

# 5 Grundlagen der Stochastik

Ein stochastisches Filter hat die Aufgabe, anhand von Messwerten den Zustand eines Systems zu schätzen. Diese Messwerte sind – ebenso wie das zur Verfügung stehende mathematische Modell des Systems – mit Unsicherheiten bzw. Fehlern behaftet, zu deren Beschreibung sich stochastische Konzepte anbieten. Ohne detailliert auf mathematische Grundlagen oder Definitionen einzugehen, werden im Folgenden einige Begriffe aus der Stochastik eingeführt.

## 5.1 Die Zufallsvariable

Eine Zufallsvariable ist mathematisch gesehen eine Funktion. Diese Funktion ordnet den Ergebnissen eines Zufallsexperiments Werte zu, die Realisationen genannt werden. Manche Autoren unterscheiden nicht immer korrekt zwischen Zufallsvariable und Realisation, sondern betrachten eine Zufallsvariable wie eine Variable, die zufällig Werte annimmt.

Man unterscheidet zwischen diskreten und kontinuierlichen Zufallsvariablen: Diskrete Zufallsvariablen ordnen Werte aus einer abzählbaren Menge zu, während kontinuierliche Zufallsvariablen auf beliebige Werte abbilden. Diskrete Zufallsvariablen werden hier nicht weiter betrachtet.

Es gibt verschiedene Möglichkeiten, die Charakteristik einer Zufallsvariablen zu beschreiben. Dazu zählen die Verteilungsfunktion oder kumulative Wahrscheinlichkeit, die Wahrscheinlichkeitsdichtefunktion, die charakteristische Funktion oder die Momente der Zufallsvariablen.

### 5.1.1 Wahrscheinlichkeitsdichte

Die Verteilungsfunktion  $F_x(\xi)$  einer Zufallsvariable  $\mathbf{x}$  gibt an, mit welcher Wahrscheinlichkeit  $P$  Realisationen kleiner als eine bestimmte Schranke  $\xi$  sind:

$$F_x(\xi) = P(\mathbf{x} \leq \xi) \quad (5.1)$$

Hierbei wurde der Begriff Wahrscheinlichkeit verwendet, für den es unterschiedliche Definitionen gibt: Die Wahrscheinlichkeit eines Ereignisses ist nach der klassischen Definition von Laplace die Anzahl der zu einem Ereignis gehörenden Elementarereignisse dividiert durch die Anzahl aller möglichen Ereignisse. Diese Definition wurde von der Definition nach Richard von Mises abgelöst, die Wahrscheinlichkeit als die relative Häufigkeit des Auftretens eines Ereignisses beschreibt. Die am breitesten akzeptierte Definition stammt von Andrey Kolmogorov, die sich auf drei Axiome stützt:

1. Die Wahrscheinlichkeit eines Ereignisses ist größer oder gleich null.
2. Die Wahrscheinlichkeit des sicheren Ereignisses ist eins.
3. Für zwei sich gegenseitig ausschließende Ereignisse ist die Wahrscheinlichkeit, dass eines von beiden eintritt, gleich der Summe der Einzelwahrscheinlichkeiten.

Die Wahrscheinlichkeit, dass eine Realisation in einem bestimmten Intervall auftritt, berechnet sich nach Gl. (5.1) zu

$$P(\xi_0 < \mathbf{x} \leq \xi_1) = F_x(\xi_1) - F_x(\xi_0) . \quad (5.2)$$

Weiterhin gilt

$$\lim_{\xi \rightarrow -\infty} F_x(\xi) = 0 \quad (5.3)$$

$$\lim_{\xi \rightarrow \infty} F_x(\xi) = 1 \quad (5.4)$$

$$\xi_0 < \xi_1 \Rightarrow F_x(\xi_0) < F_x(\xi_1) . \quad (5.5)$$

Die Verteilungsfunktion wird auch als kumulative Wahrscheinlichkeit bezeichnet, deren Ableitung ist die Wahrscheinlichkeitsdichte:

$$p_x(\xi) = \frac{dF_x(\xi)}{d\xi} \quad (5.6)$$

Die Wahrscheinlichkeitsdichte ist eine nicht-negative Funktion,

$$p_x(\xi) \geq 0 , \quad (5.7)$$

für die weiterhin gilt:

$$\int_{-\infty}^{\infty} p_x(\xi) d\xi = 1 \quad (5.8)$$

Die Wahrscheinlichkeit, dass eine Realisation in einem bestimmten Intervall auftritt, lässt sich anhand von

$$P(\xi_0 < \mathbf{x} \leq \xi_1) = \int_{\xi_0}^{\xi_1} p_x(u) du \quad (5.9)$$

berechnen, die Verteilungsfunktion kann durch Integration über die Wahrscheinlichkeitsdichte ermittelt werden:

$$F_x(\xi) = \int_{-\infty}^{\xi} p_x(u) du \quad (5.10)$$

Das erste Moment einer Zufallsvariablen  $\mathbf{x}$  ist der Erwartungswert oder Mittelwert

$$m_x = E[\mathbf{x}] = \int_{-\infty}^{\infty} \xi p_x(\xi) d\xi . \quad (5.11)$$

Die Streuung der Realisationen um den Mittelwert wird durch das zweite zentrale Moment, die Varianz  $\sigma_x^2$ , beschrieben:

$$\sigma_x^2 = E[(\mathbf{x} - m_x)^2] = \int_{-\infty}^{\infty} (\xi - m_x)^2 p_x(\xi) d\xi \quad (5.12)$$

Die Varianz ist eine nicht-negative Zahl, die Wurzel aus der Varianz wird als Standardabweichung bezeichnet.

Zwei Zufallsvariablen  $\mathbf{x}$  und  $\mathbf{y}$  können anhand der Verbundwahrscheinlichkeitsdichte  $p_{xy}$  beschrieben werden, man erhält

$$P(\xi_0 < \mathbf{x} \leq \xi_1 \wedge \eta_0 < \mathbf{y} \leq \eta_1) = \int_{\xi_0}^{\xi_1} \int_{\eta_0}^{\eta_1} p_{xy}(u, v) du dv . \quad (5.13)$$

Die Wahrscheinlichkeitsdichten der einzelnen Zufallsvariablen heißen marginale Wahrscheinlichkeitsdichten, diese können aus der Verbundwahrscheinlichkeitsdichte durch Integration über den gesamten Wertebereich der jeweils anderen Zufallsvariable berechnet werden:

$$p_x(\xi) = \int_{-\infty}^{\infty} p_{xy}(\xi, \eta) d\eta \quad (5.14)$$

$$p_y(\eta) = \int_{-\infty}^{\infty} p_{xy}(\xi, \eta) d\xi \quad (5.15)$$

Liegt eine bestimmte Realisation  $\eta$  der Zufallsvariable  $\mathbf{y}$  vor, stellt sich die Frage nach der Wahrscheinlichkeit von Realisationen  $\xi$  der Zufallsvariable  $\mathbf{x}$  unter dieser Randbedingung. Dazu wird die bedingte Wahrscheinlichkeitsdichtefunktion  $p_{x|y}$  benötigt, es gilt

$$p_{x|y}(\xi | \mathbf{y} = \eta) = \frac{p_{xy}(\xi, \eta)}{p_y(\eta)} . \quad (5.16)$$

Eine Herleitung dieses Zusammenhangs, der als Bayes'sche Regel bekannt ist, ist in [85] zu finden. Der entsprechende Zusammenhang für den Fall, dass eine bestimmte Realisation  $\xi$  vorliegt, ist gegeben durch

$$p_{y|x}(\eta | \mathbf{x} = \xi) = \frac{p_{xy}(\xi, \eta)}{p_x(\xi)} \quad (5.17)$$

Damit ergibt sich eine alternative Formulierung der Bayes'schen Regel:

$$p_{xy}(\xi, \eta) = p_{x|y}(\xi | \mathbf{y} = \eta) p_y(\eta) = p_{y|x}(\eta | \mathbf{x} = \xi) p_x(\xi) \quad (5.18)$$

Der bedingte Erwartungswert  $E[\mathbf{x}|\mathbf{y}]$  kann als Funktion der Zufallsvariable  $\mathbf{y}$  verstanden werden. Folglich handelt es sich bei  $E[\mathbf{x}|\mathbf{y}]$  ebenfalls um eine Zufallsvariable, während  $E[\mathbf{x}]$  ein deterministischer Wert ist. Weiterhin gilt

$$E[E[\mathbf{x}|\mathbf{y}]] = E[\mathbf{x}] \quad (5.19)$$

$$\text{var}(\mathbf{x}) = E[\text{var}(\mathbf{x}|\mathbf{y})] + \text{var}(E[\mathbf{x}|\mathbf{y}]) \quad (5.20)$$

wobei  $\text{var}()$  die Varianz der entsprechenden Zufallsvariable bezeichnet.

Die Eigenschaften zweier Zufallsvariablen wird durch die Begriffe Unabhängigkeit, Unkorreliertheit und Orthogonalität näher beschrieben. Zwei Zufallsvariablen  $\mathbf{x}$  und  $\mathbf{y}$  heißen

- statistisch unabhängig, wenn  $p_{xy}(\xi, \eta) = p_x(\xi)p_y(\eta)$  gilt .
- unkorreliert, wenn  $E[\mathbf{x}\mathbf{y}] = E[\mathbf{x}]E[\mathbf{y}]$  gilt .
- orthogonal, wenn  $E[\mathbf{x}\mathbf{y}] = 0$  gilt.

Aus Unabhängigkeit folgt Unkorreliertheit, der Umkehrschluss gilt im Allgemeinen aber nicht.

Schließlich kann aus mehreren skalaren Zufallsvariablen  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$  ein Zufallsvektor  $\vec{\mathbf{x}}$  gebildet werden. Der Erwartungswert dieses Zufallsvektors berechnet sich gemäß

$$\vec{m}_x = E[\vec{\mathbf{x}}] = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \vec{\xi} p_{\vec{\mathbf{x}}}(\vec{\xi}) d\xi_1 \dots d\xi_n ,$$

die Kovarianzmatrix ist gegeben durch

$$\mathbf{P}_{xx} = E[(\vec{\mathbf{x}} - \vec{m}_x)(\vec{\mathbf{x}} - \vec{m}_x)^T] .$$

Die Hauptdiagonale der Kovarianzmatrix besteht aus den Varianzen der Zufallsvariablen  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ . Sind diese unkorreliert, verschwinden die Nebendiagonalelemente. Kovarianzmatrizen sind für den pathologischen Fall, dass Komponenten des Zufallsvektors die Varianz Null besitzen positiv semi-definit, ansonsten positiv definit.

## 5.1.2 Gaußverteilung

Im Rahmen der stochastischen Filterung spielt eine bestimmte Verteilungsfunktion, die Gauß- oder Normalverteilung, eine zentrale Rolle. Ein Grund hierfür ist, dass sich eine Normalverteilung zur Beschreibung vieler Zufallsphänomene eignet. Eine beobachtbare

Größe wird häufig von vielen Störquellen beeinflusst, die als Zufallsvariablen beschrieben werden können. Vereinfacht dargestellt strebt mit wachsender Anzahl der Zufallsvariablen die Wahrscheinlichkeitsverteilung ihrer Summe gegen eine Normalverteilung. Die Zufallsvariablen dürfen dabei beliebige Verteilungsfunktionen besitzen, sie müssen jedoch unabhängig sein. Die exakte mathematische Formulierung dieses Zusammenhangs ist als zentraler Grenzwertsatz bekannt. Ein weiterer Grund für die Sonderstellung der Normalverteilung ist die gute mathematische Handhabbarkeit. Dadurch können in vielen Fällen noch analytische Lösungen gefunden werden, was bei Annahme anderer Verteilungsfunktionen nicht möglich wäre. Darüber hinaus wird die Normalverteilung durch die ersten beiden Momente bereits vollständig beschrieben. Bei realen Problemstellungen wird man häufig, auch wenn keine Normalverteilung zu Grunde liegt, kaum über mehr Informationen verfügen und daher gezwungener Maßen eine Normalverteilung annehmen.

Die Wahrscheinlichkeitsdichte ist im skalaren Fall gegeben durch

$$p_x(\xi) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\xi - m_x)^2}{2\sigma^2}}, \quad (5.21)$$

ein normalverteilter Zufallsvektor der Dimension  $n \times 1$  besitzt die Dichtefunktion

$$p_{\vec{x}}(\vec{\xi}) = \frac{1}{\sqrt{(2\pi)^n |\mathbf{P}_{xx}|}} e^{-\frac{1}{2}(\vec{\xi} - \vec{m}_x)^T \mathbf{P}_{xx}^{-1}(\vec{\xi} - \vec{m}_x)}. \quad (5.22)$$

Werden mit normalverteilten Zufallsvektoren  $\vec{x}$  und  $\vec{y}$  lineare Operationen durchgeführt, ist das Ergebnis  $\vec{z}$  ebenfalls wieder ein normalverteilter Zufallsvektor:

$$\vec{z} = \mathbf{A}\vec{x} + \mathbf{B}\vec{y} + \vec{c} \quad (5.23)$$

Bei  $\mathbf{A}$  und  $\mathbf{B}$  handelt es sich um beliebige Gewichtungsmatrizen,  $\vec{c}$  ist ein beliebiger konstanter Vektor. Der Erwartungswert und die Kovarianzmatrix des Vektors  $\vec{z}$  können direkt angegeben werden:

$$\vec{m}_z = \mathbf{A}\vec{m}_x + \mathbf{B}\vec{m}_y + \vec{c} \quad (5.24)$$

$$\mathbf{P}_{zz} = \mathbf{A}\mathbf{P}_{xx}\mathbf{A}^T + \mathbf{A}\mathbf{P}_{xy}\mathbf{B}^T + \mathbf{B}\mathbf{P}_{yx}\mathbf{A}^T + \mathbf{B}\mathbf{P}_{yy}\mathbf{B}^T \quad (5.25)$$

Sind  $\vec{x}$  und  $\vec{y}$  unkorreliert, so folgt  $\mathbf{P}_{xy} = \mathbf{0}$  und  $\mathbf{P}_{yx} = \mathbf{0}$ , die Berechnung der Kovarianzmatrix von  $\vec{z}$  vereinfacht sich entsprechend.

Im allgemeinen Fall, d.h. bei beliebigen Wahrscheinlichkeitsdichten, müssen zur Berechnung der Wahrscheinlichkeitsdichte einer Summe von Zufallsvariablen die Wahrscheinlichkeitsdichten der Summanden gefaltet werden,

$$\mathbf{z} = \mathbf{x} + \mathbf{y} \quad (5.26)$$

$$p_z(\zeta) = \int_{-\infty}^{\infty} p_x(l) p_y(\zeta - l) dl, \quad (5.27)$$

was wesentlich aufwändiger ist.

Ein möglicher Nutzen der Gleichungen (5.23)–(5.25) soll anhand eines Beispiels verdeutlicht werden: Mit einem Zufallszahlengenerator soll ein Rauschvektor mit einer definierten Kovarianzmatrix  $\mathbf{P}_{zz}$  erzeugt werden. Der Zufallszahlengenerator liefert normalverteilte, mittelwertfreie Zufallszahlen<sup>1</sup> mit der Varianz 1, ein aus diesen Zufallszahlen aufgebauter Vektor  $\vec{\xi}$  kann als Realisation eines Zufallsvektors  $\vec{\mathbf{x}}$  mit der Kovarianzmatrix  $\mathbf{P}_{xx} = \mathbf{I}$  verstanden werden. Nun wird durch Cholesky-Zerlegung von  $\mathbf{P}_{zz}$  eine Gewichtungsmatrix  $\mathbf{A}$  bestimmt:

$$\mathbf{P}_{zz} = \mathbf{A}\mathbf{A}^T \quad (5.28)$$

Durch Multiplikation von  $\vec{\xi}$  mit dieser Gewichtungsmatrix  $\mathbf{A}$  erhält man eine Realisation eines Zufallsvektors  $\vec{\mathbf{z}}$  mit der gewünschten Kovarianzmatrix  $\mathbf{P}_{zz}$ , denn es gilt:

$$\vec{\mathbf{z}} = \mathbf{A}\vec{\mathbf{x}} \quad (5.29)$$

$$\begin{aligned} E[(\vec{\mathbf{z}} - \vec{m}_z)(\vec{\mathbf{z}} - \vec{m}_z)^T] &= E[\vec{\mathbf{z}}\vec{\mathbf{z}}^T] = E[(\mathbf{A}\vec{\mathbf{x}})(\mathbf{A}\vec{\mathbf{x}})^T] \\ &= \mathbf{A}E[\vec{\mathbf{x}}\vec{\mathbf{x}}^T]\mathbf{A}^T = \mathbf{A}\mathbf{P}_{xx}\mathbf{A}^T = \mathbf{A}\mathbf{A}^T = \mathbf{P}_{zz} \end{aligned} \quad (5.30)$$

■

Liegt eine Realisation  $\vec{\eta}$  eines Zufallsvektors  $\vec{\mathbf{y}}$  vor, so sind der bedingte Erwartungswert und die bedingte Kovarianzmatrix eines Zufallsvektors  $\vec{\mathbf{x}}$  gegeben durch

$$\vec{m}_{x|\vec{\eta}} = \vec{m}_x + \mathbf{P}_{xy}\mathbf{P}_{yy}^{-1}(\vec{\eta} - \vec{m}_y) \quad (5.31)$$

$$\mathbf{P}_{xx|\vec{\eta}} = \mathbf{P}_{xx} - \mathbf{P}_{xy}\mathbf{P}_{yy}^{-1}\mathbf{P}_{yx}. \quad (5.32)$$

Die Gl. (5.31) und (5.32) können als Lösung des Schätzproblems bei normalverteilten Zufallsvektoren aufgefasst werden. In Abschnitt 6 wird unter anderem anhand dieser Gleichungen das Kalman-Filter hergeleitet.

<sup>1</sup>Liefert der Zufallszahlengenerator z.B. gleichverteilte Zufallszahlen, erhält man aufgrund des zentralen Grenzwertsatzes eine näherungsweise normalverteilte Zufallszahl durch Addition mehrerer gleichverteilter Zufallszahlen.

## 5.2 Stochastische Prozesse

Ein stochastischer Prozess kann als eine Zufallsvariable mit einem zusätzlichen Parameter, der Zeit, verstanden werden. Betrachtet man einen stochastischen Prozess zu einem festen Zeitpunkt, erhält man folglich eine Zufallsvariable.

Eine wichtige Kenngröße eines stochastischen Prozesses  $\mathbf{x}(t)$  ist die Autokorrelationsfunktion

$$r_{xx}(t_1, t_2) = E[\mathbf{x}(t_1)\mathbf{x}(t_2)] , \quad (5.33)$$

anhand derer Aussagen über die Selbstähnlichkeit des Prozesses bei zeitlichen Verschiebungen getroffen werden. Betrachtet man zwei verschiedene Prozesse  $\mathbf{x}(t)$  und  $\mathbf{y}(t)$ , so erhält man die Kreuzkorrelationsfunktion:

$$r_{xy}(t_1, t_2) = E[\mathbf{x}(t_1)\mathbf{y}(t_2)] \quad (5.34)$$

Ein stochastischer Prozess heißt stationär im strengen Sinne, wenn die Charakteristiken des Prozesses, z.B. die Wahrscheinlichkeitsdichte, keine Funktionen der Zeit sind. Von Stationarität im weiteren Sinne spricht man, wenn die ersten beiden Momente keine Funktionen der Zeit sind, dadurch hängt z.B. auch die Autokorrelation nur von der zeitlichen Verschiebung, nicht aber vom absoluten Zeitpunkt ab:

$$r_{xx}(t_1, t_2) = r_{xx}(t_1 - t_2) = r_{xx}(\tau) = E[\mathbf{x}(t)\mathbf{x}(t - \tau)] \quad (5.35)$$

Bei einem normalverteilten Zufallsprozess folgt aus Stationarität im weiteren Sinne auch Stationarität im engen Sinne, da die Normalverteilung durch die ersten beiden Momente vollständig beschrieben wird.

Im weiteren Sinne stationäre Prozesse können ergodisch sein. In diesem Fall stimmt der Scharmittelwert des Prozesses zu einem beliebigen Zeitpunkt mit dem Zeitmittelwert jeder Realisation  $\xi_i(t)$  überein:

$$E[\mathbf{x}(t)] = \int_{-\infty}^{\infty} p_x(\xi) d\xi = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \xi_i(t) dt \quad (5.36)$$

Ist ein stochastischer Prozess ergodisch, so kann z.B. die Kovarianz oder die Autokorrelation anhand des Zeitmittelwertes einer einzigen Realisation berechnet werden. Das ist von entscheidender Bedeutung, da in der Praxis häufig nur eine einzige Realisation vorliegt. Die Autokorrelationsfunktion eines ergodischen Prozesses ist beispielsweise anhand von

$$r_{xx}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \xi_i(t)\xi_i(t - \tau) dt \quad (5.37)$$

berechenbar. Die Kreuzkorrelationsfunktion kann bei Betrachtung zweier Prozesse in analoger Weise berechnet werden. Beim Übergang ins Zeitdiskrete wird das Integral

durch eine entsprechende Summe ersetzt. Da in der Praxis nur ein endliches Zeitintervall betrachtet werden kann, erhält man lediglich einen Schätzwert der Autokorrelationsfunktion.

Die spektrale Leistungsdichte  $S_{xx}(j\omega)$  eines stochastischen Prozesses ist die Fourier-transformierte der Autokorrelationsfunktion, es gilt

$$S_{xx}(j\omega) = \int_{-\infty}^{\infty} r_{xx}(\tau) e^{-j\omega\tau} d\tau . \quad (5.38)$$

Die Fourier-Rücktransformierte ist gegeben durch

$$r_{xx}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{xx}(j\omega) e^{j\omega\tau} d\omega , \quad (5.39)$$

dieser Zusammenhang ist als Wiener-Khintchine-Relation bekannt. Die Leistung eines Signals erhält man als Funktionswert der Autokorrelationsfunktion an der Stelle  $\tau = 0$ , in diesem Fall besteht der Integrand in Gl. (5.39) lediglich aus der spektralen Leistungsdichte.

Die Erweiterung der bisher eingeführten Begriffe auf vektorielle stochastische Prozesse stellt keine Schwierigkeit dar und wird daher nicht weiter betrachtet. Stattdessen soll ein spezieller stochastischer Prozess, das weiße Rauschen, angesprochen werden.

### 5.2.1 Weißes Rauschen

Weißes Rauschen ist durch eine konstante spektrale Leistungsdichte gekennzeichnet:

$$S_{xx}(j\omega) = R \quad (5.40)$$

Die Signalleistung erhält man durch Integration über die spektrale Leistungsdichte, folglich besitzt dieser stochastische Prozess unendliche Leistung. Die Existenz von weißem Rauschen ist im Zeitkontinuierlichen daher nicht möglich.

Die Autokorrelationsfunktion erhält man zu

$$r_{xx}(\tau) = R\delta(\tau) , \quad (5.41)$$

wobei  $\delta(\tau)$  den Dirac-Impuls bezeichnet. Für  $\tau = 0$  wird dieser unendlich, was in Übereinstimmung mit obigen Überlegungen auf eine unendliche Leistung schließen lässt. Da bei mittelwertfreien Prozessen die Autokorrelationsfunktion an der Stelle  $\tau = 0$  gleich der Varianz ist, ist die Varianz von weißem Rauschen im Zeitkontinuierlichen nicht definiert.

Dennoch ist weißes Rauschen als Konzept sinnvoll: Geht man davon aus, dass man von einem weißem Rauschprozess im Zeitkontinuierlichen durch Mittelwertbildung über den Abtastzeitraum  $T$  ins Zeitdiskrete übergeht, erhält man auch im Zeitdiskreten einen weißen Rauschprozess. Die Leistung dieses zeitdiskreten Rauschprozesses ist endlich,



seine Existenz ist problemlos möglich. Die Varianz  $R_k$  dieses Prozesses ist definiert und hängt, wie in Abschnitt 3.4.3 gezeigt, gemäß

$$R_k = \frac{R}{T} \quad (5.42)$$

mit der spektralen Leistungsdichte des nicht existierenden, weißen Rauschens im Zeitkontinuierlichen zusammen. Um z.B. bei der Spezifikation des Rauschens von Inertialsensoren nicht die Varianz bei einer bestimmten Abtastzeit angeben zu müssen, wird stattdessen die spektrale Leistungsdichte eines hypothetischen, zeitkontinuierlichen, weißen Rauschprozesses verwendet, die abtastzeitunabhängig ist.

## 5.2.2 Zeitkorreliertes Rauschen

Ein zeitkorrelierter Rauschprozess lässt sich mit FIR-<sup>2</sup> oder mit IIR-<sup>3</sup> Systemen beschreiben, an deren Eingang ein weißes Rauschen vorliegt. Im Rahmen der stochastischen Filterung wird üblicherweise die Beschreibung mit IIR-Systemen bevorzugt, da damit eine einfache Zustandsraumdarstellung resultiert.

Ein skalarer Gauß-Markov-Prozess 1. Ordnung ist im Zeitkontinuierlichen gegeben durch

$$\dot{w} = -\frac{1}{t_{corr}}w + \sigma\sqrt{\frac{2}{t_{corr}}}\eta. \quad (5.43)$$

Hierbei bezeichnet  $w$  den zeitkorrelierten Rauschterm,  $\eta$  ist ein weißer Rauschprozess mit der spektralen Leistungsdichte  $S_{\eta\eta} = 1$ ,  $t_{corr}$  ist die Korrelationszeit und  $\sigma$  ist die Standardabweichung des Gauß-Markov-Prozesses. Die Autokorrelationsfunktion dieses Prozesses,

$$r_{ww}(\tau) = \sigma^2 e^{-\frac{|\tau|}{t_{corr}}}, \quad (5.44)$$

ist an der Stelle  $\tau = t_{corr}$  folglich auf  $1/e$  ihres Maximalwertes, der bei  $\tau = 0$  vorliegt, abgefallen. Durch Fourier-Transformation erhält man die spektrale Leistungsdichte zu

$$S_{ww}(j\omega) = \underbrace{\frac{\sigma\sqrt{\frac{2}{t_{corr}}}}{\frac{1}{t_{corr}} + j\omega}}_{H(j\omega)} \cdot \underbrace{\frac{\sigma\sqrt{\frac{2}{t_{corr}}}}{\frac{1}{t_{corr}} - j\omega}}_{H(-j\omega)} S_{\eta\eta}, \quad (5.45)$$

dabei ist  $H(j\omega)$  die Übertragungsfunktion des IIR-Systems, offensichtlich handelt es sich hierbei um einen Tiefpass 1. Ordnung. Gauß-Markov-Prozesse 1. Ordnung eignen sich daher zur Beschreibung langsam veränderlicher Vorgänge, wie z.B. der Bias-Drift eines Inertialsensors.

Das zeitdiskrete Äquivalent dieses Rauschprozesses ist gegeben durch

$$w_{k+1} = e^{-\frac{T}{t_{corr}}} w_k + \eta_k, \quad (5.46)$$

---

<sup>2</sup>Finite impulse response

<sup>3</sup>Infinite impulse response

die Abtastzeit ist mit  $T$  bezeichnet. Für die Varianz des weißen Rauschprozesses erhält man

$$E[\eta^2] = \sigma^2(1 - e^{-\frac{2T}{\tau_{corr}}}) . \quad (5.47)$$

### Schätzung der Rauschprozessparameter

Die Schätzung der Parameter eines zeitkorrelierten Rauschprozesses soll im Folgenden beispielhaft anhand eines vektoriellen Gauß-Markov-Prozesses 1. Ordnung demonstriert werden. Ein solcher Prozess ist gegeben durch

$$\vec{w}_{k+1} = \Phi \vec{w}_k + \vec{\eta}_k . \quad (5.48)$$

Durch Multiplikation mit  $\vec{w}_k^T$  erhält man

$$\vec{w}_{k+1} \vec{w}_k^T = \Phi \vec{w}_k \vec{w}_k^T + \vec{\eta}_k \vec{w}_k^T . \quad (5.49)$$

Die Bildung des Erwartungswertes liefert

$$E[\vec{w}_{k+1} \vec{w}_k^T] = \Phi E[\vec{w}_k \vec{w}_k^T] + E[\vec{\eta}_k \vec{w}_k^T] . \quad (5.50)$$

Da das weiße Rauschen  $\vec{\eta}_k$  nicht mit dem Rauschterm  $\vec{w}_k$  korreliert ist, verschwindet der letzte Erwartungswert. Die verbleibenden Erwartungswerte können mit Hilfe der entsprechenden Stichprobenformeln geschätzt werden, liegt eine Stichprobe mit  $N$  Messwerten vor erhält man so

$$\underbrace{\frac{1}{N-1} \sum_{k=1}^N \vec{w}_{k+1} \vec{w}_k^T}_{\mathbf{A}} = \Phi \cdot \underbrace{\frac{1}{N-1} \sum_{k=1}^N \vec{w}_k \vec{w}_k^T}_{\mathbf{B}} . \quad (5.51)$$

Eine Schätzung der Transitionsmatrix ist daher gegeben durch

$$\hat{\Phi} = \mathbf{A} \mathbf{B}^{-1} . \quad (5.52)$$

Mit Hilfe dieser geschätzten Transitionsmatrix kann für jeden Zeitpunkt ein Schätzwert für den Rauschterm des weißen Rauschprozesses ermittelt werden:

$$\hat{\vec{\eta}}_k = \vec{w}_{k+1} - \hat{\Phi} \vec{w}_k \quad (5.53)$$

Die Varianz dieses weißen Rauschprozesses kann ebenfalls durch Anwendung der entsprechenden Stichprobenformel geschätzt werden.

■

Die gleiche Vorgehensweise lässt sich auch für Prozesse höherer Ordnung anwenden. Eine allgemeine Schätzvorschrift für die Parameter von Gauß-Markov-Prozessen beliebiger Ordnung ist durch die Yule-Walker-Gleichung gegeben, siehe [64].

Ein Gauß-Markov-Prozess  $n$ -ter Ordnung wird durch die Differenzengleichung

$$\vec{w}_k = \mathbf{A}_1 \vec{w}_{k-1} + \mathbf{A}_2 \vec{w}_{k-2} + \dots + \mathbf{A}_n \vec{w}_{k-n} + \vec{\eta}_k \quad (5.54)$$

beschrieben, die zugehörige Zustandsraumdarstellung ist gegeben durch

$$\begin{pmatrix} \vec{w}_k \\ \vec{w}_{k-1} \\ \vdots \\ \vec{w}_{k-n+1} \end{pmatrix} = \begin{pmatrix} \mathbf{A}_1 & \mathbf{A}_2 & \dots & \mathbf{A}_n \\ \mathbf{I} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \end{pmatrix} \begin{pmatrix} \vec{w}_{k-1} \\ \vec{w}_{k-2} \\ \vdots \\ \vec{w}_{k-n} \end{pmatrix} + \begin{pmatrix} \vec{\eta}_k \\ \vec{0} \\ \vdots \\ \vec{0} \end{pmatrix}. \quad (5.55)$$

Rauschprozesse, die sich mit diesem Ansatz modellieren lassen, werden auch als Vektor-autoregressive (VAR) Prozesse bezeichnet.