not necessarily equally spaced in time.

Kapitel 6

Nichtparametrische Schätzung von Wahrscheinlichkeitsdichten

6.1 Motivation

Aus der Wahrscheinlichkeits-Verteilung $P(x)$ einer Zufallsgröße ξ oder allgemeiner, der Verteilung $P(\mathbf{x})$ eines Zufallsvektors $\boldsymbol{\xi}$, mit

$$P(\mathbf{x}) = \text{prob}(\xi_1 \leq x_1, \dots, \xi_D \leq x_D) \text{ ,}$$

können letztlich alle Probleme der stochastischen Signalverarbeitung behandelt werden. Insofern ist die Situation ähnlich wie etwa in der statistischen Physik, wo auch die entsprechende Wahrscheinlichkeitsverteilung den Schlüssel zur thermodynamischen Beschreibung von Vielteilchensystemen bildet. In der Praxis der stochastischen Signalverarbeitung kennt man die Verteilungen jedoch in der Regel nicht. Hier hat man im Allgemeinen nur eine endliche Messreihe $\{\mathbf{x}(t)\}_{t=1}^T$ gegeben. Bei einigen Fragestellungen wie der Anpassung von AR-Prozessen (s. Kap. 5) wird das Problem der expliziten Bestimmung der relevanten Verteilungsfunktionen umgangen, indem aus den Daten zunächst die Korrelationsfunktion $\rho(\tau)$ und daraus die Regressionsparameter a_d durch Lösen von (5.3) bestimmt werden. Bei Kenntnis der relevanten zweidimensionalen Verteilungen $P(x_t, x_{t+\tau})$ des zugrunde liegenden Prozesses $\{(\xi(t), \xi(t+\tau))\}_t$, könnten die Korrelationskoeffizienten aber berechnet werden,

$$\rho(\tau) = \int_{\mathbb{R}^2} x_t x_{t+\tau} dP(x_t, x_{t+\tau}) \text{ ,}$$

und daraus letztlich auch die Regressionsparameter.

Nehmen die Daten $\mathbf{x}(t)$ Werte aus dem Kontinuum \mathbb{R}^D an, und sind die Daten verrauscht, was in der Praxis letztlich immer der Fall ist, so ist es gerechtfertigt, die Existenz einer Wahrscheinlichkeitsdichte $p(\mathbf{x}) : \mathbb{R}^D \rightarrow \mathbb{R}$ anzunehmen. Sie hängt mit der Verteilung P wie folgt zusammen,

$$P(\mathbf{x}) = P(x_1, \dots, x_D) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_D} p(\mathbf{x}^*) d\mathbf{x}^*$$

In diesem Kapitel wollen wir uns mit der Schätzung solcher Dichten aus gegebenen Zeitreihen (Messwerte, Stichprobe)

$$\{\mathbf{x}(t)\}_{t=1}^T$$

befassen.

Zunächst sei daran erinnert, dass für beliebige Wahrscheinlichkeitsdichten Folgendes gilt:

Normierung

$$\int_{\mathbb{R}^D} p(\mathbf{x}) \, d\mathbf{x} = 1$$

Berechnung der Randdichte (marginal distribution density):

$$\int_{\mathbb{R}^{n-1}} p(\mathbf{x}) \, dx_1 \dots dx_{d-1} dx_{d+1} \dots dx_n = \text{Dichte } p_d(x_i) \text{ von } \xi_d$$

6.2 Histogramm

Partitioniere den Wertebereich $[x_{1,\min}, x_{1,\max}] \times \dots \times [x_{D,\min}, x_{D,\max}]$ in paarweise disjunkte Boxen B_m . Dann ist

$$\hat{p}(\mathbf{x}) = \frac{\#\{\mathbf{x}(t) \in B_m\}}{T \cdot \lambda(B_m)} \quad \text{für } \mathbf{x} \in B_m$$

Darin ist $\lambda(B_m) = \int_{B_m} d\mathbf{x}$ das Boxvolumen. Hierin bezeichnet $\#M$ die Anzahl der Elemente (Kardinalität) der Menge M .

Im Spezialfall $D = 1$ ist $B_m =](m-1)\varepsilon, m\varepsilon[$ und wir erhalten

$$\hat{p}(x) = \frac{\#\{x(t) \in B_m\}}{T \cdot \varepsilon} \quad \text{für } x \in B_m .$$

6.3 Naiver Schätzer

$$\hat{p}(\mathbf{x}) = \frac{\#\{\|\mathbf{x}(t) - \mathbf{x}\| < \varepsilon\}}{\text{const.}}$$

Im Spezialfall $D = 1$ geht dies über in

$$\hat{p}(x) = \frac{\#\{|x(t) - x| < \varepsilon\}}{\text{const.}} = \frac{1}{T\varepsilon} \sum_{t=1}^T w\left(\frac{x(t) - x}{\varepsilon}\right) \quad (6.1)$$

mit

$$w(y) \equiv \begin{cases} 0,5 & : |y| < 1 \\ 0 & : \text{sonst.} \end{cases}$$

6.4 Kern-Schätzer

Sei $K(\mathbf{x})$, mit $K(\mathbf{x}) \geq 0$ für alle $\mathbf{x} \in \mathbb{R}^D$, eine normierte Funktion, die sogenannte *Kern-Funktion*,

$$\int_{\mathbb{R}^D} K(\mathbf{x}) \, d\mathbf{x} = 1$$

$$\hat{p}(\mathbf{x}) = \frac{1}{T\varepsilon^D} \sum_{t=1}^T K\left(\frac{\mathbf{x}^{(t)} - \mathbf{x}}{\varepsilon}\right)$$

Häufig ist K radialsymmetrisch, positiv-semidefinit und unimodal.

Beispiele:

Normalverteilung mit Varianz 1

$$K(\mathbf{x}) = \left(\frac{1}{2\pi}\right)^{D/2} \exp\left(-\frac{\mathbf{x}^\top \mathbf{x}}{2}\right)$$

Epanechnikov Kern

$$K(\mathbf{x}) = \begin{cases} \frac{D+2}{2c_D} (1 - \mathbf{x}^\top \mathbf{x}) & : \mathbf{x}^\top \mathbf{x} < 1 \\ 0 & : \text{sonst.} \end{cases}$$

Darin ist c_D das Volumen der D -dimensionalen Einheitskugel, $c_1 = 2$, $c_2 = \pi$, $c_3 = 4\pi/3$ usw., und $\mathbf{x}^\top \mathbf{x} = x_1^2 + x_2^2 + \dots + x_D^2$.

Kapitel 7

Schätzung von Zufallsgrößen

7.1 Problemstellung und grundlegende Begriffe

Gegeben sei eine Folge von T Messwerten

$$\mathbf{x}_T \equiv \{x(t)\}_{t=1}^T .$$

Sie wird in diesem Zusammenhang auch *Stichprobe* genannt. Aus den Messwerten werde nun eine neue Größe

$$y_T = f(\mathbf{x}_T)$$

nach irgend einer Vorschrift $f : \mathbb{R}^T \rightarrow \mathbb{R}$ berechnet, f heisst auch *Stichprobenfunktion* oder kurz *Statistik*. In der Praxis sind die Messwerte i. allg. mit einem Fehler behaftet, der zufälliger Natur ist. Wir fassen deshalb \mathbf{x}_T als eine Realisierung des T -dimensionalen Zufallsvektors

$$\mathbf{X}_T \equiv X(1), \dots, X(T)$$

auf. Damit ist y_T eine Realisierung der (skalaren) Zufallsgröße

$$Y_T = f(\mathbf{X}_T),$$

und es steht die Frage, wie sich die zufälligen Eigenschaften von \mathbf{X}_T auf Y_T übertragen.

7.1.1 Mittelwertschätzung

Als Beispiel betrachten wir die Schätzung eines Mittelwertes. Alle Zufallsgrößen $X(1), \dots, X(T)$ mögen den gleichen Wertebereich und die gleiche Wahrscheinlichkeitsverteilung haben. Wir schreiben deshalb $X(t) = X$ für alle $t = 1, \dots, T$. Es soll nun der *Mittelwert (Erwartungswert)*

$$\langle X \rangle$$

aus der Stichprobe \mathbf{x}_T geschätzt werden. Wir suchen also eine Stichprobenfunktion f , so dass $y_T = f(\mathbf{x}_T)$ „möglichst gut“ die gesuchte Größe $\langle X \rangle$ liefert. Später werden wir genauer definieren, was wir hier unter „möglichst gut“ verstehen. Zuvor wollen wir aber einige konkrete Stichprobenfunktion f für die beabsichtigte Mittelwertschätzung genauer angeben:

1. *Variante*: Arithmetisches Mittel,

$$f_1(\mathbf{x}_T) = \frac{1}{T} \sum_{t=1}^T x(t) \tag{7.1}$$

2. Variante: Für $T \geq 10$ definieren wir

$$f_2(\mathbf{x}_T) = \frac{1}{T} \sum_{t=1}^{10} x(t)$$

3. Variante: Für $T \geq 2$ definieren wir

$$f_3(\mathbf{x}_T) = \frac{1}{T-1} \sum_{t=1}^T x(t) \tag{7.2}$$

4. Variante:

$$f_4(\mathbf{x}_T) = 1$$

5. Variante: Geometrisches Mittel,

$$f_5(\mathbf{x}_T) = \sqrt[T]{\prod_{t=1}^T x(t)}$$

6. Variante:

$$f_6(\mathbf{x}_T) = \text{„häufigster Wert der Stichprobe“}$$

7. Variante: Wert, der in der Mitte zwischen dem minimalen und maximalen Wert der Stichprobe liegt,

$$f_7(\mathbf{x}_T) = \frac{x_{\max} + x_{\min}}{2}$$

8. Variante: Sei der Stichprobenumfang T gerade, dann definieren wir

$$f_8(\mathbf{x}_T) = \frac{2}{T} \sum_{t \text{ gerade}} x(t)$$

Wir sehen also, dass die Wahl der Stichprobenfunktion zunächst sehr willkürlich sein kann. Aber, welcher Schätzer ist am besten?

Zur weiteren Illustration wählen wir als konkretes Beispiel, das „Werfen einer Münze“. Dabei sei $X = 0$, falls „Wappen“ geworfen wurde, und $X = 1$ im Falle „Zahl“. Zum Stichprobenumfang $T = 2; 3$ sind dann folgende Ergebnisse möglich:

T	$x(1)x(2) \dots x(T)$	$f_1(\mathbf{x}_T)$
2	00	0
	01	1/2
	10	1/2
	11	1
3	000	0
	001	1/3
	010	1/3
	100	1/3
	011	2/3
	101	2/3
	110	2/3
	111	1

Wenn wir nun T -mal voneinander unabhängig die Münze werfen, dann erhalten wir eine Folge aus Nullen und Einsen, also eine Binärfolge, der Länge T . Es gibt genau 2^T verschiedene Binärfolgen. Ist p die Wahrscheinlichkeit für „Zahl=1“, und demzufolge $1-p$ die Wahrscheinlichkeit für „Wappen=0“. Dann ist die Wahrscheinlichkeit dafür, dass T -mal die 1 geworfen wurde, gleich p^T . Die Wahrscheinlichkeit, dass genau $(T-1)$ -mal die 1 geworfen wurde, ist gleich $T \cdot p^{T-1}(1-p)$, denn es gibt hierfür T viele mögliche Binärfolgen, wobei eine jede die Wahrscheinlichkeit $p^{T-1}(1-p)$ hat. Allgemein ist die Wahrscheinlichkeit, dass genau n -mal „Zahl“ geworfen wird, durch

$$P_T(n) \equiv \binom{T}{n} p^n (1-p)^{T-n} \quad (7.3)$$

gegeben. Wegen $P_T(n) > 0$, für $n = 0, 1, \dots, T$, und der Normierung

$$\sum_{n=0}^T P_T(n) = (p + (1-p))^T = 1$$

ist $\{P_T(n)\}_{n=0}^T$ in der Tat eine Wahrscheinlichkeitsverteilung, die sogenannte *Binomialverteilung*. $P_T(n)$ ist damit aber auch die Wahrscheinlichkeit, dass unser Mittelwertschätzer f_1 den Wert n/T annimmt. Die Wahrscheinlichkeit $P_T(n)$ für eine einzelne Realisierung von f_1 strebt für $T \rightarrow \infty$ und festes n gegen Null. Der Erwartungswert der Binomialverteilung ist

$$\mu_T = \sum_{n=0}^T n P_T(n) = T p ,$$

und die Varianz

$$\sigma_T^2 = \sum_{n=0}^T (n - Tp)^2 P_T(n) = Tp(1-p) .$$

Falls $n \approx \mu_T$ gilt der *lokale Grenzwertsatz von MOIVRE-LAPLACE* (s.z.B. [Gn79], S.72 ff.):

Für ein beliebiges fest gewähltes $k > 0$ und $n \in \mu_T \pm k\sigma_T$ gilt

$$\lim_{T \rightarrow \infty} \frac{\sigma_T \cdot P_T(n)}{e^{-x^2/2}/\sqrt{2\pi}} = 1 , \quad \text{mit } x \equiv \frac{n - \mu_T}{\sigma_T} . \quad (7.4)$$

Das bedeutet, die Größe $\sigma_T \cdot P_T(n)$ geht für $n \approx \mu_T$ asymptotisch in die Standard-Normalverteilung über, und es gilt hier die Näherungsformel

$$P_T(n) \approx \frac{1}{\sqrt{2\pi} \cdot \sigma_T} e^{-\frac{(n-\mu_T)^2}{2\sigma_T^2}} .$$

Aus

$$e^{-\frac{(n-\mu_T)^2}{2\sigma_T^2}} \leq 1$$

folgt für beliebiges $n \approx \mu_T$,

$$\lim_{T \rightarrow \infty} P_T(n) = 0 .$$

Unter Verwendung des lokalen Grenzwertsatzes folgt der integrale Grenzwertsatz. Danach gilt für beliebige Werte x_u und x_o , mit $-\infty \leq x_u < x_o \leq +\infty$,

$$\lim_{T \rightarrow +\infty} \text{prob}(x_u \leq x < x_o) = \frac{1}{\sqrt{2\pi}} \int_{x_u}^{x_o} e^{-\frac{z^2}{2}} dz , \quad \text{mit } x \equiv \frac{n - \mu_T}{\sigma_T} .$$

Wie gut die Näherung bei gegebenem T ist, hängt auch von der Wahrscheinlichkeit p ab. Für den fairen Münzwurf ist $p = 1/2$ und folglich $\mu_T = T/2$, $\sigma_T = \sqrt{T}/2$ sowie

$$P_T(n)|_{p=1/2} = \binom{T}{n} \frac{1}{2^T} \approx \frac{1}{\sqrt{\pi T/2}} e^{-\frac{(n-T/2)^2}{T/2}} .$$

Im Bild 7.1 ist der Quotient

$$Q_T(n) \equiv \frac{P_T(n)|_{p=1/2}}{\frac{1}{\sqrt{\pi T/2}} e^{-\frac{(n-T/2)^2}{T/2}}} \quad (7.5)$$

im Bereich $n = \mu_T \pm 3\sigma_T$ für verschiedenen Werte von T aufgetragen. Für Stichprobenlängen $T > 100$ approximiert die Normalverteilung die Binomialverteilung mit einer relativen Abweichung die kleiner als 3 % ist. Hierbei ist zu beachten, dass die Gesamtwahrscheinlichkeit in diesem Bereich konzentriert ist, denn es gilt für alle T ,

$$\sum_{n \approx \mu_T - 3\sigma_T}^{\mu_T - 3\sigma_T} P_T(n) \gtrsim 0.997 .$$

Für andere Werte von p ist jedoch die Konvergenz bei weitem nicht so „rasch“, wie im Bild für $p = 0, 2$ zu sehen ist. Für noch kleinere Werte von p verschlechtert sich das Konvergenzverhalten noch stärker, und für $p = 0$ ist diese Betrachtung vollkommen unbrauchbar. Somit stellt sich die Frage, wie die Approximation für kleine p -Werte ($p \gtrsim 0$), oder wegen der Symmetrie auch große p -Werte ($p \lesssim 1$), vorgenommen werden kann. Eine solche Formel wurde von POISSON gefunden (s.z.B. [Gn79], S. 91 ff.).

7.2 Definitionen

Ein Schätzer Y_T einer Größe a wird häufig mit einem Dach gekennzeichnet, also \hat{a} oder auch \hat{a}_T .

7.2.1 Erwartungstreuer Schätzer (unbiased estimator)

Ein Schätzer \hat{a}_T einer Größe a heißt *erwartungstreu* oder auch *unverzerrt*, wenn sein Erwartungswert gleich dem zu schätzenden wahren Wert ist,

$$\langle \hat{a} \rangle = a .$$

Die Differenz

$$\langle \hat{a} \rangle - a$$

heißt *systematischer Fehler* (engl. *bias*). Der Schätzer \hat{a}_T heißt *asymptotisch erwartungstreu*, falls¹⁾

$$\lim_{T \rightarrow \infty}^{\text{f.s.}} \langle \hat{a}_T \rangle = a .$$

¹⁾ $\lim^{\text{f.s.}}$ steht für die „fast sichere“ Konvergenz. Das bedeutet, dass zu jedem $\varepsilon > 0$ die Wahrscheinlichkeit für $|\hat{a}_T - a| < \varepsilon$ für $T \rightarrow \infty$ gegen eins strebt.

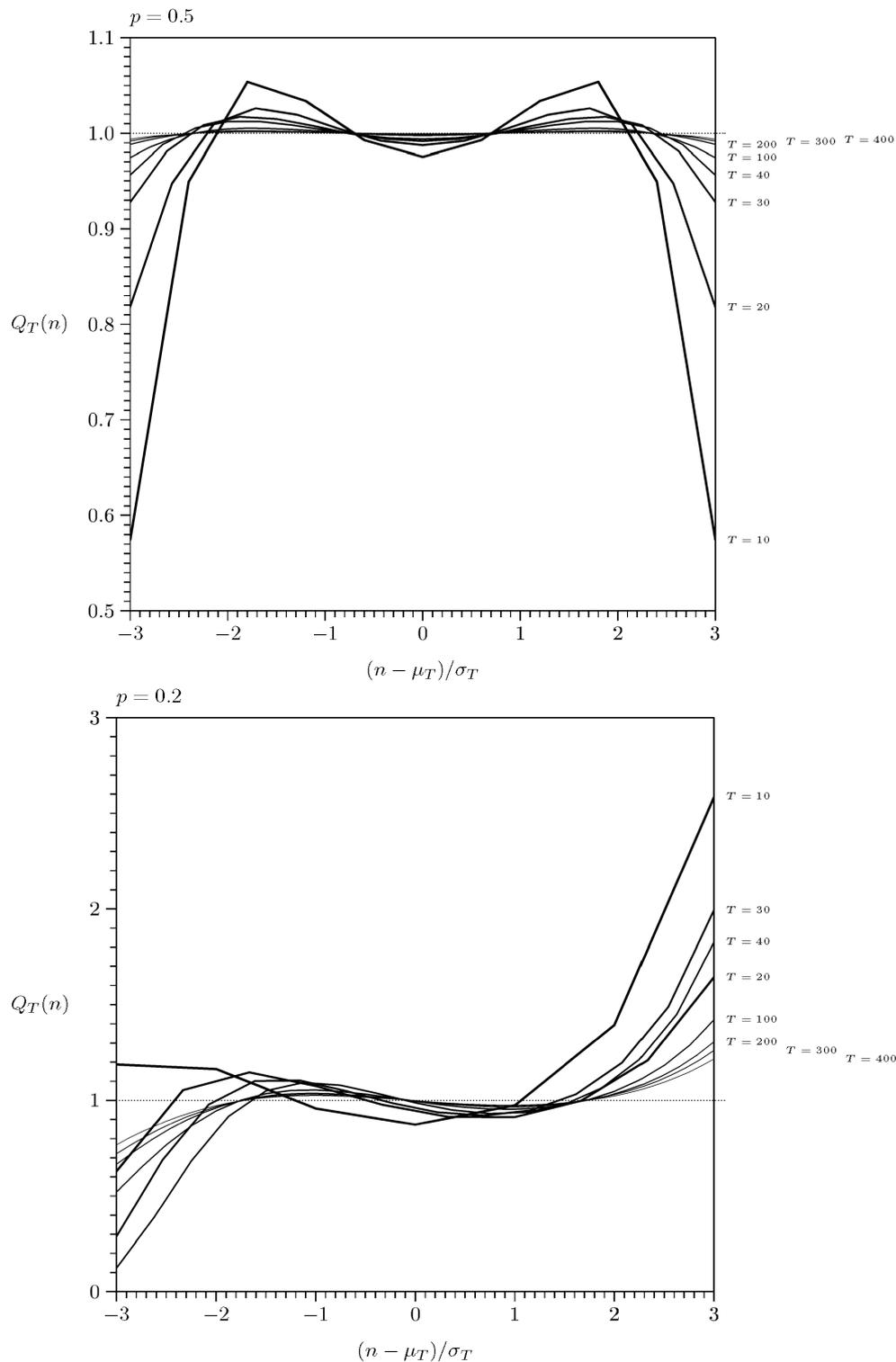


Bild 7.1: Der Quotient $Q_T(n)$ aus Gl. (7.5) im Bereich $n = \mu_T \pm 3\sigma_T$ für wachsende Stichprobenlänge T und den Wahrscheinlichkeiten $p = 0.5$ (oben) und $p = 0.2$ (unten)

Der Mittelwertschätzer f_1 in (7.1) ist erwartungstreu, denn

$$\begin{aligned} \langle \langle \widehat{X} \rangle \rangle &= \left\langle \frac{1}{T} \sum_{t=1}^T X(t) \right\rangle \\ &= \frac{1}{T} \sum_{t=1}^T \langle X \rangle \\ &= \langle X \rangle . \end{aligned}$$

Hingegen ist der Mittelwertschätzer f_3 in (7.2) nicht erwartungstreu, denn

$$\begin{aligned} \langle \langle \widehat{X} \rangle \rangle &= \frac{T}{T-1} \langle X \rangle \\ &\neq \langle X \rangle . \end{aligned}$$

7.2.2 Konsistenter Schätzer (consistent estimator)

Ein Schätzer \widehat{a}_T einer Größe a heißt (*schwach*) *konsistent*, wenn \widehat{a}_T in Wahrscheinlichkeit gegen a konvergiert,²⁾

$$\lim_{T \rightarrow \infty} \text{prob}(|\widehat{a}_T - a| > \varepsilon) = 0 , \quad \text{für beliebiges } \varepsilon > 0 .$$

Hinreichend dafür sind die Bedingungen

$$\lim_{T \rightarrow \infty} \langle \widehat{a}_T \rangle = a \quad \text{und} \quad \lim_{T \rightarrow \infty} \langle (\widehat{a}_T - a)^2 \rangle = 0 .$$

Beispielsweise ist der Mittelwertschätzer f_1 in (7.1) konsistent, denn

$$\lim_{T \rightarrow \infty} \left\langle \left(\frac{1}{T} \sum_{t=1}^T X(t) - \langle X \rangle \right)^2 \right\rangle = 0 .$$

Die Schätzer in den Varianten 1, 3 und 8 sind konsistent und die von den Varianten 2 und 4 sind offenbar inkonsistent.

Ein Schätzer \widehat{a}_T einer Größe a heißt *stark konsistent*, wenn \widehat{a}_T fast sicher gegen a konvergiert,³⁾

$$\lim_{T \rightarrow \infty}^{\text{f.s.}} \widehat{a}_T = a .$$

Beispielsweise ist der Standard-Mittelwertschätzer $\widehat{a}_T = T^{-1} \sum_{t=1}^T x_t$ stark konsistent (Gesetz der großen Zahlen).

Aus der starken Konsistenz folgt die schwache.⁴⁾

²⁾ Eine Folge X_1, X_2, \dots von Zufallsgrößen heißt *konvergent in Wahrscheinlichkeit* gegen die Zufallsgröße Y , wenn $\lim_{n \rightarrow \infty} \text{prob}(|X_n - Y| > \varepsilon) = 0$ für jedes $\varepsilon > 0$ gilt.

³⁾ Eine Folge X_1, X_2, \dots von Zufallsgrößen heißt *fast sicher konvergent* oder *konvergent mit Wahrscheinlichkeit Eins* gegen die Zufallsgröße Y , wenn $\text{prob}(\lim_{n \rightarrow \infty} X_n = Y) = 1$. Wir schreiben dafür auch $\lim_{n \rightarrow \infty}^{\text{f.s.}} X_n = Y$.

⁴⁾ Allgemein folgt aus der fast sicheren Konvergenz einer Zufallsfolge die Konvergenz in Wahrscheinlichkeit.

7.2.3 Effizienter Schätzer (efficient estimator)

Gegeben seien zwei Schätzer \hat{a}_T und \hat{a}_T^* einer Größe a . Wir betrachten die Varianzen

$$\sigma^2 \equiv \langle (\hat{a}_T - \langle \hat{a}_T \rangle)^2 \rangle$$

und

$$\sigma^{*2} \equiv \langle (\hat{a}_T^* - \langle \hat{a}_T^* \rangle)^2 \rangle$$

dieser Schätzer. Dann heißt \hat{a}_T *effizienter* als \hat{a}_T^* , falls

$$\sigma^2 < \sigma^{*2} .$$

Den Schätzer mit der kleinsten Varianz nennt man auch schlechthin *effizient*.

Beispiel: Der Schätzer von Variante 1 ist effizienter als der von Variante 8 (beide sind konsistent und erwartungstreu). Es gibt keinen effizienteren Mittelwert-Schätzer als den von Variante 1. Dieser Schätzer ist also effizient.

Im allgemeinen gibt es keinen „besten“ Schätzer. Die Auswahl hängt von der konkreten Anwendung ab. Die Effizienz eines Schätzers ändert sich i. allg. mit dem Anwendungsgebiet. Der effizienteste Schätzer muss nicht auch erwartungstreu sein. Beispielsweise kann man immer einen trivialen Schätzer benennen, der einfach konstant ist. Dieser hat die Varianz Null, würde aber i.allg. nicht erwartungstreu sein, es sei denn, diese Konstante ist die zu schätzende Größe.

7.3 Zentraler Grenzwertsatz

Seien X_t , $t = 1, 2, \dots, T$, unabhängige Zufallsgrößen mit den Mittelwerten $\langle X_t \rangle$ und den Varianzen $\sigma_t^2 < +\infty$. Dann hat die Zufallsgröße

$$Y_T \equiv \sum_{t=1}^T X_t$$

den Mittelwert

$$\langle Y_T \rangle = \sum_{t=1}^T \langle X_t \rangle$$

und die Varianz

$$\langle (Y_T - \langle Y_T \rangle)^2 \rangle = \sum_{t=1}^T \sigma_t^2 .$$

Die Verteilung P_{Y_T} von Y_T strebt für $T \rightarrow \infty$ gegen die Normalverteilung.

Die ersten beiden Aussagen lassen sich leicht zeigen: Für den Mittelwert von Y_T erhält man auch ohne der Annahme der Unabhängigkeit,

$$\langle Y_T \rangle = \left\langle \sum_{t=1}^T X_t \right\rangle = \sum_{t=1}^T \langle X_t \rangle .$$

Für die Varianz von Y_T gilt

$$\langle [Y_T - \langle Y_T \rangle]^2 \rangle = \left\langle \left[\sum_t X_t - \sum_t \langle X_t \rangle \right]^2 \right\rangle$$

$$\begin{aligned}
 &= \left\langle \left[\sum_t (X_t - \langle X_t \rangle) \right]^2 \right\rangle \\
 &= \sum_t \langle (X_t - \langle X_t \rangle)^2 \rangle + \sum_t \sum_{t^* \neq t} \langle (X_t - \langle X_t \rangle)(X_{t^*} - \langle X_{t^*} \rangle) \rangle \\
 &= \sum_t \langle (X_t - \langle X_t \rangle)^2 \rangle
 \end{aligned}$$

Zum Beweis der zentralen Aussage des Satzes, dass nämlich P_{Y_T} asymptotisch normalverteilt ist, verweisen wir auf [REN70], S. 363.

Der lokale Grenzwertsatz von MOIVRE–LAPLACE (7.4) ist ein Spezialfall des zentralen Grenzwertsatzes.

7.4 Grundlegende Schätzer

Gegeben seien T unkorrelierte Realisierungen x_1, x_2, \dots, x_T der Zufallsgröße X .

zu schätzende Größe	Schätzer Y	Merkmale
Mittelwert		
$\langle X \rangle$	$\bar{x} \equiv \frac{1}{T} \sum_t x_t$	konsistent, erwartungstreu
1. Varianz–Schätzer		
$\langle (X - \langle X \rangle)^2 \rangle$	$\frac{1}{T} \sum_t (x_t - \langle X \rangle)^2$	konsistent, erwartungstreu
2. Varianz–Schätzer		
$\langle (X - \langle X \rangle)^2 \rangle$	$\frac{1}{T} \sum_t (x_t - \bar{x})^2$	konsistent, <i>nicht</i> erwartungstreu
3. Varianz–Schätzer		
$\langle (X - \langle X \rangle)^2 \rangle$	$\frac{1}{T-1} \sum_t (x_t - \bar{x})^2$	konsistent, erwartungstreu

Mittelwert des 2. Varianz–Schätzers ($Y_T \equiv \frac{1}{T} \sum_{t^*} X_{t^*}$)

$$\begin{aligned}
 \left\langle \frac{1}{T} \sum_t (X_t - Y_T)^2 \right\rangle &= \left\langle \frac{1}{T} \sum_t [(X_t - \langle X \rangle) - (Y_T - \langle X \rangle)]^2 \right\rangle \\
 &= \langle [(X_t - \langle X \rangle) - (Y_T - \langle X \rangle)]^2 \rangle \\
 &= \langle (X_t - \langle X \rangle)^2 \rangle - 2 \cdot \langle (X_t - \langle X \rangle)(Y_T - \langle X \rangle) \rangle + \langle (Y_T - \langle X \rangle)^2 \rangle
 \end{aligned}$$

Darin ist

$$\langle (X_t - \langle X \rangle)^2 \rangle = \sigma_X^2$$

die Varianz von X_t . Wegen der vorausgesetzten Unkorreliertheit der Messwerte gilt

$$\langle (X_t - \langle X \rangle)(X_{t^*} - \langle X \rangle) \rangle = 0 \quad \text{falls } t \neq t^*$$

und somit

$$\begin{aligned}
 \langle (X_t - \langle X \rangle) \cdot (Y_T - \langle X \rangle) \rangle &= \left\langle (X_t - \langle X \rangle) \cdot \left(\frac{1}{T} \sum_{t^*} (X_{t^*} - \langle X \rangle) \right) \right\rangle \\
 &= \left\langle \frac{1}{T} (X_t - \langle X \rangle)^2 \right\rangle \\
 &= \sigma_X^2 / T .
 \end{aligned}$$

Mit

$$\langle (Y_T - \langle X \rangle)^2 \rangle = \sigma_X^2 / T$$

erhalten wir schließlich

$$\left\langle \frac{1}{T} \sum_t (X_t - Y_T)^2 \right\rangle = \left(1 - \frac{2}{T} + \frac{1}{T} \right) \sigma_X^2 = \frac{T-1}{T} \sigma_X^2 \neq \sigma_X^2 .$$

Der Faktor $\frac{T}{T-1}$ heißt *Besselsche Korrektur*. Mit ihr geht der 2. Varianz-Schätzer in den *erwartungstreuen* 3. Varianz-Schätzer über.

7.5 Maximum Likelihood Schätzung (MLS)

Sei

$$\mathbf{x}_T \equiv \{x(t)\}_{t=1}^T$$

eine Folge von Messwerten (Stichprobe), die wir als Folge von Realisierungen *unabhängiger* Zufallsgrößen

$$\mathbf{X}_T \equiv \{X(t)\}_{t=1}^T$$

auffassen. Sind die Zufallsgrößen $X(t)$ diskret, so sei $p(x(t))$ die Wahrscheinlichkeit für $X(t) = x(t)$. Sind die Zufallsgrößen $X(t)$ kontinuierlich, so interpretieren wir $p(x(t))$ als Wahrscheinlichkeitsdichte von $X(t)$ an der Stelle $x(t)$.

Wir betrachten dann die sogenannte *Likelihood-Funktion*

$$L(\mathbf{x}_T) = \prod_{t=1}^T p(x(t))$$

Im diskreten Fall ist dies die Wahrscheinlichkeit für die Stichprobe \mathbf{x}_T . Im kontinuierlichen Fall ist $L(\mathbf{x}) d\mathbf{x}$ die Wahrscheinlichkeit für $\mathbf{x}_T \in [\mathbf{x}, \mathbf{x} + d\mathbf{x}]$.

Sei nun a ein Parameter, den wir aus der Stichprobe \mathbf{x}_T schätzen wollen. Wir suchen ihn im Bereich der reellen Zahlen \mathbb{R} , im Allgemeinen kann dies aber auch nur ein Teilbereich von \mathbb{R} sein. Die Wahrscheinlichkeit bzw. Wahrscheinlichkeitsdichte $p(x(t))$ möge von diesem Parameter abhängen. Wir schreiben deshalb $p(x(t); a)$. Folglich hängt dann auch die Likelihood-Funktion von a ab, also

$$L(\mathbf{x}_T; a) = \prod_{t=1}^T p(x(t); a) .$$

Die *Maximum Likelihood Schätzung (MLS)* \hat{a} von a ist dann der Wert, für den $L(\mathbf{x}_T; a)$ maximal wird,

$$L(\mathbf{x}_T; a) \leq L(\mathbf{x}_T; \hat{a}) \quad \text{für alle } a \in \mathbb{R} .$$

Für die MLS \hat{a} des Parameters a hat also die gemessene Stichprobe \mathbf{x}_T die höchste Realisierungswahrscheinlichkeit.

Anstelle der Funktion $L(\mathbf{x}_T; a)$ können wir auch die sogenannte *Log-Likelihood-Funktion*

$$\ln L(\mathbf{x}_T; a) = \sum_{t=1}^T \ln p(x(t); a)$$

betrachten. Offenbar nimmt sie ihr Extremum an der gleichen Stelle wie $L(\mathbf{x}_T; a)$ an. Die Vorgehensweise bringt rechnerische Vorteile. Das Maximum \hat{a} erhalten wir dann aus der Gleichung

$$\frac{d}{da} \ln L(\mathbf{x}_T; a) = \sum_{t=1}^T \frac{d}{da} \ln p(x(t); a) = 0 .$$

Die MLS von mehreren zu schätzenden Parametern

$$a_1, \dots, a_n, \dots, a_N$$

verläuft vollkommen analog. Die Log-Likelihood-Funktion hängt dann von all diesen Parametern ab, und die MLS dieser Parameter ist die Lösung des Gleichungssystems

$$\frac{\partial \ln L(\mathbf{x}_T; a_1, \dots, a_N)}{\partial a_n} = 0 , \quad n = 1, 2, \dots, N .$$

Besonders interessant ist nun die folgende Tatsache: **Für große Stichprobenlängen T ist der MLS immer erwartungstreu, effizient und normalverteilt.** Es gibt dann also keinen Schätzer der Größe a mit kleinerer Varianz. In diesen Sinne gibt es also für $T \rightarrow \infty$ keinen „besseren“ Schätzer als den MLS, und seine asymptotische Verteilung ist bekannt (Normalverteilung).

Für endliche Stichprobenlängen $T < \infty$ kann es aber einen „besseren“ Schätzer geben. Insbesondere sind hier MLS in der Regel nicht erwartungstreu. Darüber hinaus muß für eine MLS die „richtige“ Familie $p(x(t); a)$ von Wahrscheinlichkeitsverteilungen als bekannt vorausgesetzt werden. In der Praxis hilft dabei oft der zentrale Grenzwertsatz.

7.5.1 Beispiel: MLS des Mittelwertes bei ungleicher Messgenauigkeit

Sei \mathbf{x}_T eine Folge von Messwerten einer bestimmten Größe, die mit unterschiedlichen Genauigkeiten gemessen wird. Ein jeder Messwert sei normalverteilt mit der Standardabweichung σ_t und Mittelwert $a \equiv \langle X(t) \rangle$. Dann ist die Wahrscheinlichkeitsdichte durch

$$p(x(t); a) = \frac{1}{\sigma_t \sqrt{2\pi}} \exp \frac{-(x - a)^2}{2\sigma_t^2}$$

gegeben, und für die Log-Likelihood-Funktion folgt

$$\begin{aligned} \ln L(\mathbf{x}_T; a) &= \sum_{t=1}^T \ln p(x(t); a) \\ &= \sum_{t=1}^T \left(-\ln \sigma_t \sqrt{2\pi} - \frac{[x(t) - a]^2}{2\sigma_t^2} \right) . \end{aligned}$$

Die Ableitung nach a ergibt

$$\left. \frac{d \ln L(\mathbf{x}_T; a)}{da} \right|_{a=\hat{a}} = \sum_{t=1}^T \frac{x(t) - \hat{a}}{\sigma_t^2} = 0 . \quad (7.6)$$

Wegen

$$\left. \frac{d^2 \ln L(\mathbf{x}_T; a)}{da^2} \right|_{a=\hat{a}} = - \sum_{t=1}^T \sigma_t^{-2} < 0$$

ist \hat{a} tatsächlich ein (relatives) Maximum. Die Maximum Likelihood Schätzung \hat{a} des Mittelwertes a folgt dann aus (7.6),

$$\hat{a} = \sum_{t=1}^T \left(\frac{\sigma_t^{-2}}{\sum_{t^*=1}^T \sigma_{t^*}^{-2}} \right) \cdot x(t) .$$

Sind alle Messgenauigkeiten gleich ($\sigma_1 = \sigma_2 = \dots = \sigma_T \equiv \sigma$), so erhalten wir für die Gewichtungsfaktoren

$$\frac{\sigma_t^{-2}}{\sum_{t^*=1}^T \sigma_{t^*}^{-2}} = \frac{\sigma^{-2}}{\sum_{t^*=1}^T \sigma^{-2}} = \frac{1}{T} ,$$

und der ML-Mittelwertschätzer geht in die hinlänglich bekannte Form über:

$$\hat{a} = \frac{1}{T} \sum_{t=1}^T x(t) .$$

7.6 Maximum Entropie Schätzung (MES)

Die Maximum-Entropie-Schätzung (MES)⁵⁾ ist ein universelles Verfahren, Wahrscheinlichkeits-Verteilungs-Dichten p unter bestimmten Nebenbedingungen zu schätzen. Die Methode geht davon aus, daß die gesuchte Dichte die Entropie

$$H(p) = - \int p(x) \ln p(x) dx , \quad (7.7)$$

maximiert, bei Einhaltung gewisser Nebenbedingungen.

Ihre Motivation findet die MES im zweiten Hauptsatz der Wärmelehre: Bezeichne $p(\mathbf{x}, \mathbf{v}, t)$ in der kinetischen Gastheorie die Einteilchen-Verteilungsdichte im Orts-Geschwindigkeits-Raum zur Zeit t . Das System sei abgeschlossen. Dann genügt p dem BOLTZMANNschen H -Theorem, wonach der Ausdruck

$$H_t \equiv -k \int p(\mathbf{x}, \mathbf{v}, t) \ln p(\mathbf{x}, \mathbf{v}, t) d\mathbf{x} d\mathbf{v}$$

bei voranschreitender Zeit t nicht abnehmen kann. Hierin ist $k = 1,38 \times 10^{-23}$ Ws/K die Boltzmann-Konstante. Im thermodynamischen Gleichgewicht nimmt H_t seinen Maximalwert an. Die Verteilungsdichte $p(\mathbf{x}, \mathbf{v}, t)$ ergibt sich als Lösung der Boltzmann-Gleichung.

7.6.1 Beispiel: MES bei gegebenen Erwartungswerten — ein-dimensionaler Fall

Sei X eine Zufallsgröße, welche die Werte $x \in \mathbb{R}$ annehme, und $p(x)$ bezeichne die zugehörige Verteilungsdichte. Gegeben seien die Erwartungswerte

$$\overline{f_n} \equiv \int f_n(x) p(x) dx \quad n = 1, 2, \dots, N. \quad (7.8)$$

⁵⁾Im Englischen: *Maximum Entropy Method* oder *Maximum Entropy Principle*.

Dabei setzen wir voraus, daß die Funktionen $f_n(x)$ bekannt sind und die Erwartungswerte existieren. Gesucht ist nun die unbekannte Dichte $p(x)$. Diese soll die Nebenbedingungen (7.8) erfüllen und nach dem MES gleichzeitig die Entropie (7.7) maximieren. Darüber hinaus soll die Dichte normiert sein,

$$1 = \int p(x) dx . \quad (7.9)$$

Wir machen für die Dichte den Ansatz

$$p(x) = A \cdot \exp[-\lambda_1 f_1(x) - \dots - \lambda_N f_N(x)] . \quad (7.10)$$

Dies ist bereits die gesuchte Dichte, was die folgende Überlegung zeigt: Die Dichte (7.10) liefert die Entropie

$$\begin{aligned} H(p) &= - \int p(x) \ln p(x) dx \\ &= - \int p(x) [\ln A - \lambda_1 f_1(x) - \dots - \lambda_N f_N(x)] dx \\ &= - \ln A + \lambda_1 \overline{f_1} + \dots + \lambda_N \overline{f_N} . \end{aligned} \quad (7.11)$$

Sei nun p^* eine möglicherweise von p verschiedene Dichte. Aus $\ln y \leq y - 1$ folgt

$$\int p^*(x) \ln \frac{p(x)}{p^*(x)} dx \leq \int p^*(x) \left(\frac{p(x)}{p^*(x)} - 1 \right) dx = \int p(x) - p^*(x) dx = 1 - 1 = 0$$

und somit

$$H(p^*) \equiv - \int p^*(x) \ln p^*(x) dx \leq - \int p^*(x) \ln p(x) dx . \quad (7.12)$$

Wenn wir nun fordern, daß p^* die gleichen Mittelwerte wie p liefert, dass also

$$\overline{f_n} = \int_{\mathbb{R}} f_n(x) p^*(x) dx$$

gelte, dann erhalten wir

$$\begin{aligned} - \int p^*(x) \ln p(x) dx &= - \int p^*(x) [\ln A - \lambda_1 f_1(x) - \dots - \lambda_N f_N(x)] dx \\ &= - \ln A + \lambda_1 \overline{f_1} + \dots + \lambda_N \overline{f_N} \\ &= H(p) . \end{aligned}$$

Unter Beachtung von (7.12) folgt somit die Behauptung

$$H(p^*) \leq H(p) .$$

Wir betrachten nun einige Beispiele:

Vorgegebener Mittelwert: Die Zufallsgröße möge nur positive Werte annehmen und der Mittelwert sei vorgegeben,

$$\bar{x} \equiv \int_0^{\infty} x p(x) dx . \quad (7.13)$$

Dann ist nach der MES die zugehörige Dichte (7.10)

$$p(x) = Ae^{-\lambda x} .$$

Die Konstanten ergeben sich aus den Bedingungen (7.13) und (7.9). Damit erhalten wir schließlich die Exponentialverteilung

$$p(x) = \frac{e^{-x/\bar{x}}}{\bar{x}} .$$

Vorgegebene Varianz: Die Zufallsgröße möge nun wieder beliebige Werte von $-\infty \dots +\infty$ annehmen. Fest vorgegeben sei die Varianz (technisch z.B. die „Leistung“ oder „Energie“)

$$\overline{x^2} \equiv \int x^2 p(x) dx . \quad (7.14)$$

Dann ist nach der MES die zugehörige Dichte (7.10)

$$p(x) = Ae^{-\lambda x^2} .$$

Die Konstanten ergeben sich aus den Bedingungen (7.14) und (7.9). Damit erhalten wir schließlich die Normalverteilung

$$p(x) = \frac{1}{\sqrt{2\pi x^2}} e^{-\frac{x^2}{2x^2}} .$$

7.6.2 MES bei gegebenen Erwartungswerten — mehr-dimensionaler Fall

Ist \mathbf{X} ein höher-dimensionaler (M -dimensionaler) Zufallsvektor, so können wir vollkommen analog vorgehen. Anstelle von (7.8) und (7.9) betrachten wir nun die Nebenbedingungen

$$\overline{f_n} \equiv \int f_n(\mathbf{x}) p(\mathbf{x}) d\mathbf{x} \quad n = 1, 2, \dots, N \quad (7.15)$$

sowie die Normierungsbedingung

$$1 = \int p(\mathbf{x}) d\mathbf{x} . \quad (7.16)$$

Eine Argumentation wie im ein-dimensionalen Fall führt auf die gesuchte Verteilungsdichte

$$p(\mathbf{x}) = A \exp[-\lambda_1 f_1(\mathbf{x}) - \dots - \lambda_N f_N(\mathbf{x})] . \quad (7.17)$$

Betrachten wir nun als Beispiel den Fall, daß die Korrelationsmatrix

$$r_{ij} \equiv \overline{x_i x_j} = \int x_i x_j p(\mathbf{x}) d\mathbf{x} \quad i, j = 1, 2, \dots, M$$

als Nebenbedingung gegeben ist. Hierbei setzen wir zur Vereinfachung der Schreibweise voraus, daß die Komponenten X_1, \dots, X_M des Zufallsvektors \mathbf{X} normiert sind auf Mittelwert null und Varianz eins. Die MES liefert dann die Normalverteilung

$$p(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^M \Delta}} \exp\left(-\frac{1}{2} \mathbf{x} R^{-1} \mathbf{x}^T\right) .$$

Hierin sind

$$\Delta \equiv \det R$$

die Determinante der Korrelationsmatrix R , welche ungleich null sei. Mit R^{-1} bezeichnen wir die Inverse von R .

Im zwei-dimensionalen Fall gilt

$$R = \begin{pmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{pmatrix} \equiv \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} ,$$

mit dem Korrelationskoeffizienten $\rho \equiv r_{12} = r_{21}$. Folglich haben wir

$$\Delta = 1 - \rho^2$$

und

$$R^{-1} = \frac{1}{\Delta} \begin{pmatrix} 1 & -\rho \\ -\rho & 1 \end{pmatrix} .$$

Damit lautet die gesuchte Dichte

$$p(x_1, x_2) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp \left[-\frac{x_1^2 - 2\rho x_1 x_2 + x_2^2}{2(1-\rho^2)} \right] .$$

7.7 Spektralschätzung

Im Kapitel 1 haben wir die Spektraldichte (1.29) einer Abtastfolge eines Signals und darauf aufbauend die diskrete Fouriertransformierte (1.41) einer Abtastfolge eingeführt. Hier fassen wir die Abtastfolge als endliche diskrete Realisierung eines stochastischen Prozesses $\{X_t\}_t$ auf. Der Prozess sei stationär und mittelwertfrei. Die Autokorrelationsfunktion (genauer: die Autokovarianzfunktion) ist dann

$$p(\tau) \equiv \lim_{T \rightarrow \infty} \int_{-\infty}^{+\infty} x(t)x(t+\tau) dt ,$$

und für das Power Spektrum gilt nach (4.12)

$$P(j\omega) = \int_{-\infty}^{+\infty} p(\tau)e^{-j\omega\tau} d\tau .$$

Es stellt sich nun die Frage, wie $P(j\omega)$ aus einer diskreten Stichprobe $\{x_m\}_{m=0}^{N-1}$ oder aus einer endlichen zeitkontinuierlichen Realisierung $\{x(t)\}_{t \in [0, T]}$ geschätzt werden kann. Für dieses Problem gibt es eine Fülle von Verfahren, darunter das „klassische“ Periodogramm.

Wir gehen von einem stationären stochastischen Prozess $\{X_t\}_t$ aus, zu dem eine endliche zeitdiskrete Realisierung $\{x_m\}_{m=0}^{N-1}$ gegeben sei. Das *Periodogramm* ist dann definiert durch

$$P_N(n) \equiv \frac{1}{2\pi N} \left| \sum_{m=0}^{N-1} x_m e^{-j2\pi mn/N} \right|^2 . \quad (7.18)$$

Die entsprechende Zufallsgröße

$$\hat{P}_N(n) \equiv \frac{1}{2\pi N} \left| \sum_{m=0}^{N-1} X_m e^{-j2\pi mn/N} \right|^2 , \quad (7.19)$$

fassen wir als zeitdiskreten Periodogramm–Spektral–Schätzer auf.

Im zeitkontinuierlichen Fall ist das *Periodogramm* durch

$$P_T(j\omega) \equiv \frac{1}{2\pi T} \left| \int_0^T x(t) e^{-j\omega t} dt \right|^2 \quad (7.20)$$

definiert, mit dem zugehörigen zeitkontinuierlichen Spektral–Schätzer

$$\hat{P}_T(j\omega) \equiv \frac{1}{2\pi T} \left| \int_0^T X(t) e^{-j\omega t} dt \right|^2 . \quad (7.21)$$

Periodogramme erscheinen als natürliche Art der Spektral–Schätzung. In der Tat sind beide Schätzer (7.19) und (7.21) asymptotisch (für N bzw. $T \rightarrow \infty$) erwartungstreue Punktschätzer der Spektraldichte $P(j\omega)$. Ist das Power Spektrum $P(j\omega)$ diskret, so dass es nur an den Stellen ω_k Masse besitzt, so konvergiert die Folge $\{\hat{P}_N(m)\}_{N=1,2,\dots}$ ($N \rightarrow \infty$) an den Stellen ω_k fast sicher gegen unendlich und gegen Null für $\omega \neq \omega_k$.

Die Spektral–Schätzer sind allerdings *nicht konsistent*. Wir zeigen dies hier nur für den zeitdiskreten Fall, in dem $\{x_m\}_{m=0}^{N-1}$ eine Realisierung eines normalverteilten i.i.d. Prozesses mit Mittelwert μ und Varianz σ^2 ist. Dazu schreiben wir (7.19) um.

Wir wollen nun den Mittelwert und die Varianz von $\hat{P}_N(n)$ in (7.19) bestimmen. Dazu schreiben wir $\hat{P}_N(n)$ zunächst um: Unter Benutzung der Eulerschen Formel, $e^{j\phi} = \cos \phi + j \sin \phi$, und mit den Abkürzungen

$$\alpha_n \equiv \sum_{m=0}^{N-1} X_m \cos(2\pi mn/N) \quad \text{und} \quad \beta_n \equiv \sum_{m=0}^{N-1} X_m \sin(2\pi mn/N)$$

erhalten wir

$$\hat{P}_N(n) \equiv \frac{\alpha_n^2 + \beta_n^2}{2\pi N} .$$

Wir charakterisieren nun zunächst α_n und β_n : Beide sind normalverteilt, da sie Linearkombinationen normalverteilter Zufallsgrößen X_m sind. Die Mittelwerte verschwinden, denn

$$\overline{\alpha_n} = \sum_{m=0}^{N-1} \overline{X_m} \cos(2\pi mn/N) = \mu \sum_{m=0}^{N-1} \cos(2\pi mn/N) = 0 .$$

Analog erhalten wir

$$\overline{\beta_n} = 0 .$$

Für die Varianz von α_n erhalten wir

$$\sigma_{\alpha_n}^2 = \overline{(\alpha_n - \overline{\alpha_n})^2} = \overline{\left(\sum_{m=0}^{N-1} (X_m - \mu) \cos(2\pi mn/N) \right)^2}$$

Die vorausgesetzte Unabhängigkeit der $\{X_m\}$ zieht die Unkorreliertheit $\overline{(X_m - \mu)(X_{m^*} - \mu)} = 0$, für $m \neq m^*$, nach sich. Damit gilt

$$\sigma_{\alpha_n}^2 = \sum_{m=0}^{N-1} \overline{(X_m - \mu)^2} \cos^2(2\pi mn/N) = \sigma^2 \sum_{m=0}^{N-1} \cos^2(2\pi mn/N) = \begin{cases} \sigma^2 N & : n = N/2 \text{ bzw. } 0 \\ \sigma^2 N/2 & : n \neq N/2 \text{ bzw. } 0 \end{cases} .$$

Analog erhält man

$$\sigma_{\beta_n}^2 = \begin{cases} 0 & : n = N/2 \text{ bzw. } 0 \\ \sigma^2 N/2 & : n \neq N/2 \text{ bzw. } 0 \end{cases} .$$

α_n und β_n sind unkorreliert, denn

$$\begin{aligned} \overline{\alpha_n \beta_n} &= \overline{\left(\sum_{m=0}^{N-1} (X_m - \mu) \cos(2\pi mn/N) \times \sum_{m^*=0}^{N-1} (X_{m^*} - \mu) \sin(2\pi mn/N) \right)} \\ &= \sum_{m=0}^{N-1} \overline{(X_m - \mu)^2} \cos(2\pi mn/N) \sin(2\pi m^* n/N) \\ &= \sigma^2 \sum_{m=0}^{N-1} \cos(2\pi mn/N) \sin(2\pi mn/N) = 0 . \end{aligned}$$

Da (α_n, β_n) bivariat normalverteilt ist, folgt aus der Unkorreliertheit auch die Unabhängigkeit von α_n und β_n .

Folglich sind die Zufallsgrößen $\sqrt{2}\alpha_n/(\sigma\sqrt{N})$ und $\sqrt{2}\beta_n/(\sigma\sqrt{N})$, für $n \neq N/2$ bzw. 0, bzw. $\alpha_n/(\sigma\sqrt{N})$ für $n = N/2$ bzw. 0, standard-normalverteilt, d.h., mittelwertfrei mit Varianz 1.

Folglich ist für $n \neq N/2$ bzw. 0 die Zufallsgröße

$$\frac{4\pi}{\sigma^2} \widehat{P}_N(n) = \frac{2\alpha_n^2}{\sigma^2 N} + \frac{2\beta_n^2}{\sigma^2 N}$$

χ^2 -verteilt mit zwei Freiheitsgraden, sie hat somit den Mittelwert 2 und die Varianz 4. Für $n = N/2$ bzw. 0 und beliebiges $c \in \mathbb{R}$ ist

$$\frac{2\pi}{\sigma^2} \widehat{P}_N(n) = \frac{\alpha_n^2}{\sigma^2 N} + c\beta_n^2$$

χ^2 -verteilt mit einem Freiheitsgrad, sie hat somit den Mittelwert 1 und die Varianz 2.

Schließlich erhalten wir für die gesuchte Varianz von $\widehat{P}_N(n)$,

$$\sigma_{\widehat{P}_N(n)}^2 = \begin{cases} \left(\frac{\sigma^2}{4\pi}\right)^2 \times 4 = \frac{\sigma^4}{4\pi^2} & : n \neq N/2 \text{ bzw. } 0 \\ \left(\frac{\sigma^2}{2\pi}\right)^2 \times 2 = \frac{\sigma^4}{2\pi^2} & : n = N/2 \text{ bzw. } 0 \end{cases}$$

Die Varianz unseres Periodogramm-Spektral-Schätzers (7.18) ist also unabhängig von der Datentlänge N und kann demzufolge mit noch so langen Stichproben ($N \rightarrow +\infty$) nicht kleiner werden. Das Periodogramm ist also im Allgemeinen kein konsistenter Spektral-Schätzer, denn für diesen müsste immer $\lim_{N \rightarrow \infty} \sigma_{\widehat{P}_N(n)}^2 = 0$ gelten.

Einen konsistenten Schätzer erhält man, indem das Signal $\{x_m\}_{m=0}^{N^*-1}$ in Abschnitte (Datenfenster) $W(k) = \{x_m\}_{m=(k-1)N}^{kN-1}$, mit $k = 1, 2, \dots, K$ und $K \cdot N = N^*$ zerlegt wird. Für ein jedes Datenfenster $W(k)$ bestimmt man dann das Periodogramm $\widehat{P}_{N,k}(n)$ nach (7.19). Mittelt man nun über alle Periodogramme so liefert dies einen konsistenten Schätzer für das Powerspektrum.

Kapitel 8

Statistisches Schließen

8.1 Grundlegende Begriffe

Grundgesamtheit

In der *induktiven Statistik* versucht man, aus einer Folge von Einzel–Beobachtungen,

$$\mathbf{x}_T \equiv x_1, x_2, \dots, x_t, \dots, x_T \quad . \quad (8.1)$$

Aussagen über eine sogenannte *Grundgesamtheit* zu machen. Unter einer Grundgesamtheit versteht man das Tripel (G, X, P) , mit den folgenden Bedeutungen:

Objektmenge G :

Das kann beispielsweise eine Menge von physikalischen Größen, von Personen oder von Gegenständen sein.

Untersuchungsmerkmal X der Objekte: Ein Untersuchungsmerkmal ordnet einer Eigenschaft der Objekte aus G eine reelle Zahl zu. Das kann beispielsweise der Messwert einer physikalischen Größe, die Körpergröße von Personen oder die Farbe von Gegenständen sein. Ein Untersuchungsmerkmal ist also eine Zufallsgröße.

Wahrscheinlichkeits–Verteilungsfunktion P : Verteilungsfunktion des Untersuchungsmerkmals X bezüglich der Objekte in G . $P(x)$ ist also die Wahrscheinlichkeit dafür, dass die Zufallsgröße X Werte kleiner x annimmt, $P(x) = \text{prob}(X < x)$.

Stichprobe und Stichprobenfunktion

Die Beobachtungen (8.1) heißen auch *Stichprobe*, *konkrete Stichprobe* oder *Datensatz*. Sie stellen eine Realisierung des Prozesses

$$\mathbf{X}_T \equiv X_1, X_2, \dots, X_t, \dots, X_T \quad (8.2)$$

dar. Der Prozess (8.2) selbst wird auch *mathematische Stichprobe* genannt, wenn Folgendes gilt:

1. $X_t = X$ für alle $t = 1, \dots, T$ gilt, das heißt, die Zufallsgrößen X_t haben alle dieselbe Verteilung, und
2. die $\{X_t\}$ vollkommen statistisch unabhängig sind.

Damit ist also eine mathematische Stichprobe ein sogenannter *i.i.d. Prozess* (*i.i.d.* steht für *independent identically distributed*).

Eine Funktion f des Zufallsvektors (8.2) ist wiederum eine Zufallsgröße,

$$Y = f(\mathbf{X}_T) . \quad (8.3)$$

Sie heißt *Stichprobenfunktion*. Die Verteilungsfunktion $Q_T(y)$ von Y berechnet sich wie folgt,

$$\begin{aligned} Q_T(y) &= \text{prob}(Y < y) \\ &= \int_{\mathbf{x}_T \in \mathbb{R}^T: f(\mathbf{x}_T) < y} dS(\mathbf{x}_T) , \end{aligned}$$

wobei $S(\mathbf{x}_T)$ die Verteilungsfunktion der mathematischen Stichprobe (8.2) ist. Weil wir eine mathematische Stichprobe als *i.i.d. process* eingeführt haben, können wir auch schreiben

$$dS(\mathbf{x}_T) = dP(x_1) \dots dP(x_T) ,$$

worin P die Verteilung von X_t ist. Existiert die Dichte p von P , so gilt weiterhin

$$dS(\mathbf{x}_T) = p(x_1) \dots dp(x_T) dx_1 \dots dx_T .$$

Wir betrachten zwei Beispiele:

1. Verteilung des Produktes zweier unabhängiger Zufallsgrößen: Die Stichprobenfunktion $Y = X_1 X_2$ hat die Verteilung

$$\begin{aligned} Q_T(y) &= \int_{x_1 x_2 < y} p(x_1) p(x_2) dx_1 dx_2 \\ &= \int_{-\infty}^{+\infty} \frac{p(x) p(y/x)}{|x|} dx . \end{aligned}$$

2. Verteilung des Quotienten zweier unabhängiger Zufallsgrößen: Die Stichprobenfunktion $Y = X_1/X_2$ hat die Verteilung

$$\begin{aligned} Q_T(y) &= \int_{x_1/x_2 < y} p(x_1) p(x_2) dx_1 dx_2 \\ &= \int_{-\infty}^{+\infty} |x| p(x) p(yx) dx . \end{aligned}$$

Median und Spannweite einer Stichprobe

Wir denken uns eine Stichprobe (8.1) der Größe nach geordnet,

$$x_1^*, x_2^*, \dots, x_t^*, \dots, x_T^* \quad \text{mit} \quad x_1^* \leq x_2^* \leq \dots \leq x_t^* \leq \dots \leq x_T^* . \quad (8.4)$$

Der *Median der Stichprobe* (auch *empirischer Median* genannt) ist dann wie folgt definiert:

$$x_{\text{Median}} \equiv \begin{cases} x_{(T+1)/2}^* & : T \text{ ungerade} \\ \frac{x_{T/2}^* + x_{T/2+1}^*}{2} & : T \text{ gerade.} \end{cases} \quad (8.5)$$

Die *Spannweite der Stichprobe* ist

$$R \equiv x_T^* - x_1^* . \quad (8.6)$$

Quantile

Sei eine Wahrscheinlichkeits-Verteilung $P(x)$ einer Zufallsgröße X gegeben. Dann heißt der Wert x_α mit der Eigenschaft

$$P(x_\alpha) \leq \alpha \leq P(x_\alpha + 0) \quad (8.7)$$

α -Quantil oder auch α -Fraktil der Verteilung P . Somit ist x_α also eine Schwelle oberhalb derer X Werte mit der Wahrscheinlichkeit $1 - \alpha$ annimmt. Ist die Verteilungsfunktion in einer Umgebung von x_α streng monoton wachsend, so können wir das Quantil auch unmittelbar aus der Umkehrfunktion P^{-1} von P erhalten,

$$x_\alpha = P^{-1}(\alpha) .$$

Einige α -Quantile habe eine besondere Bezeichnung

$$\begin{aligned} x_{\alpha=1/2} & : \text{ Median} \\ x_{\alpha=1/4} & : \text{ unteres Quartil} \\ x_{\alpha=3/4} & : \text{ oberes Quartil} \\ x_{\alpha=q/10} & : \text{ } q\text{-tes Dezantil} \\ x_{\alpha=r/100} & : \text{ } r\text{-tes Perzentil} \end{aligned}$$

8.2 Wichtige Verteilungen

Will man Punkt- oder Intervallschätzungen machen bzw. Hypothesen testen, so spielen gewisse Wahrscheinlichkeits-Verteilungen eine besondere Rolle. Einige dieser Verteilungen stellen wir im Folgenden vor.

8.2.1 χ^2 -Verteilung (HELMERT-PEARSON-Verteilung)

Sei eine jede der Zufallsgrößen X_t einer mathematischen Stichprobe (8.2) standardisiert normalverteilt (Mittelwert 0 und Varianz 1), sei also die Wahrscheinlichkeitsdichte von einem jeden X_t gegeben durch

$$p(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

Dann hat die Stichprobenfunktion

$$Y = \sum_{t=1}^T X_t^2 \quad (8.8)$$

die sogenannte *Chi-Quadrat-Verteilung mit T Freiheitsgraden*. Ihre Dichte lautet

$$q_T(y) = \frac{1}{2^{\frac{T}{2}} \Gamma\left(\frac{T}{2}\right)} y^{\frac{T}{2}-1} e^{-\frac{y}{2}} \quad \text{mit } y \geq 0 \text{ und } T = 1, 2, 3, \dots \quad (8.9)$$

Für negative Werte y verschwindet die Dichte, kann doch die Stichprobenfunktion (8.8) keine negativen Werte annehmen. Abbildung 8.1 zeigt die Chi-Quadrat-Verteilung (8.9) für einige

niedrige und höhere Freiheitsgrade. Erwartungswert und Steuung sind $\bar{Y} = T$ bzw. $\overline{(Y - \bar{Y})^2} = 2T$. Der Modalwert (Stelle y , wo die Dichte maximal wird) ist $T - 2$ für $T > 2$.

Nach dem zentralen Grenzwertsatz ist die Stichprobenfunktion

$$Y^* \equiv \frac{Y - T}{\sqrt{2T}}$$

für $T \rightarrow +\infty$ standard-normalverteilt. Schon für $T \gtrsim 30$ ähnelt die Chi-Quadrat-Verteilung sehr einer Normalverteilung mit Mittelwert T und Standardabweichung $\sqrt{2T}$.

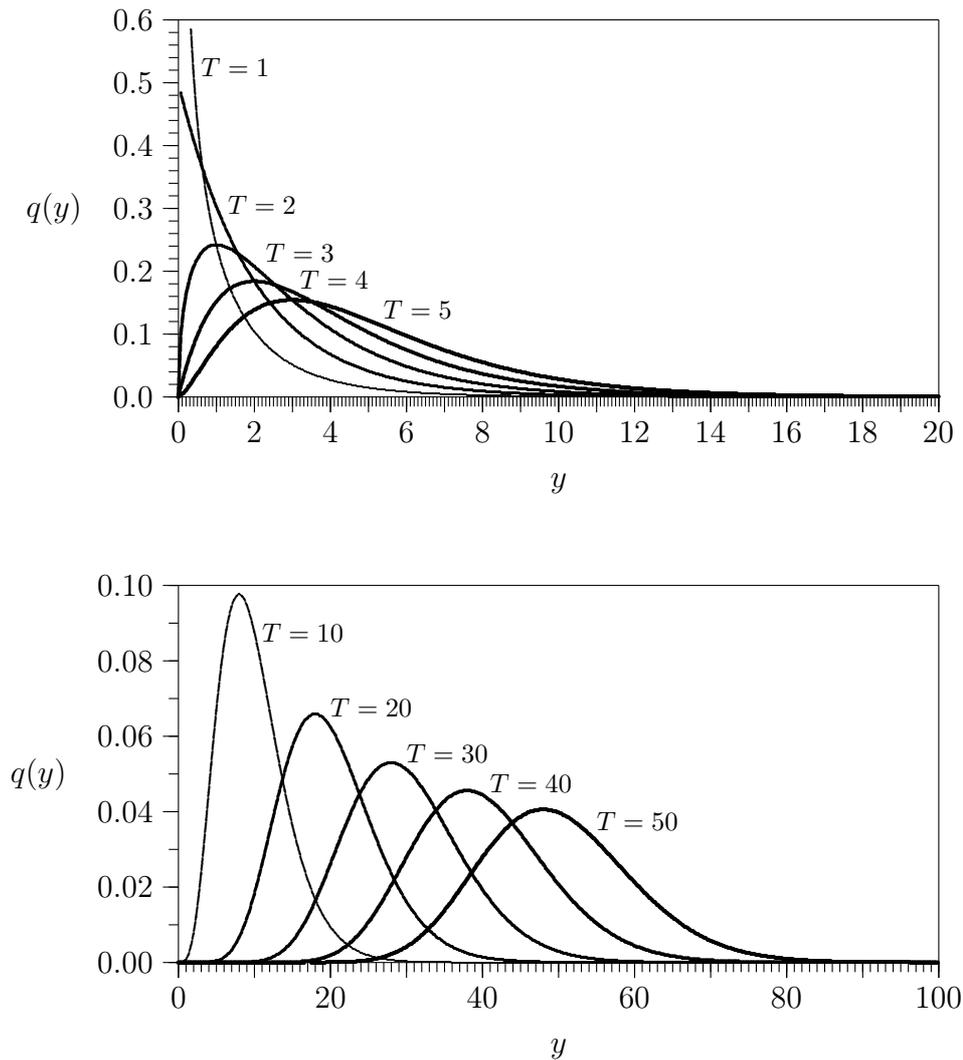
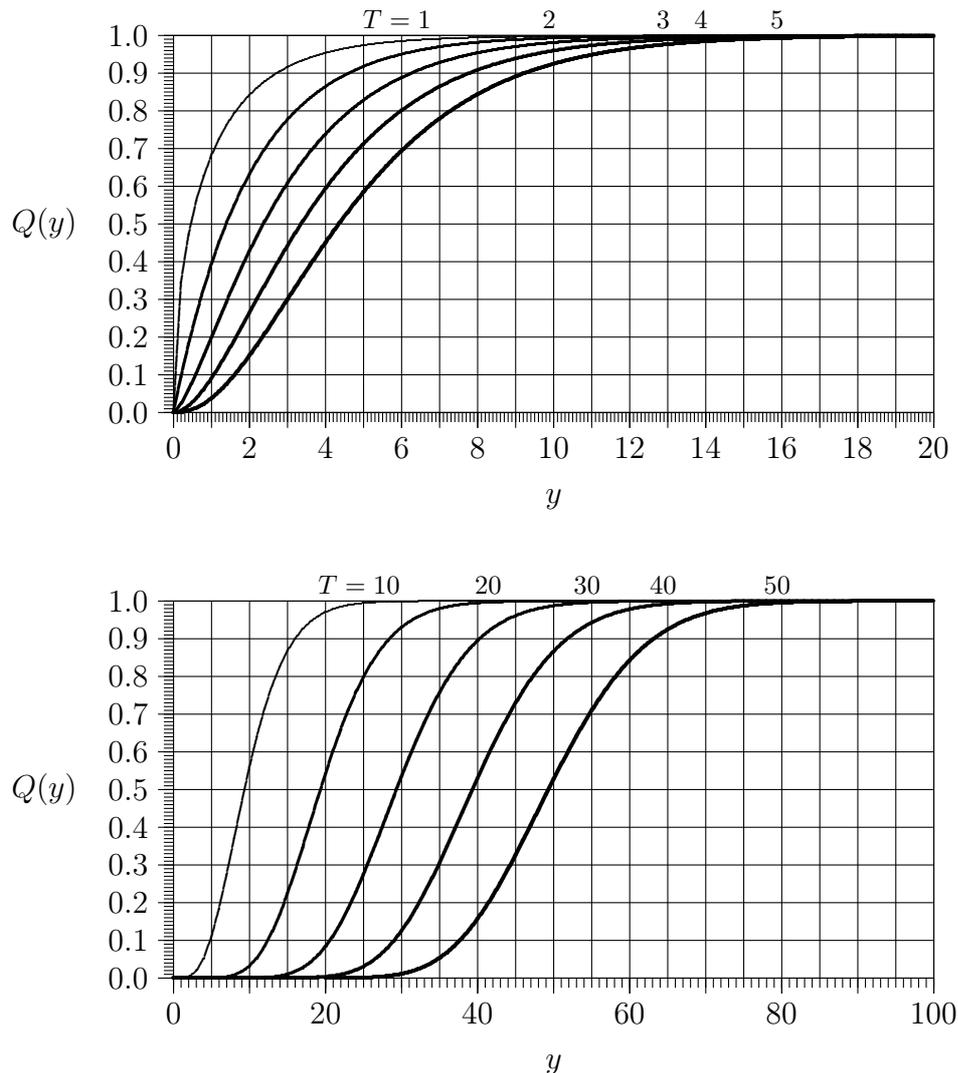


Bild 8.1: Dichte der Chi-Quadrat-Verteilung einiger niedriger und höherer Freiheitsgrade T

8.2.2 Student-Verteilung (t -Verteilung)

Sei eine jede der Zufallsgrößen X_t einer mathematischen Stichprobe (8.2) wiederum standardisiert normalverteilt. Dann hat die Stichprobenfunktion

$$Y \equiv \frac{X_1}{\sqrt{\sum_{t=2}^T X_t^2}} \tag{8.10}$$

Bild 8.2: Chi-Quadrat-Verteilung einiger niedriger und höherer Freiheitsgrade T

eine sogenannte *Student-Verteilung* mit T Freiheitsgraden, welche durch die folgende Dichte gegeben ist:

$$q_T(y) = \frac{1}{\sqrt{\pi T}} \frac{\Gamma\left(\frac{T+1}{2}\right)}{\Gamma\left(\frac{T}{2}\right)} \left(1 + \frac{y^2}{T}\right)^{-\left(\frac{T+1}{2}\right)} \quad \text{für } -\infty < y < \infty . \quad (8.11)$$

8.3 Signifikanztests

Eine *Hypothese* ist eine Vermutung bezüglich einer Grundgesamtheit oder mehrerer Grundgesamtheiten. Sie ist entweder wahr oder falsch. Leider kann man im Allgemeinen aus einer einzelnen Stichprobe vollkommen falsche Rückschlüsse auf die Grundgesamtheit ziehen. Im Rahmen der Statistik wird jedoch das Risiko für solche Fehlschlüsse genau kontrolliert. Ziel des statistischen Testens ist es, bei Kenntnis einer konkreten Stichprobe die Hypothese auf ihren Wahrheitsgehalt zu überprüfen. Dabei wird man allerdings niemals mit *absoluter* Sicherheit eine Antwort auf die Frage finden können, ob eine Hypothese wahr oder falsch ist.

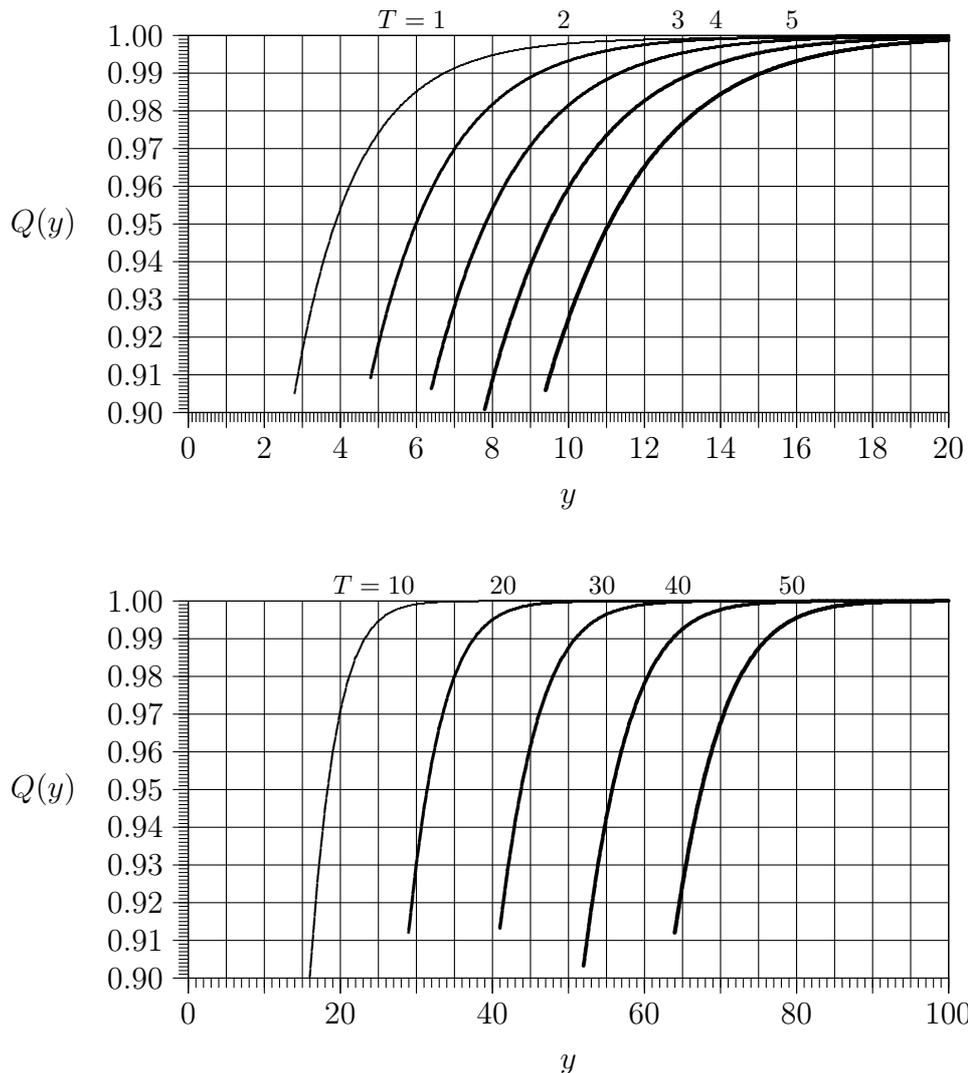


Bild 8.3: Ausschnitt aus der Chi-Quadrat-Verteilung einiger niedriger und höherer Freiheitsgrade T

Ist eine konkrete Stichprobe (8.1) gegeben, dann läuft das Testen in folgenden Schritten ab:

1. Formulierung einer *Nullhypothese* \mathcal{H}_0 für die zur Stichprobe (8.1) gehörenden Grundgesamtheit.
2. Festlegung einer *Testgröße* Y . Dies ist eine Stichprobenfunktion (8.3) $Y = f(\mathbf{X}_T)$. Die Verteilung $Q_T(y)$ von Y sei unter der Bedingung, dass \mathcal{H}_0 wahr ist, bekannt. Die Verteilung $Q_T(y)$ heißt auch *Prüfverteilung*, und die Testgröße wird auch *Teststatistik* genannt.
3. Festlegung eines Signifikanzniveaus α , $0 < \alpha \ll 1$. Bestimmung des $(1 - \alpha)$ -Quantils $y_{1-\alpha}$ der Verteilung $Q_T(y)$.
4. Berechnung einer konkreten Testgröße $TG = f(\mathbf{x}_T)$ aus einer konkreten Stichprobe (8.1) \mathbf{x}_T . TG ist also eine Realisierung von Y .
5. Für $TG(\mathbf{x}_T) > y_{1-\alpha}$ ist die Nullhypothese \mathcal{H}_0 auf dem Signifikanzniveaus α abzulehnen.

Andernfalls, also für $TG(\mathbf{x}_T) \leq y_{1-\alpha}$, kann auf der Basis des Testes nichts gegen \mathcal{H}_0 eingewendet werden.

Das Grundanliegen beim Signifikanztest ist es, die Wahrscheinlichkeit für eine Fehlentscheidung objektiv einzuschätzen. Die sogenannte *Irrtums-Wahrscheinlichkeit* p^* , die Nullhypothese \mathcal{H}_0 abzulehnen, obwohl sie wahr ist, ist nicht größer als α , $p^* \leq \alpha$. $1 - p^*$ heißt *Sicherheits-Wahrscheinlichkeit*, für sie gilt dann $1 - p^* \geq 1 - \alpha$. Das Intervall $]y_{1-\alpha}, +\infty[$ heißt *Ablehnungsbereich*.

Eigentliches Ziel eines Signifikanztestes ist es, die Nullhypothese \mathcal{H}_0 als „unglaublich“ hinzustellen, also \mathcal{H}_0 abzulehnen. Man will also nicht \mathcal{H}_0 verifizieren, sondern umgekehrt, \mathcal{H}_0 falsifizieren. Die Ablehnung von \mathcal{H}_0 nennt man *Alternativhypothese*.

Beim Testen einer Hypothese \mathcal{H}_0 müssen offenbar folgende Situationen unterschieden werden:

Nullhypothese \mathcal{H}_0	Testentscheidung	
	\mathcal{H}_0 nicht abgelehnt	\mathcal{H}_0 abgelehnt
wahr	Entscheidung korrekt	Fehler 1. Art
falsch	Fehler 2. Art	Entscheidung korrekt

Die Wahrscheinlichkeit für einen Fehler 1. Art haben wir bereits mit p^* bezeichnet, sie ist immer kleiner als das Signifikanzniveau α . Wird α nur hinreichend klein gewählt, so ist diese Fehlerart sehr gering wahrscheinlich. Im Falle der Ablehnung der Nullhypothese spricht man daher von einem *signifikanten* Testergebnis, falls $\alpha = 0.05$, sowie für $\alpha = 0.01$ und $\alpha = 0.001$ von einem *sehr signifikanten* bzw. *hochsignifikanten* Testergebnis. Wenn wir nun das Signifikanzniveau α verringern, so werden zwar seltener (unwahrscheinlicher) Fehler 1. Art begangen, allerdings wird dann auch seltener die gewünschte Ablehnung der Nullhypothese erreicht. Darüber hinaus werden dann häufiger Fehler 2. Art begangen. In der Praxis muss man mit der Festlegung des Signifikanzniveaus einen Kompromiss eingehen, man muss abschätzen, welche Fehlerart fatalere Folgen hat.

Führen entweder nur sehr kleine oder nur sehr große Werte einer Testgröße zur Ablehnung von \mathcal{H}_0 , so nennt man den Test *einseitig*. Will man aber auf sehr große und sehr kleine Abweichungen testen, so formuliert man eine *zweiseitige Nullhypothese*, die dann durch einen sogenannten *zweiseitigen Test* geprüft wird.

Unter dem Begriff χ^2 -Test werden alle Signifikanztests zusammengefasst, deren Testgröße exakt oder wenigstens näherungsweise eine χ^2 -Verteilung mit der Dichte (8.9) haben. Im Folgenden werden einige konkrete Fälle betrachtet.

8.3.1 χ^2 -Streuungstest

X sei eine normalverteilte Zufallsgröße mit Mittelwert \bar{X} und Varianz σ_X^2 .

1. Nullhypothese \mathcal{H}_0 :

$$\sigma_X^2 = \sigma_0^2.$$

2. Testgröße Y :

$$Y = \frac{(T-1)S^2}{\sigma_0^2},$$

mit der Stichprobenstreuung

$$S^2 = \frac{1}{T-1} \sum_{t=1}^T (X_t - \bar{X})^2, \quad \bar{X} = \frac{1}{T} \sum_{t=1}^T X_t.$$

Unter der Nullhypothese besitzt Y eine χ^2 -Verteilung mit $T-1$ Freiheitsgraden, $Q_{T-1}(y)$.

3. Signifikanzniveau α , Quantile :

$$Q_{T-1}^{-1}(\alpha/2) = y_{\alpha/2} \text{ und } Q_{T-1}^{-1}(1 - \alpha/2) = y_{1-\alpha/2},$$

4. Berechnung der Testgröße TG(\mathbf{x}_T) :

$$\text{TG} = \frac{(T-1)s^2}{\sigma_0^2},$$

mit der Stichprobenstreuung

$$s^2 = \frac{1}{T-1} \sum_{t=1}^T (x_t - \bar{x})^2, \quad \bar{x} = \frac{1}{T} \sum_{t=1}^T x_t.$$

5. Testaussage :

Ablehnung von \mathcal{H}_0 falls $\text{TG}(\mathbf{x}_T) < y_{\alpha/2}$ oder $\text{TG}(\mathbf{x}_T) > y_{1-\alpha/2}$.

8.3.2 χ^2 -Unabhängigkeitstest

Seien X und Y zwei diskrete Zufallsgrößen, welche die Werte $x_m, m = 1, \dots, M$ bzw. $y_n, n = 1, \dots, N$ mit den Wahrscheinlichkeiten $p_m, m = 1, \dots, M$ bzw. $q_n, n = 1, \dots, N$ annehmen. Die zum Zufallsvektor gehörigen Verbundwahrscheinlichkeiten bezeichnen wir mit $s_{mn} \equiv \text{prob}(X = x_m, Y = y_n)$. Wir wollen nun auf der Grundlage zweier Messreihen $x_1, x_2, \dots, x(t), \dots, x_T$ und $y_1, y_2, \dots, y(t), \dots, y_T$ herausfinden, ob X und Y statistisch *abhängig* sind. Dazu gehen wir wie folgt vor:

1. Nullhypothese \mathcal{H}_0 :

X und Y sind unabhängig, also $s_{mn} = p_m q_n$, für alle $m = 1, \dots, M, n = 1, \dots, N$.

2. Testgröße TG :

$$\mathcal{Y} = T \cdot \varphi^2,$$

mit der Kontingenz

$$\varphi^2 \equiv \sum_{m,n=1}^{M,N} \frac{(s_{mn} - p_m q_n)^2}{p_m q_n}.$$

(Es gilt immer $0 \leq \varphi^2 \leq \min\{M, N\} - 1$.) Für $T \rightarrow +\infty$ ist \mathcal{Y} χ^2 -verteilt mit $(M-1)(N-1)$ Freiheitsgraden.

3. Signifikanzniveaus $\alpha, (1 - \alpha)$ -Quantil $y_{1-\alpha}$:

$\alpha = 0.01$. Das Quantil $y_{0.99} = Q^{-1}(0.99)$ ist für den entsprechenden Freiheitsgrad T aus einer Tabelle oder Bild 8.3 abzulesen.

4. Berechnung der Testgröße $TG(\mathbf{x}_T)$:

$$TG(\mathbf{x}_T, \mathbf{y}_T) = T \cdot \sum_{m,n=1}^{M,N} \frac{(H_{mn} - H_{m \cdot} H_{\cdot n} / T)^2}{H_{m \cdot} H_{\cdot n}} \quad (8.12)$$

mit den absoluten Häufigkeiten

$$H_{mn} = \#\{t : (x(t), y(t)) = (x_m, y_n)\}$$

und

$$H_{m \cdot} = \sum_{n=1}^N H_{mn} \quad , \quad H_{\cdot n} = \sum_{m=1}^M H_{mn} \quad .$$

5. Testaussage :

\mathcal{H}_0 wird abgelehnt, falls $TG(\mathbf{x}_T, \mathbf{y}_T) > y_{0,99}$.

Bei kleinem T und $M = N = 2$ wird die Testgröße

$$TG_{\text{kor}}(\mathbf{x}_T, \mathbf{y}_T) = T \cdot \frac{[|H_{11}H_{22} - H_{12}H_{21}| - T/2]^2}{H_{1 \cdot} \cdot H_{\cdot 2} \cdot H_{\cdot 1} \cdot H_{2 \cdot}} \quad (8.13)$$

verwendet, die für kleine Werte T besser einer χ^2 -Verteilung mit einem Freiheitsgrad gehorcht.

Beispiel: Gegeben seien zwei binäre Zufallsfolgen der Länge $T = 6$,

$$\begin{aligned} \{x(t)\}_{t=1}^6 &= 121211 \\ \{y(t)\}_{t=1}^6 &= 212212 \end{aligned}$$

Wir wollen die Hypothese der Unabhängigkeit der beiden Binärfolgen auf dem Signifikanzniveau $\alpha = 0,01$ testen. Dazu bestimmen wir zunächst die Verbund- sowie Rand-Häufigkeiten: $H_{11} = H_{22} = H_{21} = 1$ und $H_{12} = 3$ sowie $H_{2 \cdot} = H_{\cdot 1} = 2$ und $H_{1 \cdot} = H_{\cdot 2} = 4$. Für die Testgröße (8.12) erhalten wir somit $TG = 3/8$ und für (8.13) $TG_{\text{kor}} = 3/32$. Nun lesen wir aus Bild 8.3 das 99te Perzentil (=0,99-Quantil) der χ^2 -Verteilung mit Freiheitsgrad 1 ab: $y_{0,99} \approx 6,7$. Wegen $TG < y_{0,99}$ (also $3/8 < 6,7$) können wir die Nullhypothese auf der Basis des vorliegenden Datenmaterials und beim gewählten Signifikanzniveau nicht verwerfen.

Hätten wir α deutlich größer gewählt und somit $TG > y_{1-\alpha}$ erreicht, so würden wir die Nullhypothese ablehnen können. Die Wahrscheinlichkeit für einen Fehler erster Art wäre dann aber recht groß, was letztlich die Aussage wertlos machte.

Andererseits könnten wir versucht sein, eine weitere Versuchsreihe zu gewinnen, und wiederum den Test ansetzen. Nehmen wir an, wir würden nun die sehr regelmäßig anmutenden Reihen

$$\begin{aligned} \{x(t)\}_{t=1}^6 &= 121212 \\ \{y(t)\}_{t=1}^6 &= 212121 \end{aligned}$$

erhalten. Unsere Testgröße liefert nun $TG = 6$. Dies ist immer noch kleiner als das Quantil $y_{0,99} \approx 6,7$, was also noch immer keine Ablehnung der Nullhypothese gestattet. Allerdings gilt

hier $TG > y_{0,90} \approx 3,1$. Hätten wir also von vorn herein mit dem Signifikanzniveau $\alpha = 0,10$ gearbeitet, so würde mit dem zweiten Datensatz die Nullhypothese abgelehnt werden können, im Unterschied zum ersten Datensatz, welcher die Ablehnung der Nullhypothese auch auf diesem Signifikanzniveau nicht gestattet hätte. Hier zeigt sich nun ein grundlegendes Problem: Wenn wir voraussetzen, dass die Daten unabhängig sind, dann gibt es sowohl für den ersten wie auch den zweiten Datensatz eine von Null verschiedene Wahrscheinlichkeit der Realisierung. Folglich würde es bei wiederholten Tests mit ein und den selben Parametern (Länge der Datensätze, Signifikanzniveaus), im Allgemeinen für einige Datensätze $TG > y_{1-\alpha}$ und für andere $TG < y_{1-\alpha}$ gelten können. Genau genommen ist es so, dass immer dann, wenn die Nullhypothese wahr ist und die Verteilung der Testgröße bei wahrer Nullhypothese exakt bekannt ist, für $\alpha \times 100\%$ aller Datensätze $TG > y_{1-\alpha}$ gilt. Das heißt aber, wir dürfen für unseren Test nicht immer wieder neue Datensätze verwenden, um irgendwann $TG > y_{1-\alpha}$ erhalten und dann \mathcal{H}_0 ablehnen. Dieses Verfahren ist unzulässig. Unser Test darf nur einmal, also für nur einen Datensatz, durchgeführt werden!

Darüber hinaus ist zu beachten, dass die Entscheidung $TG \geq y_{1-\alpha}$ nicht zu streng gesehen werden, wenn die Wahrscheinlichkeitsverteilung der Testgröße nur näherungsweise bekannt ist. Es sei hier daran erinnert, dass die Stichprobenfunktion (8.8) nur dann exakt χ^2 -verteilt ist, wenn die eingehenden Zufallsgrößen X_t exakt standard-normalverteilt sind. Für unseren Unabhängigkeitstest hieße dies, dass die Größen

$$\frac{|(H_{mn} - H_m \cdot H_n / T)|}{\sqrt{H_m \cdot H_n}}$$

standard-normalverteilt sein müssten. Dies trifft aber für die binären Datensätze der Länge 6 in unserem Beispiel bei weitem nicht exakt zu. Gerade für vergleichsweise kurze Datensätze und wenige mögliche Werte (hier nur die Werte 1 und 2) kann man aber anstelle der Testgrößen (8.12) oder (8.13) mit der Näherung durch die χ^2 -Verteilung auch exakte Verteilungen der folgenden Testgröße angeben,

$$TG = T \cdot \frac{H_{11} H_{22} - H_{12} H_{21}}{H_{1.} H_{.1} H_{2.} H_{.2}} .$$

Der entsprechende Test ist nach FISHER und YATES benannt. Man spricht hier auch von Fishers exaktem Test. Generell wird ein Test exakt genannt, wenn die exakte Verteilung der Testgröße verwandt wird. Oftmals ist diese Verteilung aber nicht exakt bekannt.

8.4 χ^2 -Anpassungstest

8.5 χ^2 -Homogenitätstest

Kapitel 9

Singulärwertzerlegung (SWZ)

9.1 Problemstellung

Die *Singulärwertzerlegung* (SWZ) ist auch unter den Namen *Karhune–Loewe–Transformation* oder *Hauptachsen–Transformation* bekannt. Die SWZ stellt eine Transformation einer vektorwertigen Zeitreihe auf neue Koordinaten dar, so dass diese paarweise unkorreliert sind.

9.2 Normierung

Gegeben sei ein Zeitreihenvektor der Dimension N ,

$$\mathbf{x}_t \equiv (x_{1,t}, \dots, x_{n,t}, \dots, x_{N,t}) \quad . \quad (9.1)$$

Die Zeit t durchlaufe die natürlichen Zahlen $1, 2, \dots, T$ — zeitdiskreter Fall — oder die reellen Zahlen $[1, T]$ — zeitkontinuierlicher Fall. Ein solcher Zeitreihenvektor kann im N -dimensionalen Raum \mathbb{R}^N als Bahnkurve dargestellt werden. Im zeitdiskreten Fall spricht man auch vom *scatter plot* oder allgemein vom *Phasenraumportrait*. Werden die Koordinaten von \mathbf{x}_t durch Phasenverschiebungen ϑ_n einer skalaren Zeitreihe x_t gewonnen, also

$$x_{n,t} = x_{t-\vartheta_n} \quad ,$$

so nennt man das Phasenraumportrait auch *Verzögerungsdarstellung* (*delay representation*).

Die Zeitreihen seien mittelwertfrei, d.h., es gelte

$$\overline{x_{n,t}} = 0 \quad \text{für alle } n = 1, 2, \dots, N \quad . \quad (9.2)$$

Darin ist der Mittelwert der Originalreihe durch

$$\overline{x_{n,t}} \equiv \frac{1}{T} \sum_{t=1}^T x_{n,t} \quad \text{bzw. durch} \quad \overline{x_{n,t}} \equiv \frac{1}{T} \int_0^T x_{n,t} dt$$

definiert, je nachdem, ob die Zeitreihe diskret oder kontinuierlich ist. Darüber hinaus seien die Zeitreihen so normiert, dass der Mittelwert über das Standard–Skalarprodukt¹⁾ von \mathbf{x}_t mit sich

¹⁾ Sind $\mathbf{x}, \mathbf{y} \in \mathbb{R}^N$, mit den Koordinaten $\mathbf{x} = (x_1, \dots, x_N)$ und $\mathbf{y} = (y_1, \dots, y_N)$, so gilt

$$(\mathbf{x}, \mathbf{y}) = \sum_{n=1}^N x_n y_n \quad .$$

selbst eins ergibt, also

$$\overline{(\mathbf{x}_t, \mathbf{x}_t)} = 1 . \quad (9.3)$$

Ist die Normierung bei einer experimentell erhaltenen Zeitreihe \mathbf{x}_t nicht gegeben, so müssen wir sie zunächst normieren, indem wir sie wie folgt transformieren:

$$x_{n,t} \longrightarrow \frac{x_{n,t} - \overline{x_{n,t}}}{\sqrt{(\mathbf{x}_t - \overline{\mathbf{x}}_t, \mathbf{x}_t - \overline{\mathbf{x}}_t)}} ,$$

mit

$$\overline{\mathbf{x}}_t \equiv (\overline{x_{1,t}}, \dots, \overline{x_{N,t}}) .$$

Wir nehmen im folgenden immer an, dass diese Transformation schon ausgeführt ist.

9.3 SWZ

Sei \mathbf{x}_t eine normierte N -dimensionale Zeitreihe. Wir suchen nun eine Darstellung

$$\mathbf{x}_t \approx y_{1,t} \mathbf{e}_1 . \quad (9.4)$$

Dabei möge \mathbf{e}_1 ein fester N -dimensionaler Einheitsvektor sein,

$$(\mathbf{e}_1, \mathbf{e}_1) = 1 ,$$

und y_t sei eine *skalare* Zeitreihe. Um den Approximationsfehler möglichst klein zu halten, fordern wir, dass der Winkel zwischen \mathbf{x}_t und \mathbf{e}_1 möglichst klein wird, und zwar gemittelt über alle Zeitpunkte t . Das bedeutet aber, dass das mittlere Betrags-Skalarprodukt möglichst groß ist. Um nicht mit dem Betrag rechnen zu müssen, fordern wir alternativ,

$$\overline{(\mathbf{x}_t, \mathbf{e}_1)^2} = \max!$$

Dies ist eine Bestimmungsgleichung für \mathbf{e}_1 .

Den Approximationsfehler bezeichnen wir mit

$$\Delta_{1,t} \equiv \mathbf{x}_t - y_{1,t} \mathbf{e}_1 .$$

Er kann nun ebenso approximiert werden. Dazu machen wir analog zu (9.4) den Ansatz

$$\Delta_{1,t} \approx y_{2,t} \mathbf{e}_2$$

und fordern, dass der Einheitsvektor \mathbf{e}_2 senkrecht auf \mathbf{e}_1 steht, also

$$(\mathbf{e}_1, \mathbf{e}_2) = 0 .$$

Dann ist die Forderung

$$\overline{(\Delta_{1,t}, \mathbf{e}_2)^2} = \overline{((\mathbf{x}_t - y_{1,t} \mathbf{e}_1), \mathbf{e}_2)^2} = \max!$$

äquivalent zu

$$\overline{(\mathbf{x}_t, \mathbf{e}_2)^2} = \max!$$

Der nun noch verbleibende Approximationsfehler ist

$$\begin{aligned}\Delta_{2,t} &\equiv \Delta_{1,t} - y_{2,t}\mathbf{e}_2 \\ &= \mathbf{x}_t - y_{1,t}\mathbf{e}_1 - y_{2,t}\mathbf{e}_2 .\end{aligned}$$

Dieses Verfahren setzen wir fort, bis wir die Einheitsvektoren

$$\mathbf{e}_1, \dots, \mathbf{e}_n, \dots, \mathbf{e}_N$$

konstruiert haben, die als Orthonormalbasis den \mathbb{R}^N aufspannen.

Im n -ten Konstruktionsschritt erhalten wir den Basisvektor

$$\mathbf{e}_n \equiv (e_{n,1}, \dots, e_{n,N})$$

aus der Forderung

$$\overline{(\mathbf{x}_t, \mathbf{e}_n)^2} = \max! ,$$

unter der Nebenbedingung

$$(\mathbf{e}_n, \mathbf{e}_n) = 1 .$$

Führen wir die Lagrange-Multiplikatoren λ_n ein, so können wir alternativ das Maximum der Funktion

$$\phi(e_{n,1}, \dots, e_{n,N}) \equiv \overline{(\mathbf{x}_t, \mathbf{e}_n)^2} - \lambda_n((\mathbf{e}_n, \mathbf{e}_n) - 1)$$

suchen. Das Extremum finden wir aus der Bedingung, dass die partiellen Ableitungen verschwinden,

$$\begin{aligned}0 &= \frac{\partial \phi(e_{n,1}, \dots, e_{n,m}, \dots, e_{n,N})}{\partial e_{n,m}} \\ &= 2\overline{x_{m,t}(\mathbf{x}_t, \mathbf{e}_n)} - 2\lambda_n e_{n,m} .\end{aligned}\tag{9.5}$$

Mit den Abkürzungen

$$R_{m,l} \equiv \overline{x_{m,t}x_{l,t}} \quad , \text{ mit } m, l = 1, 2, \dots, N$$

können wir anstelle von (9.5) auch

$$\sum_{l=1}^N R_{m,l}e_{n,l} = \lambda_n e_{n,m}\tag{9.6}$$

schreiben. Mit der Matrix

$$\mathbf{R} \equiv \begin{pmatrix} R_{1,1} & \dots & R_{1,l} & \dots & R_{1,N} \\ \vdots & & & & \\ R_{m,1} & \dots & R_{m,l} & \dots & R_{m,N} \\ \vdots & & & & \\ R_{N,1} & \dots & R_{N,l} & \dots & R_{N,N} \end{pmatrix}$$

können wir (9.6) noch kompakter schreiben,

$$\mathbf{R}\mathbf{e}_n = \lambda_n \mathbf{e}_n .$$

Somit können die Lagrange–Multiplikatoren λ_n als Eigenwerte der Matrix \mathbf{R} interpretiert werden, und die gesuchten Vektoren \mathbf{e}_n als zugehörige Eigenvektoren.

Offenbar ist \mathbf{R} symmetrisch,

$$\mathbf{R} = \mathbf{R}^\top ,$$

und folglich sind alle Eigenwerte λ_n positiv,

$$\lambda_n \geq 0 .$$

Für die Spur von \mathbf{R} gilt

$$\text{Tr}\mathbf{R} \equiv \sum_{m=1}^N R_{m,m} = \sum_{m=1}^N \overline{x_{m,t}x_{m,t}} = \overline{(\mathbf{x}_t, \mathbf{x}_t)} ,$$

und wegen der Normierung (9.3) folgt

$$\text{Tr}\mathbf{R} = 1 .$$

Andererseits ist aus der Algebra bekannt, dass die Spur einer Matrix gleich der Summe ihrer Eigenwerte ist, so dass wir also

$$\sum_{n=1}^N \lambda_n = 1$$

erhalten.

Die Eigenvektoren $\mathbf{e}_1, \dots, \mathbf{e}_N$ bilden eine vollständige Basis im \mathbb{R}^N , und folglich ist die Darstellung

$$\mathbf{x}_t = \sum_{n=1}^N y_{n,t} \mathbf{e}_n \tag{9.7}$$

möglich. Wegen der Orthonormalität der \mathbf{e}_n können wir die Koordinaten $y_{n,t}$ von \mathbf{x}_t in der neuen Basis \mathbf{e}_n wie folgt erhalten:

$$y_{n,t} = (\mathbf{x}_t, \mathbf{e}_n) . \tag{9.8}$$

Unkorreliertheit der neuen Koordinaten

Die neuen Koordinaten $y_{n,t}$ sind paarweise unkorreliert,

$$\overline{y_{n,t}y_{m,t}} = \begin{cases} \lambda_n & : n = m \\ 0 & : \text{sonst} \end{cases} , \tag{9.9}$$

denn es gilt:

$$\begin{aligned} \overline{y_{n,t}y_{m,t}} &= \overline{(\mathbf{x}_t, \mathbf{e}_n)(\mathbf{x}_t, \mathbf{e}_m)} \\ &= \overline{\left(\sum_{n^*} x_{n^*,t} \mathbf{e}_{n,n^*} \right) \left(\sum_{m^*} x_{m^*,t} \mathbf{e}_{m,m^*} \right)} \\ &= \sum_{n^*} \sum_{m^*} e_{n,n^*} e_{m,m^*} \overline{x_{n^*,t} x_{m^*,t}} \\ &= \sum_{n^*} \sum_{m^*} e_{n,n^*} e_{m,m^*} R_{n^*,m^*} . \end{aligned}$$

Unter Beachtung von (9.6) und $R_{n^*,m^*} = R_{m^*,n^*}$ folgt

$$\begin{aligned} \overline{y_{n,t}y_{m,t}} &= \sum_{m^*} e_{m,m^*} \left(\sum_{n^*} R_{n^*,m^*} e_{n,n^*} \right) \\ &= \sum_{m^*} e_{m,m^*} \lambda_n e_{n,m^*} \\ &= \lambda_n (\mathbf{e}_m, \mathbf{e}_n) . \end{aligned}$$

Hieraus folgt wegen der Orthogonalität der \mathbf{e}_n die Behauptung (9.9).

Approximation

Betrachten wir in der Darstellung (9.7) nur die ersten $K \leq N$ Summanden, so bekommen wir nur eine Approximation unserer Zeitreihe,

$$\mathbf{x}_t \approx \sum_{n=1}^K y_{n,t} \mathbf{e}_n . \quad (9.10)$$

Der mittlere Approximationsfehler ist gegeben durch

$$E_K \equiv \overline{\left(\mathbf{x}_t - \sum_{n=1}^K y_{n,t} \mathbf{e}_n \right)^2} .$$

Eine Rechnung analog zu der im vorigen Abschnitt ergibt

$$E_K = 1 - \sum_{n=1}^K \lambda_n .$$

Sind die Eigenwerte der Größe nach geordnet,

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N ,$$

so können wir leicht überschauen, wieviele Terme K in der Approximation (9.10) für unsere Zwecke als ausreichend angesehen werden können.

Kapitel 10

Vektorquantisierung

10.1 Problemstellung

Die Vektorquantisierung ist ein universelles Verfahren, um Folgen zufälliger Vektoren mit einer vorgegebenen mittleren Genauigkeit zu approximieren. Ausgangspunkt hierfür ist eine Trainingssequenz, aus der im sogenannten LBG-Algorithmus ein Codebuch erzeugt wird. Die Einträge in diesem Codebuch heißen *Codebuchvektoren*. Sie bilden ein Reservoir, aus dem wir für konkrete Vektoren eine „beste“ Approximation suchen.

Vektorquantisierer spielen eine Rolle bei der redundanzarmen Codierung von Sprach- oder anderen Signalen. Mit dem Verfahren können weitreichende statistische Abhängigkeiten berücksichtigt werden, indem hinreichend hoch-dimensionale Vektoren betrachtet werden. Darüber hinaus können Vektorquantisierer bei verschiedenen Aufgaben der Klassifizierung von zufälligen Merkmalsvektoren verwendet werden.

10.2 Prinzip

Gegeben sei eine Folge von Vektoren im \mathbb{R}^N ,

$$\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_t, \dots, \mathbf{x}_T . \quad (10.1)$$

Man kann sich diese zum Beispiel aus einem skalaren Signal x_t durch Delay-Konstruktion entstanden denken. Wir partitionieren nun den \mathbb{R}^N mit K vielen Boxen,

$$\begin{aligned} \beta &\equiv \{B_k\}_{k=1}^K, \\ B_{k_1} \cap B_{k_2} &= \emptyset \quad \text{falls } k_1 \neq k_2, \\ \bigcup_{k=1}^K B_k &= \mathbb{R}^N . \end{aligned}$$

Aus jeder Box B_k wählen wir nun einen Repräsentanten \mathbf{v}_k , den *Codebuchvektor*. Die Menge aller Codebuchvektoren bildet das *Codebuch*,

$$V \equiv \{\mathbf{v}_k\}_{k=1}^K . \quad (10.2)$$

Somit kann die Folge (10.1) auf eine Folge von Codebuchindizes abgebildet werden:

$$\mathbf{x}_t \xrightarrow{\mathbf{x}_t \in B_{k_t}} k_t .$$

Die Folge

$$k_1, k_2, \dots, k_t, \dots$$

kann nun beispielsweise binär codiert über einen Nachrichtekanal zum Empfänger geschickt werden. Kennt dieser das Codebuch (10.2), so kann er die Folge

$$\mathbf{v}_{k_1}, \mathbf{v}_{k_2}, \dots, \mathbf{v}_{k_t}, \dots \quad (10.3)$$

rekonstruieren. Diese Folge ist eine Approximation der ursprünglichen Folge (10.1), vorausgesetzt, wir haben die Codebuchvektoren und Partitionierung geeignet gewählt, so dass $\mathbf{x}_t \approx \mathbf{v}_{k_t}$ gilt. Dabei sollte die Anzahl K der Codebuchvektoren möglichst klein sein, um ihre Indizes mit möglichst kurzen (binären) Codeworten codieren zu können und somit letztlich die Bitrate des Vektorcoders klein zu halten. Andererseits darf diese Anzahl nicht zu gering werden, würde doch dann die Auswahl an Codebuchvektoren allzu gering werden, um noch den Approximationsfehler $d(\mathbf{x}_t, \mathbf{v}_{k_t})$ in akzeptablen Grenzen halten zu können. Bei der Entwicklung eines Vektorquantisierers steht somit die Aufgabe, ein in diesem Sinne optimales Codebuch zu finden.

Bei der Entwicklung eines Codebuches gehen wir von einem festen gegebenen Fehlermaß $\|\cdot\|$ aus. Dies kann der euklidische Abstand

$$\|\mathbf{x}_t - \mathbf{v}_{k_t}\| = \sqrt{(\mathbf{x}_t - \mathbf{v}_{k_t}, \mathbf{x}_t - \mathbf{v}_{k_t})}$$

sein, mit dem Standard-Skalarprodukt (\cdot, \cdot) . Allgemein sind aber auch andere Abstandsmaße möglich, die der konkreten Signalklasse und den Verwendungszweck des Coders angepaßt sind. Darüber hinaus muss bei der Entwicklung des Codebuches darauf geachtet werden, dass für die Entscheidung

$$\mathbf{x}_t \in B_{k_t}$$

ein möglichst effektiver Algorithmus zur Verfügung steht.

10.3 Optimales Codebuch bei gegebener Wahrscheinlichkeitsdichte

Für die Ableitung eines optimalen Codebuches machen wir den Ansatz, dass das Signal-Rausch-Verhältnis

$$\text{SNR} \equiv 10 \lg \frac{\sum_t \|\mathbf{x}_t\|^2}{\sum_t \|\mathbf{x}_t - \mathbf{v}_{k_t}\|^2}$$

maximiert werden soll. Offenbar wird dies erreicht, wenn der mittlere Quantisierungsfehler minimal ist,

$$E_\beta(V) \equiv \sum_{k=1}^K \int_{B_k} \|\mathbf{x} - \mathbf{v}_k\|^2 p(\mathbf{x}) d\mathbf{x}$$

minimal ist. Darin sei $p(\mathbf{x})$ die Wahrscheinlichkeitsdichte der Folge (10.1). Setzen wir nun die partiellen Ableitungen von $E_\beta(V)$ nach den Koordinaten von

$$\mathbf{v}_k = (v_{k1}, v_{k2}, \dots, v_{kn}, \dots, v_{kN})$$

gleich Null, so erhalten wir

$$\begin{aligned} 0 &= \frac{\partial E_{\beta}(V)}{\partial v_{kn}} \\ &= \int_{B_k} \frac{\partial \|\mathbf{x} - \mathbf{v}_k\|^2}{\partial v_{kn}} p(\mathbf{x}) \, d\mathbf{x} \\ &= - \int_{B_k} 2(x_n - v_{kn}) p(\mathbf{x}) \, d\mathbf{x} \end{aligned}$$

und somit

$$v_{kn} \int_{B_k} p(\mathbf{x}) \, d\mathbf{x} = \int_{B_k} x_n p(\mathbf{x}) \, d\mathbf{x}$$

und schließlich

$$\mathbf{v}_k = \frac{\int_{B_k} \mathbf{x} p(\mathbf{x}) \, d\mathbf{x}}{\int_{B_k} p(\mathbf{x}) \, d\mathbf{x}}.$$

Danach sind die Codebuchvektoren \mathbf{v}_k die „statistischen Schwerpunkte“ der Boxen B_k , $k = 1, 2, \dots, K$.

Die allgemeine Lösung dieses Optimierungsproblems setzt die Kenntnis der N -dimensionalen Dichte $p(\mathbf{x})$ voraus und ist selbst bei Kenntnis dieser Dichte noch recht kompliziert. Nur in sehr einfachen Modellsituationen wird eine analytische Lösung gefunden. In der Praxis werden deshalb andere Wege beschritten.

10.4 Optimales Codebuch aus Monte-Carlo-Methode: LBG-Algorithmus

LINDE, BUZO und GRAY haben 1980 den nach ihnen benannten *LBG-Algorithmus* zur Entwicklung eines Codebuches aus einer Trainingssequenz angegeben. Er kommt ohne explizite Kenntnis der Wahrscheinlichkeitsdichte $p(\mathbf{x})$ aus. Letztlich wird die entsprechende Information der Trainingssequenz entnommen.

Ausgangspunkt des LBG-Algorithmus ist eine Trainingssequenz (10.1) der Länge T und eine Norm $\|\mathbf{x}\|$, welche eine Metrik $\|\mathbf{x} - \mathbf{y}\|$ impliziert. Wir wählen in der Regel die euklidische Norm, aber auch andere, der konkreten Problematik angepasste Normen sind möglich. So werden z. B. bei der Sprachverarbeitung andere, hörphysiologisch motivierte Metriken verwendet.

Der LBG-Algorithmus ist iterativ. Dabei gehen wir in folgenden Schritten vor (Abb. 10.1)

1. Schritt (Initialisierung): Wir wählen zunächst ein Codebuch

$$V^{(0)} = \left\{ v_k^{(0)} \right\}_{k=1}^K,$$

das wir Start-Codebuch nennen, zufällig aus (auswürfeln). Die Vektoren in $V^{(0)}$ sollen alle verschieden sein, und es empfiehlt sich, sie „in die Nähe“ der Punkte der Trainingssequenz (10.1) zu legen. Man kann die Start-Vektoren $v_k^{(0)}$ auch zufällig aus der Trainingssequenz auswählen.

Der Parameter $A^{(0)}$ wird zur Konstruktion eines Abbruchkriteriums für den LBG-Algorithmus verwendet, was wir weiter unten noch erleutern.

Der Parameter m gibt uns den Iterationsschritt im LBG-Algorithmus an und ist zunächst auf 0 gesetzt.

2. Schritt (Konstruktion neuer Codebuchvektoren, $m := m + 1$): In der Box

$$B_k^{(m)} = \left\{ \mathbf{x} : \left\| \mathbf{x} - \mathbf{v}_k^{(m-1)} \right\| < \left\| \mathbf{x} - \mathbf{v}_{k^*}^{(m-1)} \right\| \quad \forall k \neq k^* \right\}$$

fassen wir alle Punkte des \mathbb{R}^N zusammen, die am Codebuchvektor $\mathbf{v}_k^{(m-1)}$ dichter, als an einem beliebig anderen Codebuchvektor $\mathbf{v}_{k^*}^{(m-1)} \in V^{(m-1)}$ liegen.

Der neue Codebuchvektor

$$\mathbf{v}_k^{(m)} = \frac{1}{\#\{\mathbf{x}_t \in B_k^{(m)}\}} \sum_{\mathbf{x}_t \in B_k^{(m)}} \mathbf{x}_t$$

ist der arithmetische Schwerpunkt aller Vektoren der Trainingssequenz, welche in die Box $B_k^{(m)}$ fallen. Dabei ist $\#\{\mathbf{x}_t \in B_k^{(m)}\}$ die Anzahl dieser Vektoren.

Die neuen Codebuchvektoren $\mathbf{v}_k^{(m)}$, $k = 1, 2, \dots, K$, werden im neuen Codebuch $V^{(m)}$ zusammengefaßt.

3. Schritt (Beurteilung der Güte des neuen Codebuches): Die Größe

$$\min_k \left\{ \left\| \mathbf{x}_t - \mathbf{v}_k^{(m)} \right\| \right\}$$

ist der Abstand des Vektors \mathbf{x}_t der Trainingssequenz zu jenem neuen Codebuchvektor $\mathbf{v}_k^{(m)}$ aus $V^{(m)}$, der \mathbf{x}_t im Sinne unserer Metrik am besten approximiert, und folglich ist

$$A^{(m)} = T^{-1} \sum_{t=1}^T \min_k \left\{ \left\| \mathbf{x}_t - \mathbf{v}_k^{(m)} \right\| \right\}$$

der über die Trainingssequenz gemittelte Wert dieser kleinsten Abstände. Je kleiner $A^{(m)}$, desto besser können wir, in einem statistischen Sinne, die Vektoren der Trainingssequenz durch Vektoren aus dem aktuellen Codebuch $V^{(m)}$ approximieren. In diesem Sinne ist $A^{(m)}$ ein „Gütemesser“ für das Codebuch.

Wir erwarten, dass $A^{(m)}$ für fortschreitende Iterationszahl fällt. Dann mißt der Quotient

$$Q \equiv \frac{A^{(m-1)} - A^{(m)}}{A^{(m-1)}} ,$$

wie stark sich die Güte des Codebuches mit dem m ten Iterationsschritt verbessert hat. Offenbar ist Q ein relatives Qualitätsmaß, und es gilt:

Verbesserung?		Güte
keine	$A^{(m-1)} = A^{(m)}$	$Q = 0$
gering	$A^{(m-1)} \gtrsim A^{(m)}$	$0 < Q \ll 1$
stark	$A^{(m-1)} \ll A^{(m)}$	$Q \lesssim 1$
schlechter	$A^{(m-1)} < A^{(m)}$	$Q < 0$

Damit kann $0 \leq Q < \varepsilon$ als Abbruchbedingung dienen. Darin ist $\varepsilon \ll 1$ eine (zunächst) willkürlich gewählte Abbruchschwelle. Für $\varepsilon < Q$ wiederholen wir das Prozedere ab dem 2. Schritt. Für $Q < 0$ sollte ein neues Startcodebuch gewählt werden.

Mit dem LBG-Algorithmus findet man im allgemeinen kein globales Optimum. In der Praxis kann man sich damit behelfen, den Algorithmus mehrere Male mit verschiedenen Startcodebüchern auszuführen, und dann das beste Codebuch zu verwenden.

10.5 Schnelle Suche

Rechenzeit-Einschränkungen gibt es für den LBG-Algorithmus in der Regel nicht, da in den meisten Anwendungsfällen die Codebücher Off-line erzeugt werden. Allerdings ist es für die Nutzung eines Codebuches V im On-line-Betrieb wichtig, möglichst schnell den besten Repräsentanten aus V für einen Vektor \mathbf{x}_t zu finden. Mit unserer bisherigen Vorgehensweise müßten wir zur Auffindung der besten Approximation \mathbf{v}_{k_t} von \mathbf{x}_t , also von

$$k_t : \|\mathbf{x}_t - \mathbf{v}_{k_t}\| \leq \|\mathbf{x}_t - \mathbf{v}_k\|, \quad k = 1, 2, \dots, K \quad (10.4)$$

genau K Abstandberechnungen $K - 1$ Vergleiche ausführen. Eine solche Vektorquantisierung (VQ) wird *full search VQ* (FSVQ) genannt. Dabei kann K einige Hundert oder bei höheren Approximations-Genauigkeiten bzw. größeren Vektordimensionen N gar einige Tausend und mehr betragen. Somit stellt sich die Frage nach schnellen Suchalgorithmen. Durch eine geeignete Strukturierung des Codebuches kann der Suchaufwand erheblich reduziert werden.

Bei der Baumsuche eines „besten“ Codebuchvektors (tree structured VQ = TSVQ) wird eine für die Suche günstige Codebuch-Struktur schon beim Entwurf des Codebuches durch geeignete Bedingungen erzwungen. Insbesondere kann hierdurch erreicht werden, dass die Anzahl der Abstandsberechnungen für alle Vektoren gleich ist.

Zur Entwicklung eines baumstrukturierten Codebuches geht man wie folgt vor:

1. Schritt: Erzeugung eines Codebuches

$$V \equiv \{\mathbf{v}_{k_1}\}_{k_1=1}^K$$

mit dem LBG-Algorithmus zur Trainingsfrequenz (10.1).

2. Schritt: Einteilung der Trainingsfrequenz (10.1) in K Teilsequenzen

$$\{\mathbf{x}_{k_1 t}\}_{t=1}^{T_{k_1}}, \quad \text{so dass } \mathbf{x}_{k_1 t} \in B_{k_1}, \quad k_1 = 1, 2, \dots, K \text{ .}$$

Dabei ist B_{k_1} die zu $\mathbf{v}_{k_1} \in V_1$ gehörige Box.

Eine jede Teilsequenz wird als Trainingssequenz des LBG-Algorithmus verwendet, was zu weiteren K Codebüchern

$$V_{k_1} \equiv \{\mathbf{v}_{k_1 k_2}\}_{k_2=1}^K$$

führt.

m -ter Schritt: Es werden immer wieder Teilsequenzen der Trainingssequenz gebildet, die aus allen Vektoren \mathbf{x}_t bestehen welche in der Box $B_{k_1 \dots k_{m-1}}$ liegen. Zu dieser Trainingssequenz wird mit dem LBG-Algorithmus das Codebuch

$$V_{k_2 \dots k_{m-1}} \equiv \{\mathbf{v}_{k_1 k_2 \dots k_{m-1} k_m}\}_{k_m=1}^K$$

gebildet.

Dabei wird der LBG-Algorithmus insgesamt

$$1 + K + K^2 + \dots + K^{m-1} = \frac{K^m - 1}{K - 1}$$

mal angewandt. Die Teil-Trainingssequenzen werden in der Regel von Schritt zu Schritt kürzer.

Auf der m -ten Stufe erhalten wir somit K^m Codebuchvektoren. Um nun einen Eingangsvektor \mathbf{x}_t zu codieren, könnten wir einerseits wie in (10.4) verfahren, auf der Grundlage des Codebuches $V_{1k_2\dots k_m}$. Dazu sind im wesentlichen K^m Vergleiche und Abstandsberechnungen nötig.

Wir können aber auch iterativ vorgehen, indem wir zunächst

$$k_1 : \|\mathbf{x}_t - \mathbf{v}_{k_1}\| \leq \|\mathbf{x}_t - \mathbf{v}_k\|, \quad k = 1, 2, \dots, K$$

bestimmen. Dann

$$k_2 : \|\mathbf{x}_t - \mathbf{v}_{k_1 k_2}\| \leq \|\mathbf{x}_t - \mathbf{v}_{k_1 k}\|, \quad k = 1, 2, \dots, K$$

und schließlich

$$k_m : \|\mathbf{x}_t - \mathbf{v}_{k_1 k_2 \dots k_{m-1} k_m}\| \leq \|\mathbf{x}_t - \mathbf{v}_{k_1 k_2 \dots k_{m-1} k}\|, \quad k = 1, 2, \dots, K .$$

Dazu sind nur noch $K \times m$ Vergleiche und Abstandsberechnungen nötig! Für $K = 2$ und $m = 10$ ist die relative Aufwandsersparnis gegenüber dem FSVQ erheblich:

$$\frac{K \times m}{K^m} = \frac{20}{1024} \approx 0,02 .$$

Bei diesem Verfahren müssen wir aber auch alle anderen Codebücher der vorherigen Stufen speichern. Das sind insgesamt

$$K + K^2 + \dots + K^m = \frac{K^{m+1} - K}{K - 1}$$

Codebuchvektoren.

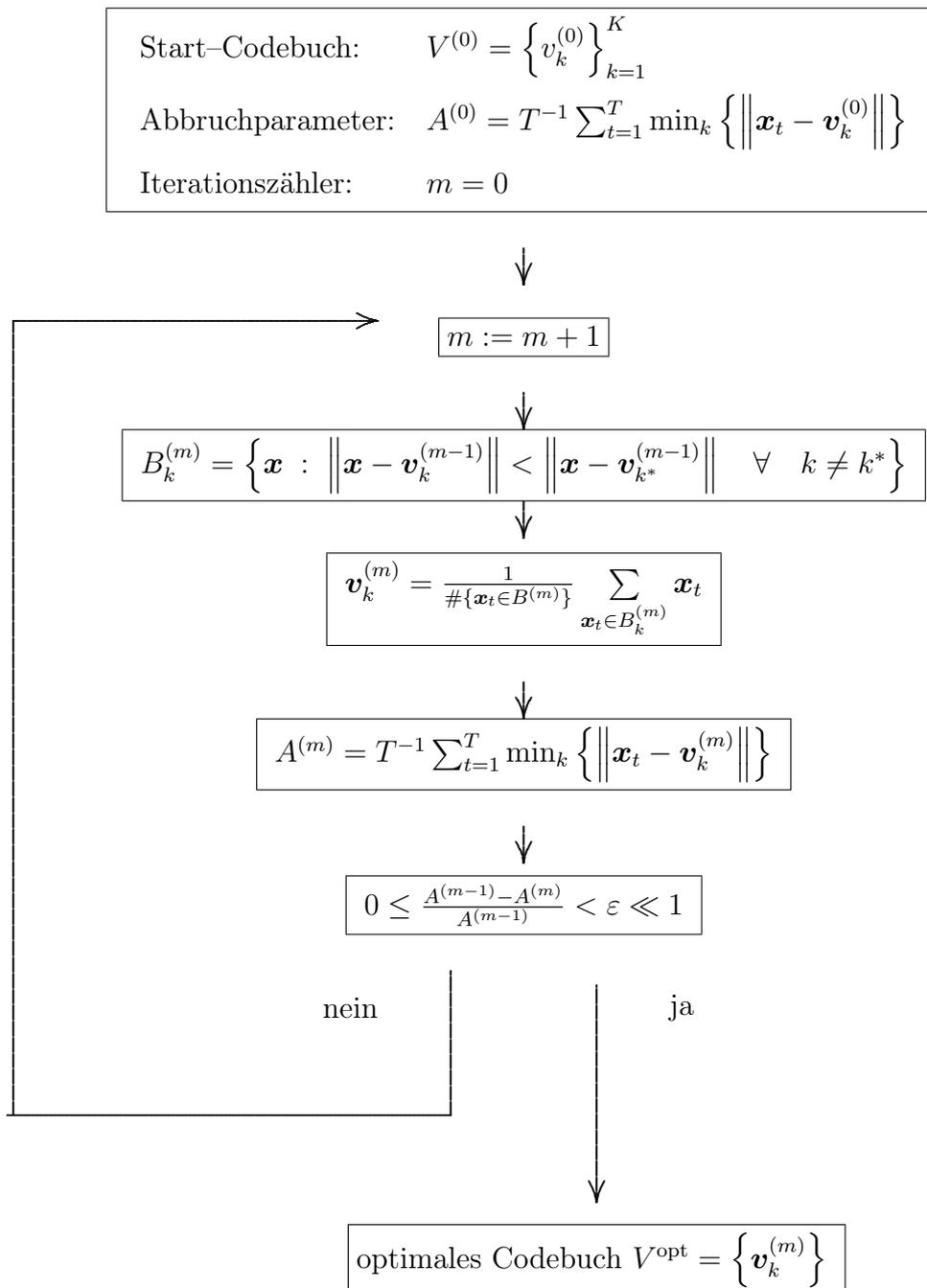


Bild 10.1: LBG-Algorithmus

Kapitel 11

Künstliche Neuronale Netze (KNN)

Zwischen der Signalverarbeitung in traditionellen Rechenmaschinen (VonNeumann–Rechnern) und der in natürlichen neuronalen Netzen (NNN) höherer Organismen, insbesondere jener im zentralen Nervensystem von Vertebraten (Wirbeltieren), gibt es einige wesentliche Unterschiede. Sie führen dazu, daß die NNN zum Beispiel Aufgaben der Mustererkennung (Sprach-, Bildererkennung u.a.) bedeutend schneller und zuverlässiger erfüllen als herkömmliche Rechenmaschinen. Selbst wenn Teilbereiche ausfallen, erfüllt das verbliebene NNN häufig noch die Funktionen des ursprünglichen, vielleicht zunächst in eingeschränkter Weise, aber nach einem Trainingsprozeß kann unter Umständen wieder die ursprüngliche Leistungsfähigkeit erreicht werden. Die NNN haben also die Fähigkeit, sich selbst zu organisieren, um einer gewissen Aufgabe auch unter veränderlichen Bedingungen gerecht zu werden.

Allerdings arbeiten NNN vergleichsweise langsam. Liegen bei Rechenmaschinen die Schaltzeiten logischer Elementarbausteine, von Transistoren bzw. Gattern, in der Größenordnung von 1 ns, so sind die Grundbausteine von NNN, die Neuronen, mit etwa 1 ms etwa eine Million mal langsamer. Diesem Nachteil steht allerdings der Vorteil einer hochgradig parallelen Verarbeitung entgegen. Ein Mensch kann innerhalb von etwa 100ms das Bild einer bekannten Person erkennen. Bei einer elementaren Schaltzeit von 1ms entspricht dies gerade mal 100 Rechenschritten. Ein konventioneller VonNeumann–Rechner kann in 100 Taktperioden nur einige wenige Befehle abarbeiten.

Mit sogenannten *Künstlichen Neuronalen Netzen* (KNN oder auch ANN für *artificial neural network*) versucht man, leistungsfähigere Rechensysteme nach dem natürlichen Vorbild zu schaffen, um komplexe praktische Aufgaben technisch zu lösen. Erste Überlegungen zu den KNN reichen bis in die vierziger Jahre (McCULLOCH und PITTS, 1943) zurück. Zwischenzeitlich ebnete das Interesse an ihnen immer wieder einmal ab. Allerdings erleben sie seit Anfang der achtziger Jahre eine gewisse Renaissance, etwa durch die Arbeiten von HOPFIELD.

11.1 Neuronen als Elementarbausteine natürlicher Neuronaler Netze

Das Nervensystem des Menschen besteht aus ca. 10^{11} Nervenzellen, den *Neuronen*, und etwa zehn– bis 50mal sovielen *Gliazellen*. Die eigentliche Signalverarbeitung bzw. –übertragung wird durch die Neuronen realisiert. Sie bestehen aus einem Zellkörper, dem *Soma*, und Zellfortsätzen, *Axone* und *Dendriten*. Ebenso wie andere Körperzellen verfügen auch Neuronen über eine vollständige molekulare Struktur, um lebensfähig zu sein, z.B. über einen Zellkern und

Mitochondrien. Allerdings ist ihre Fähigkeit zur Zellteilung nach der Embryonalentwicklung nicht mehr gegeben. Darüber hinaus besitzen Neuronen die einzigartige Fähigkeit, auch über große Entfernungen (ca. 1 Meter) präzise und schnell zu kommunizieren und sich adaptiv zu vernetzen, womit sie auf veränderte Umweltbedingungen reagieren können.

Gliazellen haben gewisse Hilfsfunktionen. Sie dienen beispielsweise als Stützelement, beseitigen die Zelltrümmer nach Zellverletzung oder -tod und bilden sogenannte *Myelinscheiden* zur elektrischen Isolation um *Axone* aus.

Axone und Dendriten sind unidirektionale Leitungen für elektrische Impulse. Axone entspringen im Soma und leiten Impulse vom Soma weg hin zu einem anderen Neuron. Über Dendriten empfängt ein Neuron Impulse. Axone können eine Länge von einigen Millimetern bis zu etwa einem Meter haben. Damit sind sie in der Regel bedeutend länger als Dendriten. Ein Neuron hat ein Axon und möglicherweise viele Dendriten. An den Enden können Axone verzweigen. Nach ihrer Funktion unterscheidet man drei Klassen von Neuronen:

sensorische Neuronen: Sie erhalten Reize der Umgebung oder des Körperinneren von einem sensorischen Element und leiten diese zu einem Interneuron weiter.

Interneuronen: Bilden seitens ihrer Axone wie auch der Dendriten Synapsen mit anderen Neuronen. Sie vermitteln also zwischen Neuronen. Sogenannte *lokale Interneurone* haben vergleichsweise kurze Axone, sie vermitteln innerhalb einer Gehirnregion. Interneurone mit langen Axonen heißen *Relais-* oder auch *Projektionsinterneuron*. Sie übermitteln Informationen zu einem relativ weit entfernten Zielneuron.

motorische Neuronen (Motoneuronen): Sie steuern Muskelzellen an, d.h., ihre Axone bilden Synapsen mit Muskelzellen. Ihre Information erhalten sie vom zentralen Nervensystem.

Die Schnittstelle zwischen zwei Zellfortsätzen verschiedener Neuronen heißt *Synapse*. Synapsen werden in der Regel zwischen dem Axon des einen und einem Dendriten eines anderen Neurons ausgebildet. Die Anzahl der Synapsen eines Neurons kann stark variieren. Ein typisches Neuron bildet zwischen 1000 und 10000 Synapsen aus und empfängt noch weitaus mehr synaptische Eingänge. Manche Neuronen, wie die vergleichsweise großen Purkinje-Zellen im Kleinhirn (Cerebellum), besitzen bis zu 150 Tausend Synapsen.

Die Reizleitung entlang der Axone erfolgt durch einen Fluß von Kalium- und Natrium-Ionen durch die etwa 5nm dicke Zellmembran der Nervenfasern. Im Ruhezustand ist die elektrische Spannung im Inneren der Faser etwa -70mV gegenüber dem Äußeren. Das Aktionspotential beträgt hingegen etwa +30mV. Die Erregungsstärke wird über die Impulsfrequenz moduliert, je stärker die Erregung desto höher ist die Frequenz. Im synaptischen Spalt sind spezielle Substanzen (Neurotransmitter) für die Reizleitung wirksam.

11.2 Die Modellierung natürlicher Neuronaler Netze

Natürliche Neuronaler Netze sind so komplex, daß bei ihrer Modellierung viele Idealisierungen und Vereinfachungen gemacht werden müssen. Dennoch können in KNN wesentliche Aspekte berücksichtigt werden. Das sind beispielsweise:

1. massive Parallelität
2. einfache Elemente

3. gerichtete Verbindungen zwischen Neuronen über gewichtete Synapsen
4. keine Kommunikation mittels komplexer Datenstrukturen, sondern allein durch die Tatsache, daß es Verbindungen gibt
5. Verbindungsgewichte sind modifizierbar, sie können in einem Lernprozeß verstärkt oder geschwächt werden.
6. hoher Grad der Konnektivität, ein Neuron kommuniziert mit sehr vielen anderen

Deutliche Abweichungen von KNN gegenüber NNN sind beispielsweise:

1. KNN haben eine viel geringere Zahl von Neuronen, etwa 100 bis 10000 gegenüber ca. 10^{11} bei NNN.
2. KNN haben eine viel weniger Verbindungen, etwa 10^5 gegenüber ca. 10^{15} bei NNN.
3. KNN haben zumeist nur einen Parameter für die Stärke der synaptischen Kopplung.
4. Bei KNN wird die Erregungsstärke nicht frequenz- sondern amplitudenmoduliert. Bislang nicht klar ist, ob dies ein entscheidender Nachteil von KNN ist.

Prinzipiell könnte man bei der Simulation von KNN auf Digitalrechnern in einem jeden Punkt sich dem biologischen Vorbild nähern, allerdings steigt dabei der Simulationsaufwand erheblich.

Ein KNN besteht aus folgenden Komponenten:

1. Zellen(Neuronen, Elemente, units)
2. Verbindungsnetzwerk
3. Propagierfunktion
4. Lernregel.

Wir stellen uns das gesamte Netz in Schichten $S_0, \dots, S_d, \dots, S_D$ unterteilt vor, wobei die Schichten seriell und unidirektional miteinander verbunden sind. Ein jedes Neuron der einen Schicht S_d ist mit einem jeden der darauffolgenden Schicht S_{d+1} verknüpft. Der Aktivierungszustand der i_{d+1} ten Zelle in S_{d+1} zur Zeit t sei $a_{d+1, i_{d+1}}(t)$. Die Zelle empfängt Information von allen Zellen aus S_d . Die i_d te Zelle in S_d möge zur Zeit t die Ausgabe $o_{d, i_d}(t)$. Diese Ausgabe wird nun entsprechend der momentanen Wichtigkeit $w_{i_d i_{d+1}}(t)$ summiert, was durch die *Propagierungsfunktion*

$$e_{i_{d+1}}(t) \equiv \sum_{i_d} w_{i_d i_{d+1}} o_{d, i_d}(t)$$

beschrieben wird. Die Aktivierung der i_{d+1} ten Zelle in S_{d+1} ist dann einen Zeitschritt später gegeben durch die *Aktivierungsfunktion*

$$a_{i_{d+1}}(t+1) = f_{\text{act}}(a_{i_{d+1}}(t), e_{i_{d+1}}(t)) \ .$$

Die Ausgabe der Zelle für die nächste Schicht S_{d+2} ist dann im $(t+1)$ ten Zeitschritt durch eine *Ausgabefunktion* gegeben,

$$o_{i_{d+1}}(t+1) = f_{\text{out}}(a_{i_{d+1}}(t+1)) \ .$$

Damit kann ein KNN auch als ein gerichteter, gewichteter Graph angesehen werden. Die Kanten sind hier die gewichteten Verbindungen zwischen den Neuronen.

Hat man sich nun auf eine gewisse Struktur des KNN festgelegt, d.h., sind Anzahl und Größe der Schichten sowie die Aktivierungs- und Ausgabefunktionen fixiert, so werden in einem *Trainingsprozeß* wiederholt gewisse *Trainingsmuster* $\{a_{i_0}\}$ für die Eingabeschicht S_0 dargeboten und das Netz D mal getaktet bis sich die Antwort des KNN in seiner Ausgabeschicht S_D zeigt. Nach bestimmten *Lernregeln* werden nun die Gewichte $w_{i_d i_{d+1}}$ bei wiederholter Präsentation der Trainingsmuster variiert, bis das KNN die gewünschten Ausgaben liefert. So können die Trainingsmuster z.B. quantisierte Schwarz-Weiß-Bilder von handgeschriebenen Ziffern sein und das KNN so trainiert werden, daß es eine Klassifizierung der Ziffern durchführt.

Die Wahl der Lernregel ist eines der zentralen Probleme beim Entwurf von KNN. Weit verbreitet ist die nach dem Psychologen DONALD O. HEBB (1949) benannte Regel, wonach zwei synaptische Verbindungen (also die Gewichte $w_{i_d i_{d+1}}$ und $w_{i_{d+1} i_{d+2}}$) verstärkt werden, wenn prä- und postsynaptisches Neuron gleichzeitig aktiv sind.¹⁾

¹⁾Hebb konnte diese Lernregel nur hypothetisch formulieren. Mittlerweile wurden derartige Lernvorgänge im Hippocampus und im visuellen System tatsächlich nachgewiesen.

Kapitel 12

Bayes Statistik

Gegeben sei eine absolut stetige Zufallsvariable X . H_i , $i = 1, 2$, seien zwei Hypothesen, welche davon ausgehen, dass p_1 bzw. p_2 die wahre Verteilungsdichte von X ist. Dann ist die Wahrscheinlichkeit für $X \in [x, x + dx]$ gegeben durch $p_1(x)dx$ bzw. $p_2(x)dx$. Wir fragen nun nach den Wahrscheinlichkeiten der beiden Hypothesen, falls ein Messwert x , also eine Realisierung von X vorliegt.

Bezeichne $W(H_i, x)$ die Verbundwahrscheinlichkeit, dass H_i wahr ist und $X \in [x, x + dx]$ gemessen wird. Solange wir keine Messung ausgeführt haben, können wir irgendwelche Vermutungen darüber anstellen, wie groß die Wahrscheinlichkeit für die Richtigkeit der Aussage „ H_i trifft zu“ ist. Diese, sogenannte *apriori-Wahrscheinlichkeit* bezeichnen wir mit $W(H_i)$. Sie ist *unabhängig* vom später beobachteten Messwert x . In der Praxis besteht eine gewisse Willkür für den Wert von $W(H_i)$, was später noch genauer diskutiert wird. Wir nehmen hier zunächst an, dass wir $W(H_i)$ kennen, und dass das Ereignis „ H_i trifft zu“ statistisch unabhängig vom später tatsächlich beobachteten Messwert ist. Dann ist die Verbundwahrscheinlichkeit $W(H_i, x)$ gegeben durch ein Produkt,

$$W(H_i, x) = W(H_i) \cdot p_i(x)dx \quad . \quad (12.1)$$

Was uns hier interessiert, ist die sogenannte *aposteri-Wahrscheinlichkeit* $W(H_i|x)$, was die Wahrscheinlichkeit für H_i bezeichnet, unter der Bedingung, dass der Wert x gemessen wurde, also dass $X \in [x, x + dx]$ gilt. Wir wollen also eine Aussage darüber machen, wie sich die apriori-Wahrscheinlichkeit unter der Bedingung gewisser Beobachtungen (Messungen) ändert. Nach der Bayesschen Formel kann die aposteri-Wahrscheinlichkeit wie folgt berechnet werden,

$$W(H_i|x) = \frac{W(H_i, x)}{W(H_1, x) + W(H_2, x)} \quad .$$

Setzen wir hier die Verbundwahrscheinlichkeiten nach (12.1) ein, so erhalten wir (das Differential dx kürzt sich weg)

$$W(H_i|x) = \frac{W(H_i) \cdot p_i(x)}{W(H_1) \cdot p_1(x) + W(H_2) \cdot p_2(x)} \quad , \quad \text{für } i = 1, 2 \quad . \quad (12.2)$$

Betrachten wir nun formal den Logarithmus des Quotienten $W(H_1|x)/W(H_2|x)$, so folgt,¹⁾

$$\log \frac{p_1(x)}{p_2(x)} = \log \frac{W(H_1|x)}{W(H_2|x)} - \log \frac{W(H_1)}{W(H_2)}$$

Die Terme auf den beiden Seiten dieser Gleichung können positiv oder negativ sein. Die Differenz auf der rechten Seite kann interpretiert werden als die Information, welche wir aus der Beobachtung (Messung) $X = x$ gewinnen. Der Term auf der linken Seite wird definiert als die *Information, welche wir aus der Beobachtung $X = x$ gewinnen, zugunsten der Hypothese H_1 und gegen H_2* . Er wird auch interpretiert als *Gewicht der Evidenz für H_1 aus der Beobachtung $X = x$* (z.B. [Ku59], S. 4) oder als Gewinn an Information, der entsteht, wenn an der Stelle x die Verteilungsdichte $p_2(x)$ durch $p_1(x)$ ersetzt wird.

Ist H_1 wahr, so fallen die Daten mit der Verteilungsdichte $p_1(x)$ an, und der mittlere Informationsgewinn, den wir aus den Messungen $X = x_1, X = x_2, X = x_3, \dots$ gegenüber H_2 gewinnen ist dann

$$\begin{aligned} I(p_1, p_2) &= \int p_1(x) \log \frac{p_1(x)}{p_2(x)} dx \\ &= \int p_1(x) \log \frac{W(H_1|x)}{W(H_2|x)} dx - \int p_1(x) \log \frac{W(H_1)}{W(H_2)} dx \\ &= \int p_1(x) \log \frac{W(H_1|x)}{W(H_2|x)} dx - \log \frac{W(H_1)}{W(H_2)}. \end{aligned}$$

¹⁾Die Basis für den Logarithmus ist für unsere Belange nahezu willkürlich. Üblich ist zum Beispiel der Logarithmus naturalis (Basis e) oder der Logarithmus dualis (Basis 2). Im ersten Fall ergeben sich zuweilen Erleichterungen beim Rechnen, etwa wenn Ableitungen gebildet werden. Bei Verwendung der Basis 2 ist die informations- bzw. codierungstheoretische Interpretation erleichtert.

Anhang A

Fourier–Entwicklungen

A.1 Reihenentwicklung einer symmetrischen Rechteckimpulsfolge

Die Rechteckimpulsfolge mit der (trivialen) Periode T_0 ist definiert durch (Abb.A.1)

$$x_R(t) = \begin{cases} A & : (k - \frac{1}{4})T_0 \leq t < (k + \frac{1}{4})T_0, \quad k \text{ ganzzahlig} \\ -A & : \text{sonst} \end{cases}$$

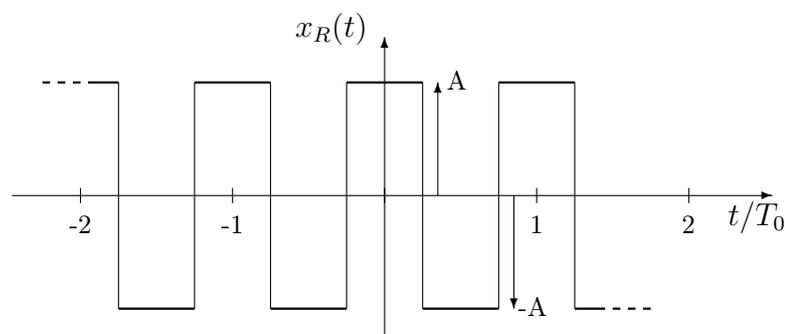


Bild A.1: Symmetrisches Rechtecksignal

Die Reihenentwicklung nach (1.7) und (1.8) liefert

$$\begin{aligned} a_n &= \frac{2}{T_0} \int_{-T_0/2}^{T_0/2} x_R(t) \cdot \cos n\omega_0 t \, dt \\ &= \frac{4A}{n\pi} \cdot \sin \frac{n\pi}{2} \\ &= \begin{cases} \frac{4A}{\pi} \cdot (-1)^{(n-1)/2} \cdot \frac{1}{n} & : n \text{ ungerade} \\ 0 & : \text{sonst} . \end{cases} \end{aligned}$$

Für Ordnungen n gerade gilt $a_n = 0$. Die Fourierkomponenten der symmetrischen Rechteckschwingung fallen mit der Ordnung n linear. Die Entwicklungsterme b_n verschwinden, so dass wir nach (1.11) die folgende Darstellung erhalten,

$$X_R(jn\omega_0) = \frac{a_n}{2} = \begin{cases} \frac{2A}{n\pi} \cdot (-1)^{(n-1)/2} & : n \text{ ungerade} \\ 0 & : \text{sonst} . \end{cases}$$

A.2 Reihenentwicklung eines symmetrischen Dreiecksignals

Die Dreieckimpulsfolge mit der (trivialen) Periode T_0 ist definiert durch (Abb.A.2)

$$x_D(t) = \begin{cases} A \left[1 - \frac{4(t-kT_0)}{T_0} \right] & : kT_0 \leq t < (k + \frac{1}{2})T_0, \quad k \text{ ganzzahlig} \\ A \left[-3 + \frac{4(t-kT_0)}{T_0} \right] & : (k + \frac{1}{2})T_0 \leq t < (k + 1)T_0 \end{cases}$$

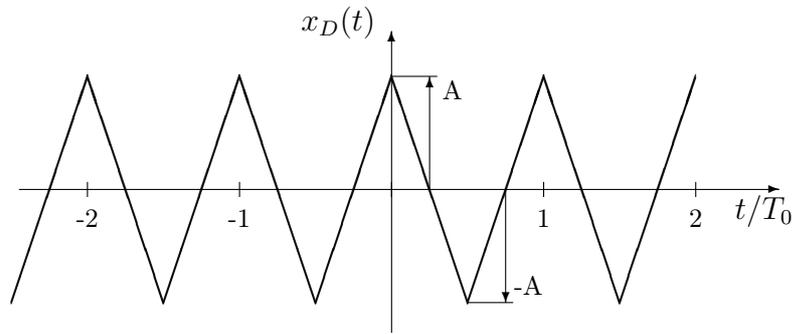


Bild A.2: Symmetrisches Dreiecksignal

Die Reihenentwicklung nach (1.7) und (1.8) liefert

$$\begin{aligned} a_n &= \frac{2}{T_0} \int_{-T_0/2}^{T_0/2} x_D(t) \cdot \cos n\omega_0 t \, dt \\ &= \frac{2}{T_0} \left[\int_{-T_0/2}^0 A(4t/T_0 + 1) \cdot \cos n\omega_0 t \, dt + \int_0^{T_0/2} A(-4t/T_0 + 1) \cdot \cos n\omega_0 t \, dt \right] \\ &= \frac{2}{T_0} \cdot \frac{4A}{T_0} \cdot \left[\int_{-T_0/2}^0 t \cdot \cos n\omega_0 t \, dt - \int_0^{T_0/2} t \cdot \cos n\omega_0 t \, dt \right] \end{aligned}$$

Der Integrand ist schiefsymmetrisch und folglich

$$a_n = -\frac{2}{T_0} \cdot \frac{4A}{T_0} \cdot 2 \cdot \int_0^{T_0/2} t \cdot \cos n\omega_0 t \, dt$$

$$\begin{aligned}
&= -\frac{16A}{T_0^2} \left[\frac{\cos n\omega_0 t}{n^2\omega_0^2} + \frac{t \cdot \sin n\omega_0 t}{n\omega_0} \right]_{0}^{T_0/2} \\
&= -\frac{16A}{T_0^2} \cdot \frac{1}{n^2\omega_0^2} [\cos(n\pi) - 1]
\end{aligned}$$

Somit erhalten wir schließlich,

$$a_n = \begin{cases} \frac{8A}{\pi^2} \cdot \frac{1}{n^2} & : n \text{ ungerade} \\ 0 & : \text{sonst} . \end{cases}$$

Die Fourierkomponenten der symmetrischen Dreiecksschwingung fallen mit der Ordnung n *quadratisch*. Die Entwicklungsterme b_n verschwinden.

A.3 Spektraldichte eines Rechteckfensters

Wir betrachten den Rechteckimpuls der Dauer T , symmetrisch um $T = 0$ gelegen, also $x_P(t) = 1$ für $|t| < T/2$ und $x_P(t) = 0$ für $|t| > T/2$. Die zugehörige Spektraldichte (1.17) ist dann,

$$\begin{aligned}
X_P &= \int_{-\infty}^{+\infty} x_P(t) \cdot e^{-j\omega t} dt \\
&= \int_{-T/2}^{T/2} e^{-j\omega t} dt \\
&= \frac{e^{j\omega T/2} - e^{-j\omega T/2}}{j\omega} .
\end{aligned}$$

Wegen $\sin x = (e^{jx} - e^{-jx})/(2j)$ folgt weiter,

$$X_P(j\omega) = \frac{2 \cdot \sin(\omega T/2)}{\omega} = T \cdot \frac{\sin(\omega T/2)}{\omega T/2} .$$

Dies ist im wesentlichen die Spaltfunktion (Abb.1.6 und Gl. 1.38), hier mit Nulldurchgängen bei $\omega = 2\pi \cdot n/T$, für alle ganzzahligen Werte von n . Die Spektraldichte ist rein reell und folglich das Phasenspektrum, das ist die Phase von $X_P(j\omega)$, Null.

A.4 Zeitfunktion eines spektralen Rechteckfensters

Wir betrachten den spektralen Rechteckimpuls der Breite $2\omega_N$, symmetrisch um $\omega = 0$ gelegen, also $X(j\omega) = 1$ für $|\omega| < \omega_N$ und $X(j\omega) = 0$ für $|\omega| > \omega_N$. Die zugehörige Zeitfunktion erhält man nach (1.19),

$$x(t) = \frac{1}{2\pi} \cdot \int_{-\infty}^{+\infty} X(j\omega) \cdot e^{j\omega t} d\omega$$

$$\begin{aligned} &= \frac{1}{2\pi} \cdot \int_{-\omega_N}^{\omega_N} e^{j\omega t} d\omega \\ &= \frac{1}{2\pi} \cdot \frac{e^{j\omega_N t} - e^{-j\omega_N t}}{jt} . \end{aligned}$$

Wegen $\sin x = (e^{jx} - e^{-jx})/(2j)$ folgt weiter,

$$x(t) = \frac{\omega_N}{\pi} \cdot \frac{\sin \omega_N t}{\omega_N t} .$$

Dies ist, bis auf den Vorfaktor $\frac{\omega_N}{\pi}$, die Spaltfunktion (Abb.1.6 und Gl. 1.38), hier mit Nulldurchgängen bei $t = \pi \cdot n/\omega_N$, für alle ganzzahligen Werte von n .

Anhang B

Programm: Fast-Fourier-Transformation

```
type
  T_FFTdata = real;
  TA_FFT = array of T_FFTdata;

procedure FFT(var ar, ai : TA_FFT;
              const m      : integer;
              const Hin    : Boolean);
// fast fourier tranformation
// Eingaben:
// ar[t], ai[t], t = 1..2^m, Real- bzw. Imaginärteil der Daten
// m = dualer Logarithmus der Blocklaenge 2^m;
// Ausgabe
// ar[t], ai[t], t = 1..2^m, Real- bzw. Imaginärteil der diskreten Fouriertransformation
// ar[1] = Gleichanteil, ai[1] irrelevant
// Normierung:
// Division durch Blocklaenge bei Rücktransformation, d.h., bei Hin = false
var
  j,n,i,nv2,nm1,k,l,le,le1,ip : Integer;
  tr,ti,ur,urr,ui,wr,wi      : extended;
begin
  n:=1;
  for i:=1 to m do n:=2*n;
  if Hin then for i:=1 to n do ai[i]:=-ai[i];
  nv2:=n div 2;
  nm1:=n-1;
  j:=1;
  for i:=1 to nm1 do begin
    if i<j then begin
      tr:=ar[j];
      ti:=ai[j];
      ar[j]:=ar[i];
      ai[j]:=ai[i];
      ar[i]:=tr;

```

```
    ai[i]:=ti;
end; // if i<j ..
k:=nv2;
while k<j do
    begin
        j:=j-k;
        k:=k div 2;
    end;
j:=j+k;
end; // for i ..
for l:=1 to m do begin
    le:=1;
    for i:=1 to l do le:=2*le;
    le1:=le div 2;
    ur:=1;
    ui:=0;
    wr:=cos(pi/le1);
    wi:=sin(pi/le1);
    for j:=1 to le1 do
        begin
            i:=j;
            while i<=n do
                begin
                    ip:=i+le1;
                    tr:=ar[ip]*ur-ai[ip]*ui;
                    ti:=ar[ip]*ui+ai[ip]*ur;
                    ar[ip]:=ar[i]-tr;
                    ai[ip]:=ai[i]-ti;
                    ar[i]:=ar[i]+tr;
                    ai[i]:=ai[i]+ti;
                    i:=i+le;
                end; (*schleife i*)
            urr:=ur*wr-ui*wi;
            ui:=ur*wi+ui*wr;
            ur:=urr;
        end;
    end; // for l ..
    if Not(Hin) then for i:=1 to n do begin
        ar[i]:=ar[i]/n;
        ai[i]:=-ai[i]/n;
    end;
end;
```

Anhang C

Programm: Interpolation

```
type
  TA_series      = array of real;
  TAA_IntFilter = array of array of extended;
  TA_IntFilter  = array of extended;

function Fak(const n : Extended) : Extended;
begin
  If n > 1 then Fak := n * Fak(n-1) else Fak := 1;
end;

procedure Get_L0(const k, n : integer; var L : TAA_IntFilter);
var p, h : extended; i, ll, j : integer;
begin
  SetLength(L,k+1,n+1);
  for i := 0 to n do begin
    if i = (n-1) div 2 then L[0,i] := 1 else L[0,i] := 0;
    if i = (n+1) div 2 then L[k,i] := 1 else L[k,i] := 0;
    for j := 1 to k-1 do begin
      h := 0.5 * (n - 1) + j / k;
      p := 1;
      for ll := 0 to n do
        if not(ll = i) then p := p * (h - ll) / (i - ll);
      L[j,i] := p;
    end; // j
  end; // i
end;

function Get_a(const n, k, i, j, u, m : integer): extended;
var tau, sum : extended; l : integer;
begin
  tau := 0.5 * (n - 1) + j / k;
  sum := 0;
  for l := 0 to n do if l <> i then sum := sum + power(tau-l,m-u-1);
  sum := sum * Fak(u) / Fak(m);
  if odd(u-m) then Get_a := - sum else Get_a := sum;
end;

procedure Get_b(const n, k, i, j, u : integer; var b : TA_IntFilter);
var m1, m2 : integer;
begin
```

```
SetLength(b,u+1);
b[0] := 1;
for m1 := 1 to u do begin
  b[m1] := 0;
  for m2 := 0 to m1-1 do b[m1] := b[m1] + Get_a(n, k, i, j, m1-1, m2) * b[m2];
end;
end;
```

```
Function LnFak(CONST n : Extended) : Extended;
Begin
  If n > 1 then LnFak := Ln(n) + LnFak(n-1)
  else LnFak := 0;
End;
```

```
function Get_InterpolDiff_Error(
  const u, n : integer; const pixf_N_div_f_s : extended): extended;
var i : integer; sum : extended;
begin
  sum := u * ln(n+1)
  + (n+2) * ln(2 * pixf_N_div_f_s)
  - ln(pi * sqrt(n + n + 3))
  - LnFak(n+1);
  for i := 0 to (n-1) div 2 do sum := sum + 2 * ln(0.5+i);
  Get_InterpolDiff_Error := exp(sum);
end;
```

```
procedure Get_LAGRANGE_Factors(const u, k, n      : integer;
                             const pixf_N_div_f_s : extended;
                             var L              : TAA_IntFilter);
// L : FIR Filter for interpolation and differentiation
var i, j : integer; fac : extended; b : TA_IntFilter;
begin
  Get_L0(k,n,L);
  if u > 0 then begin
    fac := exp(u*ln(0.5 / pixf_N_div_f_s));
    for j := 1 to k-1 do
      for i := 0 to n do begin
        Get_b(n, k, i, j, u, b);
        L[j,i] := L[j,i] * b[u] * fac;
      end;
    end; // u > 0
end;
```

```
procedure Interpol_LAGRANGE(
  const x      : TA_series;
  const TT, u, k, n : integer;
  const pixf_N_div_f_s : extended;
  var x_int    : U_SPO0.TA_series;
  var TT_int   : integer);
// LAGRANGE-Interpolation
// x : original series
// TT : number data in x
// u : differentiation u times
// for u > 8 there are remarkable rounding errors
// k : upsampling factor
```

```

// n : degree of interpolation/differentiation polynom
// x_int : interpolated series
// TT_int : number data in x_int
var t, j, m, nh : integer;
    sum          : extended;
    L            : TAA_IntFilter;
begin
  Get_LAGRANGE_Factors(u,k,n,pixf_N_div_f_s,L);
  TT_int := 1 + (TT-1) * k;
  SetLength(x_int,TT_int);
  nh := (n-1) div 2;
// at the left no interpolation/differentiation
  for t := 0 to nh - 1 do begin
    x_int[t*k] := x[t];
    for j := 1 to k-1 do x_int[t*k + j] := x_int[t*k + j - 1];
  end;
// interpolation/differentiation polynomial degree n
  for t := nh to TT - 2 - nh do begin
    x_int[t*k] := x[t];
    for j := 1 to k-1 do begin
      sum := 0;
      for m := 0 to n do sum := sum + L[j,m] * x[t-nh+m];
      x_int[t*k + j] := sum;
    end;
  end;
// at the right no interpolation/differentiation
  for t := TT - 1 - nh to TT - 2 do begin
    x_int[t*k] := x[t];
    for j := 1 to k-1 do x_int[t*k + j] := x_int[t*k + j - 1];
  end;
  x_int[(TT-1) * k] := x[TT-1];

// zero at very left and right
// linear interpolation at sample points
if (u > 0) then begin
  for t := 0 to k * nh - 1 do x_int[t] := 0;
  for t := k * nh to TT_int - 1 - k * nh do
    if t mod k = 0 then x_int[t] := 0.5 *(x_int[t-1] + x_int[t+1]);
  for t := TT_int - k * nh to TT_int - 1 do x_int[t] := 0;
end;
end;

```


Anhang D

Grundlagen der Fehlerrechnung

D.1 Problemstellung

In einem Experiment soll eine physikalische Größe gemessen werden, beispielsweise die Länge eines Stabes oder die Dauer einer Schwingung. Führen wir solch eine mehrmals Messung aus, so ist einerseits ein jeder Messwert

$$x(1), \dots x(n), \dots x(N) \quad (\text{D.1})$$

mit einem gleichbleibenden systematischen Fehler versehen, etwa weil die Skale des Messgerätes nur mit begrenzter Genauigkeit geeicht wurde. Andererseits sind die Bedingungen bei wiederholten Messungen niemals vollkommen gleichbleibend, was zu zufälligen Messwertschwankungen führt. Diese können aber nur dann wahrgenommen werden, wenn die Präzision der verwendeten Messvorrichtung groß genug. Die zufälligen Abweichungen werden also umso augenfälliger, je genauer wir messen. Allgemein stellt sich nun die Frage nach der Beziehung zwischen einem (gegebenenfalls) variierenden Messwert und dem uns unbekanntem „wahren“ Wert, x_w . Kann dieser Zusammenhang für eine bestimmte Messanordnung auch niemals vollständig aufgeklärt werden, so können doch einige allgemeingültige Aussagen zur *Unsicherheit* bei der Angabe von Messergebnissen gemacht werden, und auch darüber, wie sich Messfehler fortpflanzen, wenn mit ungenauen Messwerten weitergerechnet wird. Mit diesem Problemfeld beschäftigt sich die *Fehlerrechnung*.

Letztliches Ziel einer jeden direkten oder indirekten Messung ist es, einen Wert x und einen möglichst kleinen Unsicherheitsbereich Δx zu finden, so dass x mit großer Wahrscheinlichkeit, sagen wir 95%, im Intervall $x_w - \Delta x \dots x_w + \Delta x$ liegt. Wir schreiben dafür

$$x_w = x \pm \Delta x \quad (95\%) .$$

Ohne solch eine Fehlerangabe sind Messwerte wertlos.

Bei dieser Begriffsbildung gehen wir davon aus, dass es den wahren Wert tatsächlich gibt. Aber schon das einfache Beispiel eines realen Stabes lehrt uns, dass dies streng genommen unzulässig ist. Nur der mathematischen Idealisierung als Zylinder kann eine *bestimmte* Länge zugeordnet werden. Auf immer kleineren Skalen entspricht ein realer Stab aber immer weniger seinem Ideal. Wenn wir dennoch von einer wahren Größe ausgehen, so verstehen wir darunter eine über die geometrischen Unregelmäßigkeiten gemittelte Größe.

Eine elementare Einführung in die Fehlerrechnung gibt [TAY88]. Etwas weiterführend ist [BAR89].

D.2 Einteilung der Fehler

Ein Messfehler $\Delta x(n) = |x(n) - x_w|$ kann auf recht unterschiedliche Weise entstehen und ist deshalb im Rahmen einer Fehlerrechnung auch verschieden zu behandeln. Wir unterscheiden drei Fehlerklassen und schreiben deshalb

$$\Delta x = \Delta_g x + \Delta_s x + \Delta_z x .$$

Grober Fehler, $\Delta_g x$:

Dieser Fehler hat beispielsweise ein grob falsches Ablesen des Messinstrumentes oder eine Verwechslung zur Ursache. Man erkennt ihn zumeist als *Ausreißer* in der Messreihe (D.1). Bei gewissenhaftem Experimentieren sind grobe Fehler in der Regel vermeidbar. Ausreißer werden durch eine kritische Hinterfragung der Plausibilität der Messwerte erkannt. Bei der Datenauswertung werden sie nicht berücksichtigt.

Systematischer Fehler, $\Delta_s x$:

Dieser Fehler ist unter gleichen Versuchsbedingungen bei einem jeden Messwert $x(n)$ gleich, sowohl im Betrag, d.h. $\Delta_s x = \text{const.}$, wie auch in der Richtung, d.h. entweder $x(n) - x_w > 0$ oder $x(n) - x_w < 0$. Seinen genauen Wert kennen wir allerdings nicht. Systematische Fehler können beispielsweise durch eine Nullpunktverschiebung der Messskale oder eine Eichungenauigkeit der Messapparatur verursacht werden, also durch Unvollkommenheiten der Messgeräte und Maßverkörperungen. Ändern sich die Versuchsbedingungen, so ändert sich im Allgemeinen auch der systematische Fehler. Wird darüber hinaus mit den Messwerten weitergerechnet, so können auch verwendete Näherungsformeln zum systematischen Fehler beitragen.

Die Abschätzung des systematischen Fehlers verlangt eine gründliche Analyse der Messapparatur, des –verfahrens und der Datenauswertung. Dafür sind viele Erfahrungen nötig. Werden mehrere unabhängige Apparaturen oder Verfahren angewandt, so sind einige systematische Fehleranteile abschätzbar. Der Messwert kann somit teilweise korrigiert werden. Bei teuren Präzisions–Messinstrumenten ist der durch die Apparatur bedingte systematische Fehler vergleichsweise klein.

Der nicht erfassbare Anteil des systematischen Fehlers wird schließlich durch den sogenannten *absoluten Größtfehler* charakterisiert. Darunter versteht man den geschätzten Betrag der maximal möglichen Abweichung des Messwertes vom wahren Wert. Er kann sich aus verschiedenen Anteilen additiv zusammensetzen:

- größter Fehler des Messinstrumentes (vom Hersteller angegebene oder geschätzte Präzision)
- geschätzter größter Fehler durch den Experimentator und durch Einflüsse aus der Umwelt
- größter Fehler infolge der Anwendung von Näherungsformeln bei der Datenauswertung
- geschätzter größter Fehler des Messverfahrens.

Zufälliger Fehler, $\Delta_z x$:

Unter vermeintlich gleichen Versuchsbedingungen ist dieser Fehler bei verschiedenen Messwerten sowohl im Betrag wie auch in der Richtung unterschiedlich. Er stellt also eine zufällige

Abweichung des Messwertes $x(n)$ von x_w dar. Seine Ursachen sind zum einen Unvollkommenheiten des Messobjektes und zum anderen zufällige äußere Umstände, die vom Experimentator immer nur teilweise beeinflusst werden können. Der zufällige Fehler kann durch die mehrfache Wiederholung der Messung und einer anschließenden statistischen Analyse der Messreihe quantifiziert werden. Dies wird weiter unter noch ausführlich dargelegt.

D.3 Abschätzung und Fortpflanzung des Größtfehlers

D.3.1 Allgemeine Betrachtung

In der Regel wird eine physikalische Größe nicht direkt gemessen, sondern aus Messwerten anderer Größen berechnet, beispielsweise die Geschwindigkeit eines Körpers aus dem Verhältnis von zurückgelegtem Weg und zugehöriger Zeitspanne. Allgemein gehen wir von K physikalischen Größen

$$x_1, \dots, x_k, \dots, x_K$$

aus, deren Messwerten

$$x_{M1}, \dots, x_{Mk}, \dots, x_{MK}$$

und zugehörigen Größtfehlern

$$\Delta x_1, \dots, \Delta x_k, \dots, \Delta x_K .$$

Ist die interessierende Größe y als Funktion

$$y = f(x_1, \dots, x_k, \dots, x_K)$$

darstellbar, dann ergibt sich aus den Messwerten der mittelbare Messwert für y ,

$$y_M = f(x_{M1}, \dots, x_{Mk}, \dots, x_{MK}) .$$

Es stellt sich nun die Frage, wie sich die Größtfehler Δx_k auf den Größtfehler von y fortpflanzen.

Wir ermitteln diesen Größtfehler Δy , indem wir zunächst die Funktion f an den Stellen x_{Mk} in eine Taylor-Reihe nach Potenzen von Δx_k entwickeln, und schon nach dem linearen Glieder abbrechen,

$$f(x_{M1} + \Delta x_1, \dots, x_{MK} + \Delta x_K) \approx f(x_{M1}, \dots, x_{MK}) + \sum_{k=1}^K \left. \frac{\partial f}{\partial x_k} \right|_{x_{M1}, \dots, x_{MK}} \Delta x_k .$$

Um den Größtfehler Δy zu erhalten, müssen wir die in der Praxis mögliche Kompensation der Einzelfehler ausschließen. Deshalb sind die Absolutbeträge der partiellen Ableitungen zu verwenden. Wir erhalten somit die gesuchte Gleichung für die Fortpflanzung des Größtfehlers:

$$\Delta y = \sum_{k=1}^K \left\| \left. \frac{\partial f}{\partial x_k} \right|_{x_{M1}, \dots, x_{MK}} \right\| \Delta x_k \quad (\text{D.2})$$

Die Formel verdeutlicht, dass sich die einzelnen Messfehler Δx_k sehr unterschiedlich auf die berechnete Größe y_M auswirken können. Wie stark der Einfluss ist, hängt demzufolge von der Größe der jeweiligen partiellen Ableitung an der Stelle x_{Mk} ab.

D.3.2 Beispiel

Wir verdeutlichen nun die Arbeit mit dem Fehlerfortpflanzungsgesetz (D.2) an einem Beispiel.

Zwischen der Fadenlänge L und der Schwingungsdauer T eines sogenannten *mathematischen Pendels*¹⁾ besteht die Beziehung

$$T = 2\pi\sqrt{\frac{L}{g}} .$$

Darin ist g die *Fallbeschleunigung* (Tabellenwert: $9,81 \text{ ms}^{-2}$), die wir nach Umstellung der obigen Gleichung erhalten,

$$g = 4\pi^2 \frac{L}{T^2} .$$

Zur Messung von L verwenden wir einen Maßstab mit Millimeterteilung, dessen nicht erfasster systematischer Fehler $0,5 \text{ mm}$ je Meter beträgt. Darüber hinaus ist ein Ablesefehler von $0,5 \text{ mm}$ für jeden Endpunkt zu berücksichtigen. Die Messung möge einen Wert von $L_M = 974 \text{ mm}$ ergeben. Der Größtfehler ist dann mit

$$\Delta L = \frac{0,974 \text{ m}}{1 \text{ m}} \times 0,5 \text{ mm} + 2 \times 0,5 \text{ mm} \approx 1,5 \text{ mm}$$

hinreichend genau abgeschätzt.

Für die Zeitmessung verwenden wir eine mechanische Stoppuhr mit einem kleinsten ablesbaren Zeitintervall von $0,2 \text{ s}$. Da es sich um einen periodischen Vorgang handelt, stoppen wir besser nicht nur die Dauer einer Schwingungsperiode, sondern gleich die von 100. Dabei ergebe sich ein Wert von $100T_M = 196,8 \text{ s}$, das heisst, $T_M = 1,968 \text{ s}$.

Es sei weiterhin durch Vergleich mit einer hoch präzisen Quarzuhr bekannt, dass die Stoppuhr in $1 \text{ h} = 3600 \text{ s}$ um etwa $25,6 \text{ s}$ nachgeht. Die Kenntnis dieses erfassbaren systematischen Fehleranteils ermöglicht eine Korrektur der Zeitmessung:

$$T_C = \frac{3600}{3600 - 25,6} T_M = 1,982 \text{ s} .$$

Der Größtfehler der Zeitmessung wird offenbar nahezu ausschließlich durch das kleinste ablesbare Zeitintervall und die Reaktionsfähigkeit des Experimentators (zufälliger Stoppfehler) bestimmt. Beträgt bei einer Messung über 100 Schwingungsperioden der Gesamtfehler $0,2 \text{ s} + 0,2 \text{ s} = 0,4 \text{ s}$ (Start- und Stoppfehler), so folgt für den Größtfehler bei der Bestimmung der Dauer einer Schwingung

$$\Delta T = \frac{0,4 \text{ s}}{100} = 0,004 \text{ s} .$$

Für die Schwerebeschleunigung ergibt sich also der mittelbare Messwert

$$g_M = 4\pi^2 \frac{L_M}{T_C^2} = 9788 \text{ mm s}^{-2} .$$

Zur Bestimmung der Fortpflanzung der Größtfehler nach (D.2) benötigen wir die partiellen Ableitungen von $g = f(L, T)$,

$$\begin{aligned} \frac{\partial f}{\partial L} &= 4\pi^2 \frac{1}{T^2} \quad \longrightarrow \quad \left\| \frac{\partial f}{\partial L} \Big|_{L_M, T_C} \right\| = 4\pi^2 \frac{1}{T_C^2} \approx 10 \text{ s}^{-2} \\ \frac{\partial f}{\partial T} &= -8\pi^2 \frac{L}{T^3} \quad \longrightarrow \quad \left\| \frac{\partial f}{\partial T} \Big|_{L_M, T_C} \right\| = 8\pi^2 \frac{L_M}{T_C^3} \approx 9877 \text{ mm s}^{-3} . \end{aligned}$$

¹⁾Ein mathematisches Pendel ist die Idealisierung eines realen, sogenannten *physikalischen* Pendels. Beispielsweise ist das Pendel einer großen Standuhr in guter Näherung als mathematisches Pendel beschreibbar.

Damit erhalten wir für den Größtfehler bei der Ermittlung der Schwerebeschleunigung

$$\Delta g = \left\| \frac{\partial f}{\partial L} \Big|_{L_M, T_C} \right\| \Delta L + \left\| \frac{\partial f}{\partial T} \Big|_{L_M, T_C} \right\| \Delta T \approx 55 \text{ mm s}^{-2} .$$

Das Ergebnis lautet daher

$$g = (9788 \pm 55) \text{ mm s}^{-2} \approx (9,79 \pm 0,06) \text{ m s}^{-2} .$$

Die Fehlerrechnung führt nun zu der Aussage, dass der unbekannte wahre Wert g_w mit Sicherheit im Intervall zwischen den Werten $9,73 \text{ m s}^{-2}$ und $9,85 \text{ m s}^{-2}$ zu finden ist. Mit anderen Worten, $g = 9,79 \text{ m s}^{-2}$ unterscheidet sich von g_w mit Sicherheit um weniger als $0,06 \text{ m s}^{-2}$.

D.4 Statistische Analyse und Fortpflanzung des zufälligen Fehlers

D.4.1 Fehlerstatistik

Die statistische Analyse des zufälligen Fehlers einer gegebenen Messreihe (D.1) hat zur Voraussetzung, dass die Messwerte *unabhängig* gewonnen werden. Praktisch heisst dies, jeder Messvorgang muss von einer möglichst willkürlichen Einstellung der Messapparatur ausgehen. Folglich sind alle relevanten Einstellungen für eine jede Messung erneut vorzunehmen.²⁾ Letztlich soll mit diesem Vorgehen gewährleistet werden, dass die Messreihe *statistisch unabhängig* ist und somit dem statistischen Modell einer *mathematischen Stichprobe* möglichst nahe kommt. Alle folgenden Aussagen basieren auf dieser Annahme.

Zur Vereinfachung der Schreibweise gehen wir im Folgenden davon aus, dass der systematische Fehler verschwindet, $\Delta_s x = 0$. Andernfalls wäre x_w durch $x_w + \Delta_s x$ zu ersetzen. Darüber hinaus nehmen wir zunächst an, dass wir viele Werte in unserer Messreihe (D.1) gesammelt haben, $N \gtrsim 30$. Die Auflösung unserer Messapparatur möge so groß sein, dass viele verschiedene Messwerte erhalten werden, wenigstens aber zwei. Zur statistischen Analyse des zufälligen Fehlers sind dann nacheinander folgende Größen zu berechnen:

$$\text{Mittelwert} \quad \bar{x} = \frac{1}{N} \sum_{n=1}^N x(n) \quad (\text{D.3})$$

$$\text{Standardabweichung} \\ \text{der Einzelmessung} \quad s_x = \sqrt{\frac{1}{N-1} \sum_{n=1}^N (x(n) - \bar{x})^2} \quad (\text{D.4})$$

$$\text{Standardabweichung} \\ \text{des Mittelwertes} \quad s_{\bar{x}} = s_x / \sqrt{N} . \quad (\text{D.5})$$

²⁾ Damit kann unter Umständen eine Gegebenheit, die zunächst einen systematischen Fehleranteil verursacht, nun zum zufälligen Fehler beitragen, was sie einer statistischen Analyse zugänglich macht. Aus dieser Sicht sollten also möglichst viele Einstellungen verändert werden. Allerdings kann dadurch der Messvorgang auch sehr aufwendig werden.

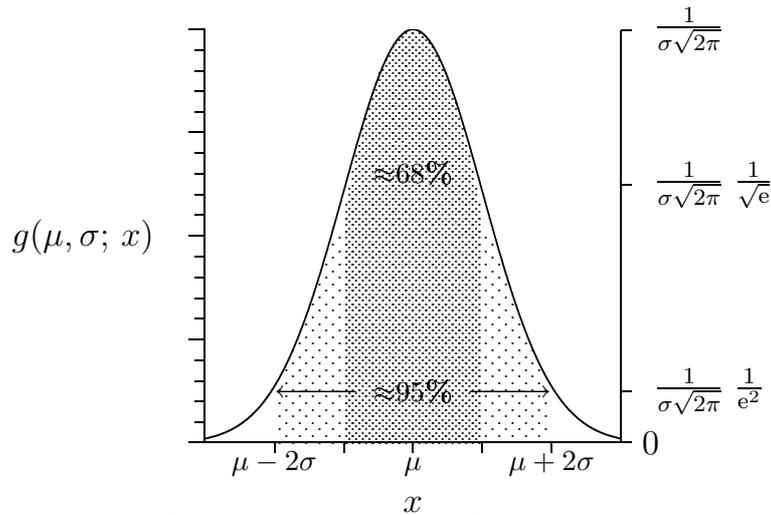


Bild D.1: Dichte der Normalverteilung (GAUSSsche Glockenkurve) mit Mittelwert μ und Standardabweichung σ

Das Quadrat s^2 der Standardabweichung s ist die sogenannte *Streuung* oder *Dispersion*. Der zufällige Fehler unserer Messung kann nun wie folgt charakterisiert werden: *Die Wahrscheinlichkeit, dass eine Einzelmessung im Intervall $x_w \pm \Delta x$ liegt, beträgt etwa*

$$p(\Delta x) = \int_{x_w - \Delta x}^{x_w + \Delta x} g(x_w, s_x; x) \, dx \quad . \quad (D.6)$$

Darin bezeichnet

$$g(\mu, \sigma; x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (D.7)$$

die *Wahrscheinlichkeitsdichte der Normalverteilung* mit dem Mittelwert μ und der Standardabweichung σ (Abb. D.1).³⁾ An den Stellen $x = \mu \pm \sigma$ hat sie gerade ihre Wendepunkte. Die Wahrscheinlichkeit (D.6) kann man sich als Fläche unter der Glockenkurve veranschaulichen. Mit wachsender Intervalllänge $2\Delta x$ strebt sie gegen den Wert 1. Häufig wählt man Δx als ein niedriges ganzzahliges Vielfaches der Standardabweichung σ . Dafür erhält man die folgenden Wahrscheinlichkeiten:

Δx	$p(\Delta x) \times 100\%$
σ	$\approx 68\%$
2σ	$\approx 95\%$
3σ	$\approx 99\%$

Wir nehmen nun an, dass nicht nur eine, sondern mehrere Messreihen vorliegen, beispielsweise von den verschiedenen studentischen Praktikumsgruppen. Dann kann zu einer jeden Messreihe der Mittelwert nach (D.3) berechnet werden. Diese Mittelwerte werden ihrerseits auch zufällig schwanken, und zwar ebenfalls näherungsweise normalverteilt, was man aus theoretischen

³⁾Im Zusammenhang mit der Behandlung von Messfehlern wurde die Normalverteilung von den Mathematikern ADRIEN MARIE LEGENDRE (1752–1833) und CARL FRIEDRICH GAUSS (1777–1855) eingeführt. Sie spielt in der gesamten Statistik eine fundamentale Rolle. Wegen ihrer markanten Form wird die zugehörige Verteilungsdichte auch *Glockenkurve* genannt.

Immer dann, wenn der zufällige Fehler einer Einzelmessung als additive Überlagerung einer Vielzahl von zufälligen, untereinander unabhängigen und in ihrer Variabilität beschränkten Einflüssen (Zufallsgrößen) angesehen werden kann, ist der resultierende Fehler normalverteilt, was durch den *Zentralen Grenzwertsatz* der Wahrscheinlichkeitsrechnung genauer beschrieben wird. Dies ist umso bemerkenswerter, als wir hierbei die Verteilung der einzelnen Einflüsse oder gar ihren physikalischen Ursprung nicht zu kennen brauchen.

Überlegungen findet. Die entsprechende Standardabweichung kann aber schon aus allein einer Messreihe nach Gl. (D.5) geschätzt werden. Darüber hinaus ist der Mittelwert von diesen Mittelwerten der einzelnen Messreihen gerade gleich dem gesuchten wahren Wert, x_w .⁴⁾ Man kann somit von einer Normalverteilung $g(x_w, s_{\bar{x}}; x)$ des Mittelwertschätzers (D.3) ausgehen. Folglich beträgt die Wahrscheinlichkeit, dass der Abstand des aus allein einer Messreihe berechneten Mittelwertes \bar{x} vom wahren Wert x_w kleiner als Δx ist, etwa

$$\int_{x_w - \Delta x}^{x_w + \Delta x} g(x_w, s_{\bar{x}}; x) dx .$$

Damit können wir, ohne x_w zu kennen, beispielsweise Folgendes sagen: *Mit der Wahrscheinlichkeit von etwa 95% liegt \bar{x} dichter als $\Delta x = 2s_{\bar{x}}$ am wahren Wert.* Man schreibt

$$x = \bar{x} \pm 2s_{\bar{x}} \quad (95\%) \quad (\text{D.8})$$

und sagt dafür auch: *Mit der Irrtumswahrscheinlichkeit $\alpha = 1 - 0,954 = 0,046$ liegt \bar{x} im Intervall $x_w \pm 2s_{\bar{x}}$.*

Der Fehler kann auch relativ angegeben werden,

$$x = \bar{x} \left(1 \pm \frac{2s_{\bar{x}}}{\bar{x}} \right) \quad (95\%) ,$$

oder relativ in Prozent,

$$x = \bar{x} , \quad \pm \frac{2s_{\bar{x}}}{\bar{x}} \times 100\% \quad (95\%) .$$

Die beiden Werte $\bar{x} \pm 2s_{\bar{x}}$ heißen auch *Vertrauensgrenzen des Mittelwertes zum Signifikanzniveau $1 - \alpha = 0,954$.*

Tritt auch ein systematischer Fehler $\Delta_s x$ auf, so liegt \bar{x} mit der Wahrscheinlichkeit 95% dichter als $2s_{\bar{x}}$ an $x_w + \Delta_s x$. Ein systematischer Fehler wird durch die beschriebene statistische Analyse *nicht* erfasst.

Will man mit gleicher Sicherheit, beispielsweise den üblichen⁵⁾ 95%, den Fehlerbereich weiter einschränken, so muss man nach Gl. (D.5) die Anzahl N der Messwerte vergrößern. Allerdings fällt $s_{\bar{x}}$ nur proportional zu $1/\sqrt{N}$. Wollen wir beispielsweise $s_{\bar{x}}$ halbieren, so müssen wir gleich 4-mal so viele Messwerte erfassen. Steigt der dafür notwendige experimentelle Aufwand auch enorm, so kann doch zumindest prinzipiell mit wachsender Anzahl N der Vertrauensbereich des Mittelwertes beliebig verringert werden. Eine Vergrößerung von N ist aber spätestens dann nicht mehr sinnvoll, wenn $s_{\bar{x}} = s_x/\sqrt{N}$ deutlich kleiner als der systematische Fehler $\Delta_s x$ ist. Es sollte also $s_x/\sqrt{N} \gtrsim \Delta_s x/10$ gelten, woraus

$$N \lesssim \frac{100 s_x^2}{\Delta_s x^2} \quad (\text{D.9})$$

⁴⁾ Stimmt der Mittelwert eines Schätzers einer Größe mit dieser Größe überein, so nennt man den Schätzer *erwartungstreu*. \bar{x} in Gl. (D.3) ist also ein erwartungstreuer Schätzer für die gesuchte Größe x_w .

Der Schätzer s_x in Gl. (D.4) für die „wahre“ Standardabweichung σ_x der Einzelmessung ist ebenfalls erwartungstreu. Würden wir hier allerdings anstelle des Faktors $1/(N-1)$ unter der Wurzel $1/N$ setzen, so wäre dieser Schätzer nicht mehr erwartungstreu. Letztlich rührt dies daraus, dass bei der Berechnung von s_x der geschätzte Mittelwert, \bar{x} , anstelle des unbekannt wahren Mittelwertes, x_w , verwendet wird. Für sehr lange Messreihen, $N \gtrsim 30$, ist der Unterschied dieser beiden Vorfaktoren allerdings praktisch zumeist bedeutungslos.

⁵⁾ Die Wahl des Signifikanzniveaus unterliegt immer einer gewissen Willkür und muss letztlich aus dem Kontext der Messaufgabe begründet werden. Man wird es umso größer wählen, je folgenreicher eine große Abweichung des angegebenen Wertes vom wahren Wert ist. In der Physik und der Vermessungstechnik rechnet man zuweilen auch mit der statistischen Sicherheit von 68%, in der Industrie sind 95% üblich, in der Biologie 99,7%.

folgt. Generell sollten eher weniger Messwerte verarbeitet werden, die aber mit großer experimenteller Sorgfalt gewonnen wurden, als viele sorglos gewonnene. Insbesondere sollte großes Augenmerk auf die Erkennung bzw. Vermeidung grober Messfehler gelegt werden. Ein einzelner grober Fehler wird zwar für $N \rightarrow \infty$ ein immer geringeres statistisches Gewicht bekommen, allerdings muss man, je nach Größe des Ausreißers, unter Umständen äußerst lange Messreihen haben, um seinen Einfluß auszumerzen.

Wir hatten bislang für die statistische Analyse vorausgesetzt, dass möglichst viele Messwerte vorliegen, jedenfalls $N \gtrsim 30$. Ist die Gewinnung so vieler Messdaten zu aufwendig oder gar unmöglich, oder wollen wir die Relation (D.9) einhalten, so können die wenigen verfügbaren Daten dennoch einer statistischen Analyse unterzogen werden. Dazu geht man zunächst analog vor, indem man die Größen (D.3), (D.4) und (D.5) berechnet. Der Mittelwertschätzer ist auch hier wieder normalverteilt. Allerdings machen sich für $N \lesssim 30$ die zufälligen Schwankungen von \bar{x} im Zusammenspiel mit den Schwankungen unseres Schätzers für die Standardabweichung $s_{\bar{x}}$ bemerkbar, was letztlich die Angabe der Vertrauensgrenzen $\bar{x} \pm s_{\bar{x}}$ beeinflusst. Konkret führt dies dazu, dass die Verteilung der Größe $(\bar{x} - x_w)/s_{\bar{x}}$ für $N \lesssim 30$ nicht länger durch eine Normalverteilung ausreichend gut angenähert werden kann.⁶⁾ Die Vertrauensgrenzen des Mittelwertes müssen nun korrigiert werden. Zum Datenumfang N und zum Signifikanzniveaus $1 - \alpha$ erhalten wir

$$x = \bar{x} \pm t_{\alpha/2; N-1} s_{\bar{x}} \quad ((1 - \alpha) \times 100\%) . \quad (\text{D.10})$$

Damit tritt beispielsweise für $\alpha = 0,05$ an die Stelle von Gl. (D.8) nun

$$x = \bar{x} \pm t_{0,025; N-1} s_{\bar{x}} \quad (95\%) . \quad (\text{D.11})$$

Die Faktoren $t_{\alpha/2; N-1}$ sind für einige typische Werte von $1 - \alpha$ und N der Tabelle D.1 zu entnehmen. Beispielsweise ist für $N = 10$ in (D.11) $t_{0,025; 9} = 2,26$ einzusetzen.⁷⁾

⁶⁾Die zufällige Größe $(\bar{x} - x_w)/s_{\bar{x}}$ ist für $N < \infty$ nicht normalverteilt mit der Dichte $g(0, 1; x)$, sondern t -verteilt mit $N - 1$ Freiheitsgraden. Die entsprechende Dichte ist

$$t_{N-1}(x) = \frac{1}{\sqrt{\pi(N-1)}} \frac{\Gamma(N/2)}{\Gamma((N-1)/2)} \left(1 + \frac{x^2}{N-1}\right)^{-N/2} .$$

Darin bezeichnet Γ die Gamma-Funktion, für die Folgendes gilt: $\Gamma(N) = (N-1)!$ sowie $\Gamma(N+1/2) = 2N! \sqrt{\pi}/(N! 2^{2N})$ falls $N = 1, 2, 3, \dots$. Die t -Verteilung wird auch *Student-Verteilung* genannt, nach dem englischen Statistiker W. S. GOSSET, der diese Verteilung als erster auf empirischem Weg fand und 1901 unter dem Pseudonym „Student“ publizierte. Die Dichte der t -Verteilung ähnelt der Gaußschen Glockenkurve $g(0, 1; x)$. Insbesondere ist sie auch symmetrisch um den Mittelwert 0. Die Standardabweichung der t -Verteilung mit dem Freiheitsgrad $N - 1 > 2$ ist $\sqrt{(N-1)/(N-3)}$. Für kleine Werte von N fällt die Dichte der t -Verteilung langsamer als die der Normalverteilung mit $x \rightarrow \pm\infty$ ab. Für $N \rightarrow \infty$ geht $t_{N-1}(x)$ in die Dichte der Standard-Normalverteilung $g(0, 1; x)$ über.

⁷⁾Die Größe $t_{\alpha/2; N-1}$ in Gl. (D.10) ist das sogenannte $\alpha/2$ -Quantil der t -Verteilung mit dem Freiheitsgrad $N - 1$, das die folgende Gleichung erfüllt,

$$\frac{\alpha}{2} = \int_{t_{\alpha/2; N-1}}^{+\infty} t_{N-1}(x) dx .$$

Für unendlich große Messreihen, $N \rightarrow \infty$, geht die t -Verteilung in die Normalverteilung über, und folglich sind die Quantile $t_{\alpha/2; \infty}$ (s. letzte Spalte in der Tabelle) aus der Normalverteilung (D.7) zu berechnen,

$$1 - \alpha = \int_{-t_{\alpha/2; \infty}}^{t_{\alpha/2; \infty}} g(0, 1; x) dx .$$

Für das übliche Signifikanzniveau $\alpha = 0,05$ und für $N > 30$ ist die relative Abweichung zwischen $t_{\alpha/2; N}$ und $t_{\alpha/2; \infty}$ kleiner als 5%, weshalb wir hier in guter Näherung mit der Normalverteilung der rechnen können.

Tabelle D.1: Quantile $t_{\alpha/2;N-1}$ der t -Verteilung für ausgewählte Werte des Signifikanzniveaus $1 - \alpha$ und der Messdatenanzahl N

	$N \rightarrow$	5	10	15	20	30	60	120	∞
$1 - \alpha \downarrow$	$t_{\alpha/2;N-1}$								
0,6827		1,14	1,06	1,04	1,03	1,02	1,01	1,00	1,00
0,950		2,78	2,26	2,15	2,09	2,05	2,00	1,98	1,96
0,9545		2,88	2,32	2,20	2,14	2,09	2,05	2,02	2,00
0,990		4,66	3,28	3,00	2,88	2,77	2,68	2,63	2,59
0,9974		6,97	4,23	3,74	3,54	3,36	3,20	3,13	3,00

D.4.2 Fortpflanzung zufälliger Fehler

Zu jeder Größe x_k wird eine Messreihe (D.1) aufgenommen. Alle Messwerte $x_k(n)$, $n = 1, \dots, N_k$ und $k = 1, \dots, K$, seien vollkommen unabhängig voneinander, sowohl innerhalb einer wie auch zwischen verschiedenen Messreihen.

Nach Gleichung (D.3) wird nun der Mittelwert, \bar{x}_k , berechnet, nach (D.4) und (D.5) die Standardabweichung des Mittelwertes $s_{\bar{x}_k}$. Wir nehmen des Weiteren an, dass die Bereiche $\bar{x}_k \pm \Delta x_k$, mit

$$\Delta x_k = \Delta x_{\max,k} + t_{\alpha/2;N-1} s_{\bar{x}_k} \quad , \quad (\text{D.12})$$

klein genug sind, so dass sich y hierin näherungsweise nur linear mit x_k ändert.⁸⁾ Dann ist der Mittelwert der gesuchten Größe näherungsweise durch

$$\bar{y} = y(\bar{x}_1, \dots, \bar{x}_k, \dots, \bar{x}_K) \quad (\text{D.13})$$

gegeben.⁹⁾ Wie sich die zufälligen Fehler der einzelnen Messgrößen x_k auf y fortpflanzen, beschreibt das *Gauss'sche Fehlerfortpflanzungsgesetz*: Die Standardabweichung von \bar{y} ist

$$s_{\bar{y}} = \sqrt{\sum_{k=1}^K a_k^2 s_{\bar{x}_k}^2} \quad . \quad (\text{D.14})$$

Darin sind

$$a_k = \left. \frac{\partial y(x_1, \dots, x_k, \dots, x_K)}{\partial x_k} \right|_{\bar{x}_1, \dots, \bar{x}_k, \dots, \bar{x}_K}$$

die Werte der partiellen Ableitungen der Funktion y nach den Variablen x_k an der Stelle der Mittelwerte. Je größer $|a_k|$ ist, umso stärker ist der Einfluss von Fehlern in der Messung von

⁸⁾ Wir sagen, dass sich y in einem Bereich $\bar{x}_k \pm \Delta x_k$ linear ändert, wenn für $\Delta x_k = -\Delta x_k \dots \Delta x_k$ folgende Näherung gilt:

$$y(\bar{x}_1 + \Delta x_1, \dots, \bar{x}_K + \Delta x_K) \approx y(\bar{x}_1, \dots, \bar{x}_K) + \sum_{k=1}^K \left. \frac{\partial y}{\partial x_k} \right|_{\bar{x}_1, \dots, \bar{x}_K} \Delta x_k \quad .$$

Dies kann als Entwicklung der Funktion y an der Stelle $\bar{x}_1, \bar{x}_2, \dots, \bar{x}_K$ in eine Taylor-Reihe nach Δx_k angesehen werden, die nach dem in Δx_k linearen Term abgebrochen wurde.

⁹⁾ Der Schätzer (D.13) für y_w ist im Allgemeinen nicht erwartungstreu, selbst dann, wenn alle \bar{x}_k erwartungstreu geschätzt werden. Allerdings kann man davon ausgehen, dass er *konsistent* ist, also für immer längere Messreihen, $N_k \rightarrow \infty$, mit großer Wahrscheinlichkeit immer dichter bei y_w liegt. Allerdings, je besser die Linearitätsbedingung (D.16) erfüllt ist, umso besser approximiert (D.13) einen erwartungstreuen Schätzer.

x_k auf die Berechnung von y . Folglich müssen gerade jene Größen x_k mit besonderer Sorgfalt gemessen werden, bei denen $|a_k|$ vergleichsweise groß ist.¹⁰⁾ Die Fortpflanzung der Größtfehler $\Delta x_{\max,k}$ ist durch

$$\Delta x_{\max,y} = \sum_{k=1}^K |a_k| \Delta x_{\max,k} \quad (\text{D.15})$$

beschrieben.

Die für die Gültigkeit der Fehlerfortpflanzungsgesetze (D.14) und (D.15) vorausgesetzte Linearität ist gewährleistet, wenn

$$\left| \frac{y(\bar{x}_1, \dots, \bar{x}_K) + \sum_{k=1}^K a_k \Delta x_k}{y(\bar{x}_1 + \Delta x_1, \dots, \bar{x}_K + \Delta x_K)} - 1 \right| \ll 1 \quad (\text{D.16})$$

im Bereich

$$-\Delta x_k < \Delta x_k < \Delta x_k, \quad (\text{D.17})$$

gilt. Die Größe Δx_k ist in Gl. (D.12) definiert. Der Gesamtfehler ergibt sich dann aus der Fortpflanzung des zufälligen und des Größtfehlers,

$$y = \bar{y} \pm (\Delta x_{\max,y} + t_{\alpha/2; N_1, \dots, N_K}^* s_{\bar{y}}) \quad ((1 - \alpha) \times 100\%) .$$

Die Faktoren $t_{\alpha/2; N_1, \dots, N_K}^*$ sind für große Messreihen, also für $N_k \gtrsim 30$, durch die letzte Spalte in der Tabelle D.1 auf S. 145 gegeben, denn dann ist \bar{y} näherungsweise normalverteilt. Für endliche Messreihen geben wir diese Faktoren nicht explizit an. Allerdings, für das Signifikanzniveau $1 - \alpha = 0,95$ und $N_k = 10$, für alle $k = 1, \dots, K$, kann $t_{\alpha/2; N_1, \dots, N_K}^* \approx 3$ gesetzt werden.

Für die beschriebene statistische Fehleranalyse und die Fehlerfortpflanzung mussten wir viele Voraussetzungen machen, die in der Praxis niemals vollständig erfüllt sein können. Ist beispielsweise eine physikalische Größe wie die Länge immer positiv, so kann für endlich viele Messwerte niemals eine Gauss- oder auch t -Verteilung streng gelten. Deshalb sollten wir die Angaben zum Vertrauensbereich immer nur als Näherungen ansehen und eine sinnvolle Rundung der Zahlenwerte vornehmen. *Eine noch so „hochgestochene“ statistische Auswertung kann nicht den gesunden Menschenverstand ersetzen.*

¹⁰⁾ Als Erläuterung zum Fehlerfortpflanzungsgesetz (D.14) merken wir Folgendes an: Die zufälligen Schwankungen von \bar{y} um den gesuchten wahren Wert $y_w = y(x_{w,1}, \dots, x_{w,K})$, resultieren aus der additive Überlagerung der Schwankungen der einzelnen Mittelwerte \bar{x}_k . (Tritt auch ein systematischer Fehler auf, so sind im Folgenden $x_{w,k}$ und y_w durch $x_{w,k} + \Delta x_{s,x_k}$ bzw. $y_w + \Delta x_{s,y}$ zu ersetzen.) Je stärker sich y an der Stelle

$$(x_{w,1}, \dots, x_{w,K}) \approx (\bar{x}_1, \dots, \bar{x}_K)$$

mit \bar{x}_k ändert, also je größer $|a_k|$ ist, umso stärker ist der Einfluss der Änderungen von \bar{x}_k auf \bar{y} . Schwankt nun \bar{x}_k normalverteilt mit der Dichte $g(x_{w,k}, s_{\bar{x}_k}; x_k)$, so ist der zufällige Beitrag von \bar{x}_k zur Schwankung von \bar{y} um y_w gerade durch die Verteilungsdichte $g(y_w, |a_k| s_{\bar{x}_k}; x_k)$ beschrieben. Werden diese K zufälligen und untereinander vollständig unabhängigen Beiträge additiv überlagert, so ist \bar{y} näherungsweise normalverteilt mit der Dichte $g(y_w, s_{\bar{y}}; y)$. Diese Näherung gilt umso besser, je mehr Messwerte N_k in den Messreihen zu den x_k vorliegen. Zum einen liegt das daran, dass die Schwankungen von $(\bar{x}_k - x_{w,k})/s_{\bar{x}_k}$ für $N_k \rightarrow \infty$ immer genauer durch eine Normalverteilung beschrieben werden, zum anderen werden dann die Standardabweichungen $s_{\bar{x}_k} = s_{x_k}/\sqrt{N_k}$ und damit die Bereiche (D.17) immer kleiner, so dass hierin die Änderung von y immer besser als linear angesehen werden kann.

Tabelle D.2: Messreihen für den Stabdurchmesser d und die Stablänge l

n	d/mm	l/mm
1	9,99	99,74
2	9,97	100,06
3	9,97	100,26
4	9,95	99,92
5	10,00	100,08
6	9,96	100,18
7	10,01	99,54
8	10,00	99,96
9	9,98	100,01
10	9,98	99,80
11	9,12	97,60

D.5 Beispiel

Wir wollen das Volumen V eines Stabes bestimmen. Gemessen werden sein Durchmesser d sowie die Länge l . Daraus ist

$$V(d, l) = \frac{\pi}{4} d^2 l \quad (\text{D.18})$$

zu berechnen. Wir gehen davon aus, dass d mit einer Schraublehre der Präzision 0,01mm gemessen wird und l mit einer Schieblehre der Präzision 0,05mm. Bei der Schraublehre sei der Nullpunkt in der Weise verschoben, dass alle Durchmesser um 0,03mm zu klein gemessen werden. Die erhaltenen Messreihen geben wir in der Tabelle D.2 an. Zur Auswertung gehen wir in folgenden Schritten vor:

Aussonderung von Ausreißern:

Der jeweils elfte Wert der beiden Messreihen ist offenbar ein Ausreißer und wird deshalb bei der folgenden Auswertung weggelassen.

Nullpunktkorrektur:

Von den Messwerten $d(n)$ wird der uns bekannte systematische Nullpunktfehler

$$\Delta x_{s,d} = -0,03\text{mm}$$

abgezogen. Das heisst konkret, alle Werte $d(n)$ in der Tabelle D.2 werden um 0,03mm vergrößert.

Mittelwerte und Standardabweichungen:

Die Mittelwerte und Standardabweichungen werden nach (D.3), (D.4) und (D.5) berechnet. Wir erhalten dann für den Durchmesser

$$\bar{d} = 10,011 \text{ mm} \quad , \quad s_d \approx 0,019 \text{ mm} \quad , \quad s_{\bar{d}} \approx 0,006 \text{ mm}$$

und für die Länge

$$\bar{l} = 99,945 \text{ mm} \quad , \quad s_l \approx 0,244 \text{ mm} \quad , \quad s_{\bar{l}} \approx 0,077 \text{ mm} \quad .$$

Hierbei haben wir auf drei Stellen hinter dem Komma gerundet. Das ist sinnvoll, weil die Messwerte bis auf zwei Stellen hinter dem Komma vorliegen und die Ungenauigkeit des Mittelwertes etwa um $1/\sqrt{N} = 1/\sqrt{10} \approx 0,3$ kleiner als die einer Einzelmessung ist.

Vertrauensbereiche zu den zufälligen Fehlern:

Zur Irrtumswahrscheinlichkeit $\alpha = 0,05$ wollen wir nun den Vertrauensbereich angeben. Nach Gl. (D.11) brauchen wir dazu den Korrekturfaktor $t_{0,025;9}$, den wir aus der Tabelle D.1 auf S. 145 zu $t_{0,025;9} = 2,26$ ablesen. Damit erhalten wir

$$\begin{aligned} d &= (10,011 \pm 2,26 \times 0,006) \text{ mm} \quad (95\%) \\ &= (10,011 \pm 0,014) \text{ mm} \quad (95\%) \quad . \end{aligned}$$

und analog

$$l = (99,945 \pm 0,174) \text{ mm} \quad (95\%) \quad .$$

Das bedeutet, mit der Wahrscheinlichkeit von 95% liegt $\bar{d} = 10,011 \text{ mm}$ dichter als 0,006 mm am wahren Durchmesser und $\bar{l} = 99,945 \text{ mm}$ dichter als 0,174 mm an der wahren Länge.

Gesamtfehler:

Die vom Hersteller angegebene Präzision der Messinstrumente wird hier allein als Größtfehler angesehen,

$$\Delta x_{\max,d} = 0,01 \text{ mm} \quad , \quad \Delta x_{\max,l} = 0,05 \text{ mm} \quad .$$

Zusammen mit den zufälligen Fehlern ergeben sich dann nach Rundung die folgenden Gesamtfehler:

$$\begin{aligned} d &= (10,010 \pm 0,025) \text{ mm} \quad (95\%) \\ l &= (99,94 \pm 0,23) \text{ mm} \quad (95\%) \quad . \end{aligned}$$

Wenn wir davon ausgehen, dass d mit der recht präzisen Schraublehre gemessen wurde, so könnten wir geneigt sein, die Angaben zum Durchmesser bis auf $1 \mu\text{m}$ (3 Stellen hinter dem Komma) genau zu machen. Die Andruckkraft der Messspindel unterliegt aber einer gewissen Willkür, die wir mit $\pm 5 \mu\text{m}$ abschätzen. Deshalb haben wir die Angaben zum Durchmesser auf ein ganzzahliges Vielfaches von $5 \mu\text{m}$ gerundet. Generell sollte bei solchen Rundungen der Vertrauensbereich eher vergrößert als verkleinert werden. Beispielsweise hätten wir den Gesamtfehler $\Delta x_d = 0,022 \text{ mm}$ auch auf $\Delta x_d = 0,025 \text{ mm}$ aufgerundet.

Prüfung auf Linearität für Fehlerfortpflanzung:

Wir prüfen hier, ob die Linearitätsbedingung (D.16) im Bereich (D.17) erfüllt ist und damit die Fehlerfortpflanzungsgesetze (D.14) und (D.15) anwendbar sind. Dazu berechnen wir zunächst die Ableitungen

$$a_d = \left. \frac{\partial V(d,l)}{\partial d} \right|_{\bar{d},\bar{l}} = \frac{\pi}{4} 2\bar{d}\bar{l} \approx 1572 \text{ mm}^2$$

und

$$a_l = \left. \frac{\partial V(d, l)}{\partial l} \right|_{\bar{d}, \bar{l}} = \frac{\pi}{4} \bar{d}^2 \approx 79 \text{ mm}^2 .$$

Der Vergleich beider Anstiege liefert $a_d \approx 20 a_l$. Dies zeigt, dass wir mit größter Sorgfalt gerade d messen müssen, um den Gesamtfehler für das Volumen gering zu halten. Setzen wir nun die Anstiege a_d und a_l in (D.16) ein, so erhalten wir für

$$\begin{aligned} \Delta x_d &= -(\Delta x_{\max, d} + t_{0,025;9} s_{\bar{d}}) \dots (\Delta x_{\max, d} + t_{0,025;9} s_{\bar{d}}) \\ &\approx (-0,024 \dots 0,024) \text{ mm} \end{aligned}$$

und

$$\begin{aligned} \Delta x_l &= -(\Delta x_{\max, l} + t_{0,025;9} s_{\bar{l}}) \dots (\Delta x_{\max, l} + t_{0,025;9} s_{\bar{l}}) \\ &\approx (-0,225 \dots 0,225) \text{ mm} \end{aligned}$$

die folgenden Relationen:

$$\left| \frac{V(\bar{d}, \bar{l}) + (a_d \Delta x_d + a_l \Delta x_l)}{V(\bar{d} + \Delta x_d, \bar{l} + \Delta x_l)} - 1 \right| < 0,00002 \ll 1 .$$

Die Linearitätsbedingung (D.16) ist also hervorragend erfüllt. Damit können wir die Fehlerfortpflanzungsgesetze anwenden.

Mittleres Stabvolumen:

Für das Stab-Volumen (D.18) erhalten wir den Mittelwert

$$\bar{V} = V(\bar{d}, \bar{l}) = \frac{\pi}{4} \bar{d}^2 \bar{l} \approx 7866,9 \text{ mm}^3 .$$

Fortpflanzung des zufälligen Fehlers:

Die Standardabweichung von \bar{V} ergibt sich nach dem Fehlerfortpflanzungsgesetz (D.14) zu

$$\begin{aligned} s_{\bar{V}} &= \sqrt{a_d^2 s_{\bar{d}}^2 + a_l^2 s_{\bar{l}}^2} \\ &\approx 11,3 \text{ mm}^3 . \end{aligned}$$

Vertrauensbereich des Volumens:

Wegen der wenigen Messwerte, $N = 10$, können wir nicht davon ausgehen, dass \bar{V} normalverteilt ist. Allerdings können wir in guter Näherung davon ausgehen, dass \bar{V} mit einer Wahrscheinlichkeit $> 95\%$ dichter als $3s_{\bar{V}}$ an V_w liegt. Damit erhalten wir gerundet

$$\begin{aligned} V &= (7867 \pm 33) \text{ mm}^3 \quad (95\%) \\ &= 7867 \text{ mm}^3, \quad \pm 0,4\% \quad (95\%) . \end{aligned}$$

Größtfehler des Volumens:

Nach (D.15) berechnen wir den Größtfehler des Volumens,

$$\begin{aligned} \Delta x_{\max, V} &= |a_d| \Delta x_{\max, d} + |a_l| \Delta x_{\max, l} \\ &= (1572 \times 0,01 + 79 \times 0,05) \text{ mm}^3 \\ &\approx 20 \text{ mm}^3 . \end{aligned}$$

Gesamtfehler des Volumens:

Den zufälligen Fehler und den Größtfehler des Volumens fassen wir schließlich wie folgt zusammen:

$$\begin{aligned} V &= (7867 \pm 53) \text{ mm}^3 \quad (95\%) \\ &= 7867 \text{ mm}^3, \quad \pm 0,7\% \quad (95\%) . \end{aligned}$$

Wir sagen: *Das Stabvolumen misst 7867 mm³. Mit einer Wahrscheinlichkeit von 95% weicht es um nicht mehr als 53 mm³ vom wahren Wert ab.*¹¹⁾

D.6 Gewichtete Mittelung

Wird eine Größe x mit unterschiedlichen Methoden gemessen, so erhält man im Allgemeinen unterschiedliche Werte \hat{x}_n und Vertrauensbereiche $s_{x,n}$, $n = 1, 2, \dots, N$. Mit der gewichteten Mittelung kann man aus diesen Werten einen einzigen Schätzwert \hat{x} berechnen, so dass Messwerte mit einem kleineren Fehler ein größeres Gewicht bekommen gegenüber den ungenaueren Werten,

$$\hat{x} = \sum_{n=1}^N w_n \cdot \hat{x}_n \quad (\text{D.19})$$

mit den Gewichten

$$w_n = \frac{1}{\sum_{m=1}^N \frac{1}{s_{x,m}^2}} \cdot \frac{1}{s_{x,n}^2} .$$

Der Vertrauensbereich von \hat{x} ist

$$s_x = \frac{1}{N} \sqrt{\sum_{m=1}^N s_{x,m}^2} \quad (\text{D.20})$$

Sind alle Ungenauigkeiten $s_{x,n}$ gleich groß, folgen $w_n = 1/N$ und $s_x = s_{x,n}/\sqrt{N}$, was der gewöhnlichen Mittelung entspricht.

¹¹⁾Die Werte der Messreihen in der Tabelle D.2 wurden auf einem Rechner erzeugt, mit normalverteilten zufälligen Fehlern für den Durchmesser $d_w = 10 \text{ mm}$ und die Länge $l_w = 100 \text{ mm}$. Die „wahren“ Standardabweichungen sind $\sigma_d = 0,02 \text{ mm} \approx s_d$ und $\sigma_l = 0,2 \text{ mm} \approx s_l$. Das „wahre“ Volumen ist somit auch exakt bekannt, $V_w = (\pi/4)10^4 \text{ mm}^3 \approx 7854 \text{ mm}^3$, was unsere Fehlerabschätzung recht plausibel scheinen lässt.

Anhang E

Regression

Gegeben seien zwei Messreihen, $x(n)$ und $y(n)$, $n = 1, 2, \dots, N$, die mit Messfehlern behaftet sind. Aus theoretischen Überlegungen sei bekannt, dass zwischen den fehlerfreien Größen y und x ein funktionaler Zusammenhang einer bestimmten Form (Funktionen-Klasse) besteht, $y = f_{a,b,\dots}(x)$, mit frei wählbaren Parameter a, b, \dots . Die *Regression* befasst sich mit der Bestimmung (Schätzung) dieser Parameter aus den gegebenen Messwerten. Dabei wird häufig die Methode der kleinsten quadratischen Abweichung angewandt, die für normalverteilte Messfehler äquivalent zur Maximum-Likelihood-Methode ist, die für große N beste (erwartungstreue, effiziente) Schätzer liefert.

E.1 Lineare Regression

Zwischen den fehlerfreien Größen x und y möge ein linearer Zusammenhang bestehen, es gelte also die Geradengleichung

$$y = a + b \cdot x \ .$$

Der Anstieg b und der Schnittpunkt a der Geraden mit der Ordinate (y -Achse) sind aus den Messwerten x_n und y_n , $n = 1, 2, \dots, N$, zu schätzen. Je nachdem, ob nur y oder auch x fehlerbehaftet sind, und ob jeder Messwert mit einem gleichem absoluten Fehler versehen ist, werden verschiedene Fälle betrachtet.

E.1.1 1. Fall: x fehlerfrei, y mit festem Fehler σ_y

Die Messwerte x_n , $n = 1, 2, \dots, N$, seien fehlerfrei, und die Messfehler y_n seien bei beliebigem n normalverteilt mit der Standardabweichung σ_y . Dann ist die Log-Likelihood-Funktion (s. S. 89 ff.) gegeben durch

$$\begin{aligned} \ln L(y_1, \dots, y_N) &= \ln \prod_{n=1}^N \frac{1}{\sqrt{2\pi}\sigma_y} \exp\left(-\frac{[y_n - (a + bx_n)]^2}{2\sigma_y^2}\right) \\ &= \sum_{n=1}^N \left[\ln\left(\frac{1}{\sqrt{2\pi}\sigma_y}\right) + \left(-\frac{[y_n - (a + bx_n)]^2}{2\sigma_y^2}\right) \right] \\ &= \sum_{n=1}^N \ln\left(\frac{1}{\sqrt{2\pi}\sigma_y}\right) - \frac{\chi^2}{2} \ . \end{aligned}$$

darin haben wir

$$\chi^2(a, b) \equiv \frac{1}{\sigma_y^2} \sum_{n=1}^N (y_n - b \cdot x_n - a)^2 \quad (\text{E.1})$$

gesetzt. Die Likelihood-Funktion nimmt ihren Maximalwert beim Minimum von $\chi^2(a, b)$ an. Dies ist die sogenannte *Methode der kleinsten Fehlerquadrate*,

$$\chi^2(a, b) = \min!$$

Aus

$$\frac{\partial \chi^2(a, b)}{\partial a} = 0 \quad \text{und} \quad \frac{\partial \chi^2(a, b)}{\partial b} = 0$$

erhalten wir folgende Schätzwerte für die Parameter a und b unserer Regressionsgeraden:

$$\boxed{\hat{b} = \frac{\overline{xy} - \bar{x}\bar{y}}{S_x^2}} \quad (\text{E.2})$$

mit

$$\bar{x} = \frac{1}{N} \sum_{n=1}^N x_n, \quad \bar{y} = \frac{1}{N} \sum_{n=1}^N y_n, \quad \overline{xy} = \frac{1}{N} \sum_{n=1}^N x_n y_n, \quad \overline{x^2} = \frac{1}{N} \sum_{n=1}^N x_n^2, \quad (\text{E.3})$$

und

$$S_x^2 = \overline{x^2} - \bar{x}^2,$$

sowie

$$\boxed{\hat{a} = \bar{y} - \hat{b} \cdot \bar{x}} \quad (\text{E.4})$$

Die Standardabweichung von \hat{b} ist

$$s_b = \sqrt{\frac{1}{N} \frac{\sigma_y^2}{S_x^2}}. \quad (\text{E.5})$$

Wenn die Gerade gut den Zusammenhang zwischen x und y widerspiegelt, dann gilt $\chi^2/(N-2) \approx 1$, wobei χ^2 in (E.1) an der Stelle $a = \hat{a}$ und $b = \hat{b}$ zu nehmen ist. Folglich gilt dann

$$\begin{aligned} \sigma_y^2 \approx s_y^2 &= \frac{1}{N-2} \sum_{n=1}^N (y_n - \hat{b} \cdot x_n - \hat{a})^2 \\ &= \frac{N}{N-2} \left(\overline{y^2} - \bar{y}^2 - \hat{b}^2 (\overline{x^2} - \bar{x}^2) \right) \\ &= \frac{N}{N-2} S_y^2 (1 - \rho_{xy}^2), \end{aligned}$$

mit dem Korrelationskoeffizienten

$$\rho_{xy} = \frac{\overline{xy} - \bar{x}\bar{y}}{S_x S_y}, \quad S_y^2 = \overline{y^2} - \bar{y}^2, \quad \overline{y^2} = \frac{1}{N} \sum_{n=1}^N y_n^2.$$

Setzt man dies in (E.5) ein, folgt schließlich,

$$\boxed{s_b \approx \sqrt{\frac{1}{N-2} \cdot \frac{S_y^2}{S_x^2} \cdot (1 - \rho_{xy}^2)}} \quad (\text{E.6})$$

Auf der rechten Seite stehen nur noch Größen, die aus den Messwertpaaren (x_n, y_n) berechnet werden können.

Für die Standardabweichung von \hat{a} erhält man

$$s_a = s_b \cdot \sqrt{x^2} \quad (\text{E.7})$$

Der nach (E.2) geschätzte Anstieg \hat{b} liegt mit ca. 68% iger Wahrscheinlichkeit dichter als s_b am (unbekannten) wahren Anstieg b ,

$$b = \hat{b} \pm s_b \quad (68\%) .$$

Eine analoge Aussage gilt für den Schnittpunkt a ,

$$a = \hat{a} \pm s_a \quad (68\%) .$$

Ersetzt man in den Gleichungen σ_y durch $2 \cdot \sigma_y$, erhöht sich die Signifikanz auf etwa 95%. Es sei jedoch erinnert, dass hierbei nur zufällige Fehler der Messwerte y_n erfasst werden.

E.1.2 2. Fall: x_n und y_n mit Fehler σ_x bzw. σ_y

Die Messwerte x_n und y_n seien mit dem Fehler (Standardabweichung) σ_x bzw. σ_y gemessen, für jede Messung n also gleich. Man kann dies näherungsweise auf den ersten Fall zurückführen, indem man x_n wiederum als fehlerfrei ansieht, dafür aber die Varianz von y_n um $\hat{b}^2 \sigma_x^2$ erhöht. Die Parameter a und b sowie deren Unsicherheiten werden also wieder nach (E.2), (E.4), (E.6) und (E.7) geschätzt, wobei in (E.5)

$$\sigma_y^2 \longrightarrow \sigma_{y^*}^2 = \sigma_y^2 + \hat{b}^2 \sigma_x^2$$

zu ersetzen ist. Insbesondere gilt hier für den Fall, dass die Gerade die Daten gut approximiert,

$$\sigma_y^2 + \hat{b}^2 \sigma_x^2 \approx s_{y^*}^2 = \frac{1}{N-2} \sum_{n=1}^N (y_n - \hat{b} \cdot x_n - \hat{a})^2 .$$

Folglich kann die Ungenauigkeit des Anstiegs genau wie im ersten Fall nach (E.6) bestimmt werden.

E.1.3 3. Fall: x_n fehlerfrei, y_n mit variablem Fehler $\sigma_{y;n}$

Die Messwerte x_n seien fehlerfrei, und die Messfehler y_n , $n = 1, 2, \dots, N$, seien normalverteilt mit der Standardabweichung $\sigma_{y;n}$.

Die Methode der kleinsten gewichteten Fehlerquadrate,

$$\chi^2 = \sum_{n=1}^N \frac{(y_n - b \cdot x_n - a)^2}{\sigma_{y;n}^2} = \min! \quad (\text{E.8})$$

liefert dann dieselben Gleichungen wie (E.2), (E.4), (E.6) und (E.7), nur mit dem Unterschied, dass anstelle der gewöhnlichen (E.3) nun gewichtete Mittelwerte treten,

$$\bar{x} = \sum_{n=1}^N w_n x_n , \quad \bar{y} = \sum_{n=1}^N w_n y_n , \quad \overline{xy} = \sum_{n=1}^N w_n x_n y_n , \quad \overline{x^2} = \sum_{n=1}^N w_n x_n^2 , \quad (\text{E.9})$$

mit den Gewichten

$$w_n = \frac{\frac{1}{\sigma_{y;n}^2}}{\sum_{m=1\dots N} \frac{1}{\sigma_{y;m}^2}} = \left(\sum_{m=1\dots N} \frac{\sigma_{y;n}^2}{\sigma_{y;m}^2} \right)^{-1} .$$

Darüber hinaus ist in (E.6) σ_y^2 durch

$$\overline{\sigma_y^2} = \frac{N}{\sum_{m=1\dots N} \frac{1}{\sigma_{y;m}^2}}$$

zu ersetzen.

Sind alle Fehler gleich $\sigma_{y;n}^2 = \sigma_y^2$, folgt $w_n = 1/N$, was wieder der gewöhnlichen Mittelung (E.3) entspricht, sowie $\overline{\sigma_y^2} = \sigma_y^2$.

Ist die Ungenauigkeit $\sigma_{y;n}$ ein gewisser Bruch α des Messwertes y_n , z.B. $\alpha = 0,005$ bei einem relativen Fehler von 0,5%, also $\sigma_{y;n} = \alpha \cdot y_n$, dann kann für die Gewichte wie folgt geschrieben werden,

$$w_n = \left(\sum_{m=1\dots N} \frac{y_n^2}{y_m^2} \right)^{-1} .$$

E.1.4 4. Fall: x_n und y_n mit variablem Fehler $\sigma_{x;n}$ bzw. $\sigma_{y;n}$

Die Messwerte x_n und y_n seien mit dem Fehler (Standardabweichung) $\sigma_{x;n}$ bzw. $\sigma_{y;n}$ gemessen. Man kann dies näherungsweise auf den obigen Fall zurückführen, indem man x_n wiederum als fehlerfrei ansieht, dafür aber die Varianz von y_n um $\widehat{b}^2 \sigma_{x;n}^2$ erhöht. Die Parameter a und b sowie deren Unsicherheiten werden also wieder nach (E.2), (E.4), (E.6) und (E.7) geschätzt, wobei gewichtete Mittelwerte (E.9) zu verwenden sind, nun aber mit den Gewichten

$$w_n = \frac{\frac{1}{\sigma_{y;n}^2 + \widehat{b}^2 \sigma_{x;n}^2}}{\sum_{m=1\dots N} \frac{1}{\sigma_{y;m}^2 + \widehat{b}^2 \sigma_{x;m}^2}} .$$

Damit hängt aber die rechte Seite von (E.2) auch von \widehat{b} ab, und eine Auflösung nach \widehat{b} ist im Allgemeinen schwierig.

Index

- AR(D), 66
- MA(D), 65

- Abbildung, logistische, 64
- AKF, 50
- aposteri-Wahrscheinlichkeit, 125
- apriori-Wahrscheinlichkeit, 125
- Ausreiser, 138
- Autokorrelationsfunktion, 50, 60
- Autokorrelationsfunktion, eines harm. Signals, 55

- Bayes Statistik ff., 125
- Besselsche Korrektur, 89
- bias, 84
- Binomialverteilung, 83
- Borelmenge, 58

- chapter-programm-Fast-Fourier-Transformation ff., 131
- charakteristische Gleichung von AR(D), 69

- delay plot, 45
- Dichte-Schätzer, Histogramm, 78
- Dichte-Schätzer, Kern-, 78
- Dichte-Schätzer, naiver, 78
- Dispersion, 142

- Energiespektraldichte, 60
- Entfaltung, 14

- Fehler, grober, 138
- Fehler, systematischer, 84, 138
- Fehler, zufälliger, 138
- Fehlerfortpflanzung, Größtfehlers ff., 139
- Fehlerfortpflanzung, zufälliger Fehler ff., 145
- Fehlerrechnung, 137
- Formantfrequenz, 26
- Fourier-Reihe, Dreieck, 128
- Fourier-Reihe, Rechteck, 127
- Fourier-Reihenentwicklungen ff., 127

- Glockenkurve, 142
- Grenzwertsatz, zentraler, 87
- Größtfehler, 138

- Histogramm, 78

- Interpolation Code ff., 133
- Interpolation ff., 29
- Irrfahrt, 67
- Irrtumswahrscheinlichkeit, 143

- Kern-Funktion, 78
- KKF, 50
- Korrelation, ff., 45
- Korrelationsfunktion, 50
- Korrelationskoeffizient, 48
- Kovarianz, 48
- Kreuzenergie, 60
- Kreuzenergiespektraldichte, 60
- Kreuzkorrelationsfunktion, 50, 60
- Kreuzkorrelationsfunktion, zeitkontinuierliche, 52
- Kreuzleistungsspektrum, 53

- least mean square (LMS) fit, 45

- Maximum Entropie Schätzung, ff., 91
- Maximum Likelihood Schätzung, ff., 89
- Mittlung, gewichtete, 150
- Mittelwert, 141
- Mittelwert einer Funktion, 58
- Mittelwert, zeitkontinuierlicher, 53
- mittlere Kreuzleistung, 54
- moving average, 65

- Normalverteilung, 142

- Periodogramm, 26
- Periodogramm, zeitdiskretes, 94
- Periodogramm, zeitkontinuierliches, 95
- Pitchfrequenz, 26
- Power Spektrum, 60

- Prozess, δ -korrelierter, 61
- Prozess, autoregressiver-, 66
- Prozess, diskreter, 57
- Prozess, farbiger, 62
- Prozess, ff., 57
- Prozess, Gleitmittel-, 65
- Prozess, i.i.d., 61
- Prozess, kontinuierlich, 57
- Prozess, stationärer, 57
- Prozess, stochastischer, 57
- Prozess, vollkommen unabhängiger, 61
- Prozess, weiser, 62
- Prozess, zeitdiskreter, 57
- Prozess, zeitkontinuierlicher, 57
- Prozess, zufälliger, 57
- Prädiktor, Nächster-Nachbar-, 75

- Randverteilung, 58
- Regression ff., 151
- Regression, lineare ff., 151

- scatter plot, 45
- Schätzer, effizienter, 87
- Schätzer, erwartungstreuer, 84
- Schätzer, konsistenter, 86
- Schätzer, unverzerrter, 84
- Schätzer, Varianz, 88
- Schätzung ff., 81
- Schätzung, Mittelwert-, 81
- Signal, ergodisches, 59
- Signal, stationäres, 51
- Signale, linear unabhängige, 49
- Signale, negativ korrelierte, 49
- Signale, positiv korrelierte, 49
- Signale, unkorrelierte, 49
- Spektraldichte, Rechteckfenster, 129
- Standardabweichung, 47
- Standardabweichung, der Einzelmessung, 141
- Standardabweichung, des Mittelwertes, 141
- Statistik, 81
- Stichprobe, 81
- Stichprobe, mathematische, 141
- Stichprobenfunktion, 81
- Streuung, 47, 142
- Student-Verteilung, 144

- t-Verteilung, 144

- Varianz, 47

- Varianz, zeitkontinuierliche, 53
- Verschiebeoperator, 66
- Verteilung, 57
- Verteilungsdichte, 58
- Vertrauensgrenzen, 143
- Vorhersage, beste, 74
- Vorhersageproblem, allgemeines, 73

- Wahrscheinlichkeitsdichte, Schätzung ff., 77
- Wahrscheinlichkeitsverteilung, 57
- Wiener-Chintchin-Theorem, 60

- Yule-Walker-Gleichungen, 69

- Zeitfunktion, spektrales Rechteckfenster, 129
- Zeitreihe, stationäre, 51
- Zentraler Grenzwertsatz, 87
- Zufallsgröße, 57

Literaturverzeichnis

- [BoJe76] Box G. E. P., Jenkins G. M. (1976): *Time Series Analysis — Forecasting and Control* (Prentice–Hall, Inc. Englewood Cliffs, New Jersey)
- [Si86] Silverman B. W. (1986): *Density Estimation for Statistics and Data Analysis* (Chapman & Hall, London [u.a.])
- [AKA85] Akaike, Hirotugu. *Prediction and entropy*. In *A Celebration of Statistics: the ISI Centenary Volume*, Anthony C. Atkinson, Stephen E. Fienberg, eds. New York: Springer–Verlag, 1985.
- [BAR89] Barlow, R. J. , *Statistics, A Guide to the Use of Statistical Methods in the Physical Sciences*, (John Wiley Sons, Chichester u.a.,1989)
- [BDS86] Brock W.A., Dechert W.D., Scheinkmann J. A Test for Independence Based On the Correlation Dimension. (Department of Economics, University of Wisconsin, Madison, University of Houston, and University of Chicago, 1986)
- [BRO96] Brockwell, Peter J., Davis Richard A. *Introduction to Time Series and Forecasting*. New York, Springer–Verlag, 1996.
- [BOX76] Box, George E.P., Jenkins Gwilym M. *Time Series Analysis, forecasting and control*. London, Prentice–Hall, 1976.
- [CAO97] Cao L. Practical method for determining the minimum embedding dimension. *Physica D* 1997; 110:43–50
- [CAO99] Cao L., Soofi A.S. Forecasting daily dollar exchange rates. *International Journal of Forecasting* (preprint, 1999)
- [CAS89] Casdagli M. Nonlinear prediction of chaotic time series. *Physica D* 1989; 35:335–356
- [FAR87] Farmer J.D., Sidorovich J.J. Predicting chaotic time series. *Phys. Rev. Lett.* 1987; 59:845–848
- [FRA86] Fraser A.M., Swinney H.L. Using mutual information to find independent coordinates for strange attractors. *Phys. Rev. A* 1986; 33:1134–1140
- [FRA89] Fraser A. Information and entropy in strange attractors. *IEEE Trans. Inform. Theory* 1989; IT–35:245–262
- [GRA83] Grassberger P., Procaccia I. Estimation of the Kolmogorov entropy from a chaotic signal. *Phys. Rev. A* 1983; 28:2591–2593
- [Gn79] Gnedenko B. W. (1979): *Lehrbuch der Wahrscheinlichkeitsrechnung* (Akademie–Verlag, Berlin)
- [HUR89] Hurvich C.M., Tsai C.L. Regression and time series model selection in small samples. *Biometrika* 1989; 76:297–307
- [KaEtA196] Kandel, E. R.; Schwartz, J. H.; Jessell, T. M. (Hrsg.) (1996): *Neurowissenschaften: Eine Einführung* (Spektrum Akademischer Verlag, Heidelberg [u.a.])
- [Ku59] Solomon Kullback (1968). *Information Theory and Statistics* (Dover Publications, Mineola, New York) (Erstauflage 1959)
- [LED85] Ledrappier F., Young L.S. The metric entropy of diffeomorphisms. Part I: Characterization of measures satisfying Pesin’s entropy formula. Part II: Relations between entropy, exponents and dimension. *Ann. Math.* 1985; 122:509–539, 540–574

- [LEV85] Leven R.W., Pompe B., Wilke C., Koch B.P. Experiments on periodic and chaotic motions of a parametrically forced pendulum. *Physica D* 1985; 16:371–384
- [OSB04] Oppenheimer A. V., Schafer R. W., und Buck J. R. (2004): *Time Series Analysis — Forecasting and Control*, 2., überarbeitete Auflage (Pearson, München u.a.)
- [PAC80] Packard N.H., Crutchfield J.P., Farmer J.D., Shaw R.S. Geometry from a time series. *Phys. Rev. Lett.* 1980; 45:712–716
- [POM93] Pompe B. Measuring statistical dependencies in a time series. *J. Stat. Phys.* 1993; 73:587–610
- [POM97] Pompe, Bernd. *Die Messung von Informationsflüssen mit einer verallgemeinerten Transinformation*. Greifswald, Ernst–Moritz–Arndt–Universität, Habilitationsschrift, 1997.
- [POM98] Pompe, Bernd. "Ranking and entropy estimation in nonlinear time series analysis." In *Nonlinear Analysis of Physiological Data* Holger Kantz, Jürgen Kurths, Gottfried Mayer–Kress, eds. Berlin: Springer–Verlag, 1998.
- [REN70] Rényi, Alfréd. *Probability Theory*. Amsterdam, North–Holland, 1970
- [SAU91] Sauer T., Yorke J.A., Casdagli M. Embedology, *J. Stat. Phys.* 1991; 65:579–616
- [SHA48] Shannon C.E. A mathematical theory of communication. *The Bell System Technical Journal* 1948; XXVII:379–423, 623–656
- [SIL85] Silverman, Bernard W. *Density Estimation for Statistics and Data Analysis*. London: Monographs on Statistics and Applied Probability 26, Chapman & Hall, 1985.
- [TAK80] Takens, Floris. "Detecting strange attractors in turbulence." In *Dynamical Systems and Turbulence*, Lecture Notes Mathematics, volume 898, p. 366. Berlin: Springer–Verlag, 1980,
- [TAY88] Taylor, J. R., *Fehleranalyse, Eine Einführung in die Untersuchung von Unsicherheiten in physikalische Messungen* (VCH Verlagsgesellschaft, Weinheim u.a., 1988)
- [VER88] Veres S. Asymptotic distributions of likelihood ratios for overparametrized ARMA processes. *J. Time Ser. Anal.* 1988; 8:345–357
- [Zel94] Zell, A. (1994): *Simulation neuronaler Netze* (Addison–Wesley, Bonn u.a.)