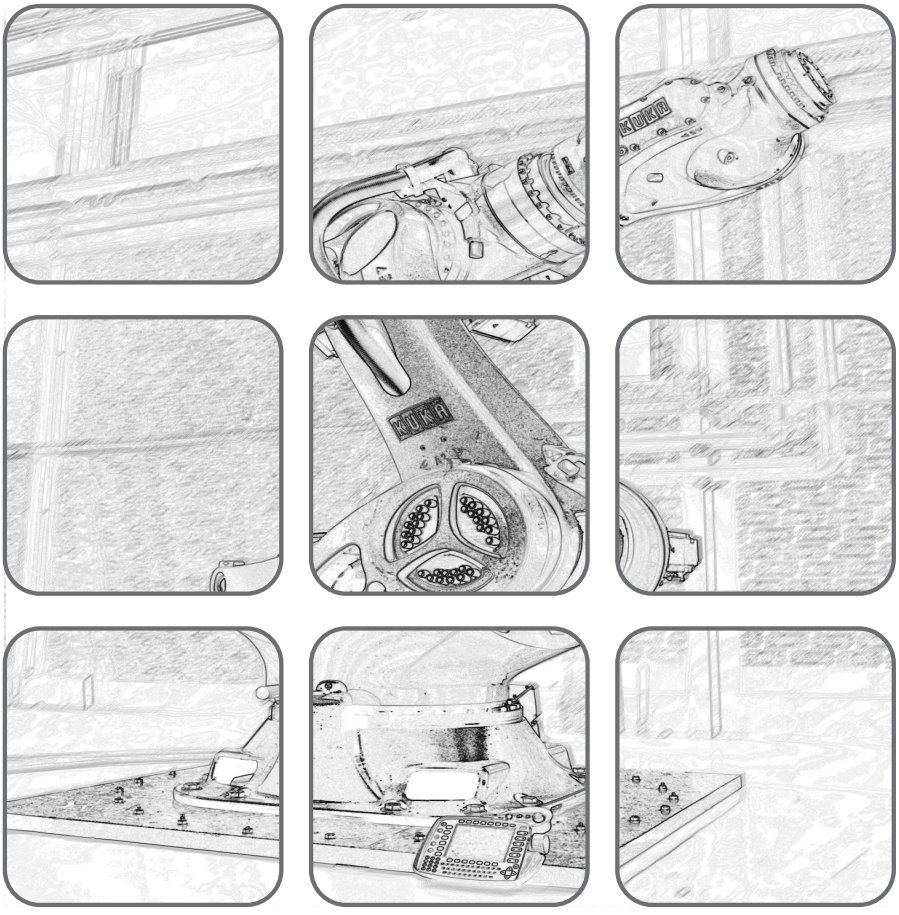
TECHNISCHE UNIVERSITÄT WIEN

AUTOMATION & CONTROL INSTITUTE
INSTITUT FÜR AUTOMATISIERUNGS- & REGELUNGSTECHNIK

Vorlesung
WS 2024/2025

REGELUNGSSYSTEME

Assoc.-Prof. Dr.-Ing. Wolfgang Kemmetmüller
Univ.-Prof. Dr. techn. Andreas Kugi



Regelungssysteme

Vorlesung
WS 2024/2025

Assoc.-Prof. Dr.-Ing. Wolfgang Kemmetmüller
Univ.-Prof. Dr. techn. Andreas Kugi

TU Wien
Institut für Automatisierungs- und Regelungstechnik
Gruppe für komplexe dynamische Systeme

Gußhausstraße 27–29
1040 Wien
Telefon: +43 1 58801 – 37615
Internet: https://www.acin.tuwien.ac.at

Inhaltsverzeichnis

1 Identifikationsverfahren

Dieses Kapitel beschäftigt sich mit der Prozessidentifikation linearer dynamischer Systeme. Nach einer kurzen Einführung in das Konzept der Identifikation werden einige wesentliche Vertreter von nicht-parametrischen und parametrischen Identifikationsmethoden diskutiert. Es sei bereits an dieser Stelle angemerkt, dass sich die verwendeten Begriffe sehr stark an dem Buch von L. Ljung [1.1] orientieren, da dieses Buch auch die Grundlage der MATLAB-Identification-Toolbox, in der alle in diesem Kapitel diskutierten und darüberhinaus noch wesentlich mehr Algorithmen zur Identifikation implementiert sind, bildet.

1.1 Allgemeines

Abbildung 1.1 zeigt die prinzipielle Aufgabe der Prozessidentifikation. Auf den Prozess

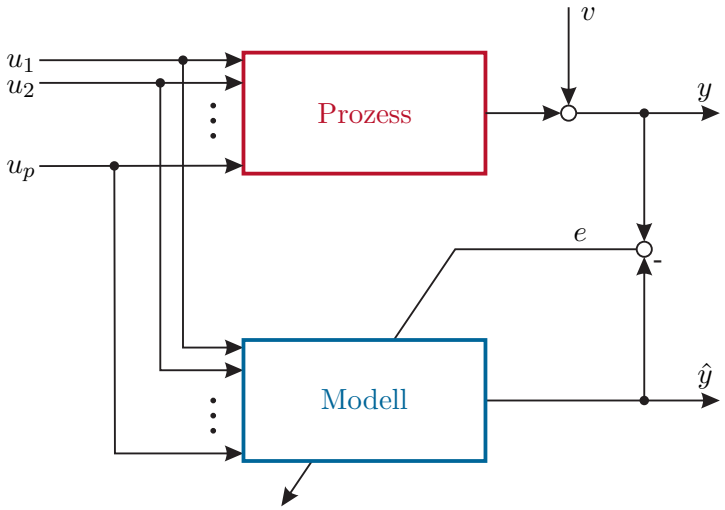


Abbildung 1.1: Zum Konzept der Prozessidentifikation.

wirken die Eingangsgrößen $u_1, \ldots, u_p$ und es wird angenommen, dass nur eine Ausgangsgröße y gemessen wird. Diese gemessene Ausgangsgröße y ist durch das Rauschsignal v gestört. Das Prozessmodel soll nun den Prozess so gut wie möglich beschreiben, wobei hier natürlich möglichst viel a priori Wissen über den Prozess berücksichtigt werden soll. Die Qualität des Modells wird dann beispielsweise anhand des Fehlers e zwischen dem gestörten gemessenen Ausgang y und dem Modellausgang $\hat{y}$ bewertet. Dieser Fehler dient in weiterer Folge dazu, das Modell zu verbessern. Die verschiedenen Schritte, die im Rahmen der Identifikationsaufgabe durchzuführen sind, sind in Abbildung 1.2 zusammengefasst und müssen je nach Problem mehrmals iterativ durchlaufen werden:

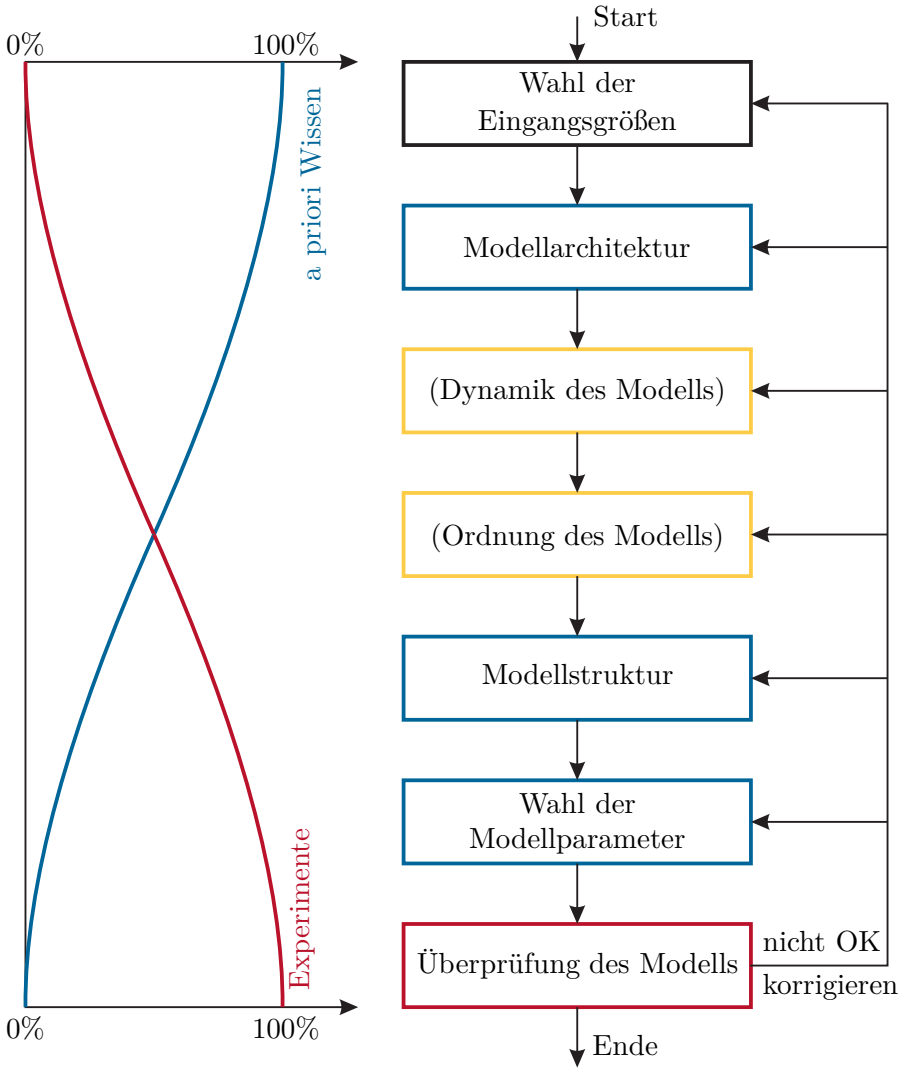


Abbildung 1.2: Ablauf der Identifikation.

(A) **Wahl der Eingangsgrößen:** In diesem Schritt muss entschieden werden, welche Eingangsgrößen für die Prozessidentifikation verwendet werden und welche Signale sich für die einzelnen Eingangsgrößen eignen.

(B) **Modellarchitektur:** Hier sollte festgelegt werden, ob es sich um ein statisches oder dynamisches Modell handelt, für welchen Zweck das Modell benötigt wird (Simulation, Reglerentwurf, Fehlererkennung etc.), der daraus resultierende Dynamikbereich, ob die Identifikation on- oder off-line erfolgen soll usw.

(C,D) **Dynamik des Modells** und **Ordnung des Modells:** In diesen Schritten erfolgt eine Beschreibung der Systemdynamik, beispielsweise in Form einer Übertragungsfunktion oder eines Zustandmodells, und es sollte eine Systemordnung festgelegt werden, die aber, falls nicht durch a priori Wissen bekannt, nur im Rahmen von gezielten Experimenten sinnvoll geschätzt werden kann.

(E) **Modellstruktur:** Im Rahmen dieses Schrittes wird die der Identifikationsaufgabe zugrundeliegende Modellstruktur festgelegt. Es sei an dieser Stelle auf Abschnitt 1.3.1 für die unterschiedlichen Modelle von parametrischen Identifikationsverfahren linearer dynamischer Systeme (ARMA, ARX, ARMAX etc.) verwiesen.

(F) **Wahl der Modellparameter:** Oft liegen die Modellparameter durch die vorigen Schritte, insbesondere die Wahl der Modellstruktur sowie der Systemordnung, bereits fest. Es besteht jedoch hier noch die Möglichkeit, gezielt Parameter angepasst an die Identifikationsaufgabe auszuwählen.

(G) **Überprüfung des Modells:** Je nach Zweck des Modells (Schritt (B)) muss überprüft werden, ob das Modell die entsprechende Güte aufweist. Hier ist besonders darauf zu achten, dass zur Überprüfung des Modells nicht der gleiche Datensatz wie zur Lösung der Identifikationsaufgabe verwendet werden darf.

Man unterscheidet in der Literatur noch zwischen *white-box*, *black-box* und *grey-box* Modellen:

- Bei white-box Modellen lassen sich sämtliche Gleichungen und Parameter auf Basis physikalischer Überlegungen gewinnen. Man spricht auch dann noch von white-box Modellen, wenn das Modell komplett über physikalische Gesetze hergeleitet wird und einige so genannte konstitutive Parameter (Reibungsparameter, Leckageparameter, Streuinduktivitäten, usw.) aus Experimenten ermittelt werden. Die Vorteile dieser Modelle bestehen in einer sehr guten Extrapolierbarkeit des Modells über die durch Experimente gewonnenen Daten hinaus, einer hohen Zuverlässigkeit, einer guten Einsicht in das Modell, sowie in der Skalierbarkeit des Modells, womit es auch für noch nicht realisierte Systeme (Prototyping) anwendbar ist. Als Nachteil von white-box Modellen kann angegeben werden, dass die Erstellung im Allgemeinen relativ zeitintensiv ist und man eine genaue Kenntnis des Systems benötigt.

- Black-box Modelle stützen sich lediglich auf experimentelle Ergebnisse und haben kein (oder sehr wenig) a priori Wissen des Systems. Natürlich sollte man sich bewusst

sein, dass das so gewonnene Modell nur in dem durch die Identifikation abgedeckten Datensatz Gültigkeit hat. Der Hauptvorteil besteht darin, dass man relativ wenig Wissen über das System benötigt. Sämtliche Vorteile der white-box Modelle können hier als Nachteile angegeben werden.

- Die grey-box Modelle kombinieren die Modellbildung auf Basis physikalischer Überlegungen mit Prinzipien der Prozessidentifikation.

1.2 Nicht-parametrische Verfahren

Lineare, zeitinvariante Systeme lassen sich durch ihre Übertragungsfunktion (s-Übertragungsfunktion im Zeitkontinuierlichen und z-Übertragungsfunktion im Zeitdiskreten) bzw. ihre Impulsantwort charakterisieren. Das Ziel der *nicht-parametrischen Identifikationsmethoden* ist nun die Bestimmung des Frequenzgangs bzw. der Impulsantwort direkt über Messungen der Ein- und Ausgangsgrößen, ohne dabei auf eine bestimmte der Identifikationsaufgabe zugrundeliegende Modellstruktur zurückzugreifen. Da diese Methoden nicht auf eine finite Anzahl von zu identifizierenden Parametern ausgelegt sind, werden sie als nicht-parametrisch bezeichnet. Den nachfolgenden Betrachtungen liegt ein System der Form von Abbildung 1.3 mit der deterministischen Eingangsfolge (u_k), der stochastischen Störung (Rauschen) (v_k), der ungestörten Ausgangsfolge $(\bar{y}_k)$, der Ausgangsfolge (y_k) sowie der BIBO-stabilen z-Übertragungsfunktion $G(z)$ mit der Impulsfolge $(g_k) = \mathcal{Z}^{-1}\{G(z)\}$ zugrunde.

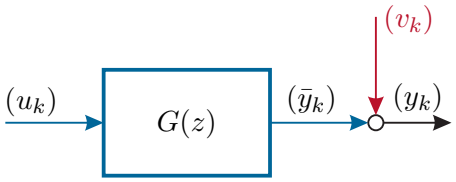


Abbildung 1.3: Betrachtetes System für die nicht-parametrische Identifikation.

Für die Ausgangsfolge gilt nun

$$y_k = \sum_{i=0}^{k} g_{k-i} u_i + v_k \tag{1.1}$$

bzw. im z-Bereich

$$y_z(z) = G(z) u_z(z) + v_z(z) \tag{1.2}$$

mit $y_z(z)$, $v_z(z)$ bzw. $u_z(z)$ als z-Transformierte der Folgen (y_k), (v_k) und (u_k).

1.2.1 Impulsantwortanalyse

Wählt man als Eingangsgröße (u_k) die Impulsfolge

$$u_k = \delta_k = \begin{cases} \alpha & \text{für } k = 0 \\ 0 & \text{für } k > 0, \end{cases} \tag{1.3}$$

dann folgt die Ausgangsfolge zu

$$y_k = \alpha g_k + v_k \ . \tag{1.4}$$

Wenn das Rauschen klein gegenüber der Impulsantwort ist, also $|v_k| \ll |\alpha g_k|$, dann kann die Impulsantwort in der Form

$$\hat{g}_k = \frac{y_k}{\alpha} \tag{1.5}$$

aus den gemessenen Folgenwerten y_k ermittelt werden. Das Problem dieser Methode besteht darin, dass sehr viele technische Systeme keine impulsförmigen Eingänge zulassen.

1.2.2 Frequenzbereichsanalyse - Anregung mit einer harmonischen Funktion

Für das System von Abbildung 1.3 erhält man im eingeschwungenen Zustand als Antwort auf die harmonische Eingangsfolge

$$(u_k) = (U \sin(\omega_0 k T_a)) \tag{1.6}$$

mit der Abtastzeit T_a die harmonische Ausgangsfolge (y_k)

$$y_k = Y \sin(\omega_0 k T_a + \varphi) + v_k \tag{1.7}$$

mit

$$Y = U\left|G\left(\mathrm{e}^{\mathrm{I}\omega_0 T_a}\right)\right| \quad \text{und} \quad \varphi = \arg\left(G\left(\mathrm{e}^{\mathrm{I}\omega_0 T_a}\right)\right) \tag{1.8}$$

bzw.

$$y_k = Y_c \cos(\omega_0 k T_a) + Y_s \sin(\omega_0 k T_a) + v_k \tag{1.9}$$

mit

$$Y_c = Y \sin(\varphi) \quad \text{und} \quad Y_s = Y \cos(\varphi) \ . \tag{1.10}$$

In Anbetracht von (1.7)-(1.10) ist es naheliegend, für den Schätzwert $\hat{y}_k$ der gemessenen Ausgangsgröße y_k einen Ansatz der Form

$$\hat{y}_k = \hat{Y} \sin(\omega_0 k T_a + \hat{\varphi}) = \hat{Y}_c \cos(\omega_0 k T_a) + \hat{Y}_s \sin(\omega_0 k T_a) \tag{1.11}$$

mit den zu schätzenden Parametern $\hat{Y}_c = \hat{Y} \sin(\hat{\varphi})$ und $\hat{Y}_s = \hat{Y} \cos(\hat{\varphi})$ zu wählen. Offensichtlich können dann die Parameter $\hat{Y}$ und $\hat{\varphi}$ sehr einfach über die Beziehungen

$$\hat{Y} = \sqrt{\hat{Y}_c^2 + \hat{Y}_s^2} \quad \text{und} \quad \hat{\varphi} = \arctan\left(\frac{\hat{Y}_c}{\hat{Y}_s}\right) \tag{1.12}$$

bestimmt werden. Es ist nun zu klären, durch welche Wahl von $\hat{Y}_c$ und $\hat{Y}_s$ der quadratische Fehler

$$L = \frac{1}{N} \sum_{k=0}^{N-1} (y_k - \hat{y}_k)^2 \tag{1.13}$$

für N Messungen von y_k minimal wird. Durch Einsetzen von (1.11) in (1.13), Ableiten nach $\hat{Y}_c$ und Nullsetzen erhält man

$$\frac{\partial L}{\partial \hat{Y}_c} = -\frac{2}{N} \sum_{k=0}^{N-1} \Big(y_k - \hat{Y}_c \cos(\omega_0 k T_a) - \hat{Y}_s \sin(\omega_0 k T_a)\Big) \cos(\omega_0 k T_a) = 0 \tag{1.14}$$

bzw. mit $(\cos(\alpha))^2 = \frac{1}{2}(1+\cos(2\alpha))$ und $\cos(\alpha)\sin(\alpha) = \frac{1}{2}\sin(2\alpha)$

$$\hat{Y}_c - \frac{2}{N}\sum_{k=0}^{N-1} y_k \cos(\omega_0 k T_a) + \frac{\hat{Y}_c}{N}\sum_{k=0}^{N-1}\cos(2\omega_0 k T_a) + \frac{\hat{Y}_s}{N}\sum_{k=0}^{N-1}\sin(2\omega_0 k T_a) = 0\ . \tag{1.15}$$

Mit Hilfe der Euler-Formeln $\cos(\alpha) = \frac{1}{2}\left(\mathrm{e}^{\mathrm{I}\alpha} + \mathrm{e}^{-\mathrm{I}\alpha}\right)$ und $\sin(\alpha) = \frac{1}{2\mathrm{I}}\left(\mathrm{e}^{\mathrm{I}\alpha} - \mathrm{e}^{-\mathrm{I}\alpha}\right)$ lassen sich die zweite und dritte Summe von (1.15) wie folgt

$$\sum_{k=0}^{N-1}\cos(2\omega_0 k T_a) = \frac{1}{2}\sum_{k=0}^{N-1}\left(\mathrm{e}^{\mathrm{I}2\omega_0 k T_a} + \mathrm{e}^{-\mathrm{I}2\omega_0 k T_a}\right) \tag{1.16a}$$

$$\sum_{k=0}^{N-1}\sin(2\omega_0 k T_a) = \frac{1}{2\mathrm{I}}\sum_{k=0}^{N-1}\left(\mathrm{e}^{\mathrm{I}2\omega_0 k T_a} - \mathrm{e}^{-\mathrm{I}2\omega_0 k T_a}\right) \tag{1.16b}$$

umschreiben. Setzt man nun für die Kreisfrequenz ω_0 nur diskrete Werte der Form

$$\omega_0 = \frac{2\pi l}{N T_a} \quad \text{mit} \quad l = 1, 2, \ldots \tag{1.17}$$

an, dann erhält man unter Zuhilfenahme der Beziehung

$$\sum_{k=0}^{N-1} z^{-k} = \frac{1 - z^{-N}}{1 - z^{-1}} \tag{1.18}$$

folgenden Ausdruck für die Summe

$$\sum_{k=0}^{N-1}\mathrm{e}^{-\mathrm{I}2\omega_0 k T_a} = \sum_{k=0}^{N-1}\mathrm{e}^{-\mathrm{I}\frac{4\pi l k}{N}} = \frac{1-\mathrm{e}^{-\mathrm{I}4\pi l}}{1-\mathrm{e}^{-\mathrm{I}\frac{4\pi l}{N}}} = \begin{cases} N & \text{für } l = \frac{r}{2}N, r = \pm 1, \pm 2, \ldots \\ 0 & \text{für } l = \pm 1, \pm 2, \ldots \end{cases} \tag{1.19}$$

Das analoge Ergebnis erhält man natürlich, wenn man I durch $-\mathrm{I}$ ersetzt. Für ω_0 gemäß (1.17) errechnen sich demnach die Ausdrücke (1.16) zu

$$\sum_{k=0}^{N-1}\cos(2\omega_0 k T_a) = \begin{cases} N & \text{für } l = \frac{r}{2}N, r = \pm 1, \pm 2, \ldots \\ 0 & \text{für } l = \pm 1, \pm 2, \ldots \end{cases} \tag{1.20a}$$

$$\sum_{k=0}^{N-1}\sin(2\omega_0 k T_a) = 0 \tag{1.20b}$$

Wählt man die Testfrequenz ω_0 gemäß (1.17) so, dass $l < N/2$ ist, dann ist auch der Ausdruck in (1.20a) Null und die optimale Lösung $\hat{Y}_c$ errechnet sich aus (1.15) in der Form

$$\hat{Y}_c = \frac{2}{N}\sum_{k=0}^{N-1} y_k \cos\left(\frac{2\pi l}{N}k\right)\ . \tag{1.21}$$

Bemerkung 1.1. Man beachte, dass bei vorgegebener Abtastzeit T_a die maximal mögliche Frequenz einer eindeutig darstellbaren harmonischen Funktion echt kleiner der Grenzfrequenz $\omega_{\max} = \pi/T_a$ sein muss. Es ist unmittelbar zu erkennen, dass der Wert $l = N/2$ eingesetzt in ω_0 von (1.17) gerade dieser Grenzfrequenz entspricht.

Aufgabe 1.1. Zeigen Sie, dass sich das optimale $\hat{Y}_s$ wie folgt

$$\hat{Y}_s = \frac{2}{N}\sum_{k=0}^{N-1} y_k \sin\left(\frac{2\pi l}{N}k\right) \tag{1.22}$$

errechnet.

Dies lässt sich nun wie folgt zusammenfassen: Regt man das System von Abbildung 1.3 mit der harmonischen Folge

$$(u_k) = (U\sin(\omega_0 k T_a)) \quad \text{mit} \quad \omega_0 = \frac{2\pi l}{NT_a}, l = 1, 2, \ldots, \frac{N}{2} - 1 \tag{1.23}$$

an, dann kann man aus den N Messwerten y_k, $k = 0, \ldots, N-1$, den diskreten Frequenzgang $G\left(\mathrm{e}^{\mathrm{I}\omega_0 T_a}\right)$ an den Frequenzen $\omega_0 = \frac{2\pi l}{NT_a}$, $l = 1, 2, \ldots, \frac{N}{2} - 1$ über die Beziehungen (siehe (1.7),(1.8))

$$\hat{Y}_c = U\left|\hat{G}\left(\mathrm{e}^{\mathrm{I}\omega_0 T_a}\right)\right| \sin\left(\arg\left(\hat{G}\left(\mathrm{e}^{\mathrm{I}\omega_0 T_a}\right)\right)\right) \tag{1.24a}$$

$$\hat{Y}_s = U\left|\hat{G}\left(\mathrm{e}^{\mathrm{I}\omega_0 T_a}\right)\right| \cos\left(\arg\left(\hat{G}\left(\mathrm{e}^{\mathrm{I}\omega_0 T_a}\right)\right)\right) \tag{1.24b}$$

wie folgt

$$\left|\hat{G}\left(\mathrm{e}^{\mathrm{I}\omega_0 T_a}\right)\right| = \frac{\sqrt{\hat{Y}_s^2 + \hat{Y}_c^2}}{U} \tag{1.25a}$$

$$\arg\left(\hat{G}\left(\mathrm{e}^{\mathrm{I}\omega_0 T_a}\right)\right) = \arctan\left(\frac{\hat{Y}_c}{\hat{Y}_s}\right) \tag{1.25b}$$

mit $\hat{Y}_c$ und $\hat{Y}_s$ aus (1.21) bzw. (1.22) approximieren.

Die obigen Beziehungen (1.21) und (1.22) hängen eng mit der *diskreten Fourier-Transformation (DFT)* der Folgen (u_k) und (y_k) zusammen. Zur Wiederholung sei die diskrete Fourier-Transformierte $F_n(\omega)$ einer Folge (f_k)

$$F_n(\omega) = \sum_{k=0}^{N-1} f_k \mathrm{e}^{-\mathrm{I}\omega k T_a} \quad \text{mit} \quad \omega = \frac{2\pi n}{NT_a}, \quad n = 0, 1, \ldots, N-1 \tag{1.26}$$

sowie die inverse diskrete Fourier-Transformierte (IDFT)

$$f_k = \frac{1}{N}\sum_{n=0}^{N-1} F_n \mathrm{e}^{\mathrm{I}\omega k T_a} \quad \text{mit} \quad \omega = \frac{2\pi n}{NT_a}, \quad k = 0, 1, \ldots, N-1 \tag{1.27}$$

angegeben. Die DFT der Eingangsfolge (u_k) nach (1.23) sowie die DFT der gemessenen Ausgangsfolge (y_k) lautet unter Berücksichtigung von (1.21) und (1.22) mit $\omega_0 = \frac{2\pi l}{NT_a}$

$$\begin{aligned} U_n &= \sum_{k=0}^{N-1} u_k \mathrm{e}^{-\mathrm{I}\omega k T_a} = \frac{U}{2\mathrm{I}} \sum_{k=0}^{N-1} \left(\mathrm{e}^{\mathrm{I}\omega_0 k T_a} - \mathrm{e}^{-\mathrm{I}\omega_0 k T_a}\right) \mathrm{e}^{-\mathrm{I}\omega k T_a} \\ &= \frac{U}{2\mathrm{I}} \sum_{k=0}^{N-1} \left(\mathrm{e}^{\mathrm{I}\frac{2\pi k}{N}(l-n)} - \mathrm{e}^{-\mathrm{I}\frac{2\pi k}{N}(l+n)}\right) = \begin{cases} \frac{NU}{2\mathrm{I}} & \text{für } l = n \\ 0 & \text{sonst} \end{cases} \end{aligned} \tag{1.28}$$

und

$$Y_n = \sum_{k=0}^{N-1} y_k \mathrm{e}^{-\mathrm{I}\omega k T_a} = \sum_{k=0}^{N-1} y_k \left(\cos\left(\frac{2\pi n}{N} k\right) - \mathrm{I}\sin\left(\frac{2\pi n}{N} k\right)\right) = \frac{N}{2}\left(\hat{Y}_c - \mathrm{I}\hat{Y}_s\right) . \tag{1.29}$$

Aufgabe 1.2. Beweisen Sie die Beziehung von (1.28).

Damit sieht man, dass die Schätzung des diskreten Frequenzganges $G\left(\mathrm{e}^{\mathrm{I}\omega_0 T_a}\right)$ an den Frequenzen $\omega_0 = \frac{2\pi l}{NT_a}$, $l = 1, 2, \ldots, \frac{N}{2} - 1$ nach (1.25) sehr einfach mit Hilfe der diskreten Fourier-Transformierten der Eingangs- und Ausgangsfolgen gemäß (1.28) und (1.29) berechnet werden kann – es gilt nämlich

$$\frac{Y_n(\omega_0)}{U_n(\omega_0)} = \frac{\mathrm{DFT}((y_k))}{\mathrm{DFT}((u_k))} = \frac{\frac{N}{2}\left(\hat{Y}_c - \mathrm{I}\hat{Y}_s\right)}{\frac{NU}{2\mathrm{I}}} = \frac{\sqrt{\hat{Y}_c^2 + \hat{Y}_s^2}}{U} \mathrm{e}^{\mathrm{I}\arctan\frac{\hat{Y}_c}{\hat{Y}_s}} . \tag{1.30}$$

Für die praktische Anwendung ist es natürlich zweckmäßig, die Berechnung der diskreten Fourier-Transformationen mit Hilfe des *effizienteren FFT (Fast Fourier Transformation)* Algorithmus durchzuführen.

1.2.3 Frequenzbereichsanalyse - ETFE

Die Beziehung (1.30) legt nun nahe, als Eingangsfolge (u_k) nicht nur ein harmonisches Signal mit konstanter Frequenz ω_0 zu wählen, sondern ein Signal, welches mehrere Frequenzen beinhaltet, um sich damit das Superpositionsprinzip linearer System zu Nutze zu machen. Man nennt dann die Schätzung des Frequenzganges nach Beziehung (1.30) in der englischsprachigen Literatur *empirical transfer function estimate (ETFE)* – man beachte dazu auch den gleichnamigen MATLAB-Befehl. Die Schätzung wird deshalb als empirisch bezeichnet, da außer der Annahme der Linearität und Zeitinvarianz des Systems keine weiteren Annahmen über die Modellstruktur getroffen werden. Ist für bestimmte Frequenzen ω zufolge der Anregung (u_k) die Fourier-Transformierte $U(\omega) = 0$, dann ist die ETFE an diesen Frequenzen nicht definiert. Beispiele für geeignete Eingangsfolgen sind Impulsfolgen, der 3-2-1-Sprung, Chirp-Signale oder PRBS-Signale.

Ein lineares Chirp-Signal mit trapezförmiger Fensterung ist dabei in der Form

$$u_k = U_0 + r_k \sin\left(\omega_{start} k T_a + \frac{(\omega_{end} - \omega_{start})}{NT_a} \frac{(kT_a)^2}{2}\right) \tag{1.31a}$$

$$r_k = U \mathrm{sat}\left(\frac{10k}{N}\right) \mathrm{sat}\left(\frac{10(N-k)}{N}\right) \tag{1.31b}$$

mit

$$\mathrm{sat}(x) = \begin{cases} 1 & \text{für } x \geq 1 \\ x & \text{für } -1 < x < 1 \\ -1 & \text{für } x \leq -1 \end{cases} \tag{1.32}$$

und $k = 0, 1, \ldots, N-1$, für die Abtastzeit T_a, der Anzahl der Abtastpunkte N, der Konstanten U_0 zur Kompensation eines Offsets der Amplitude U sowie der unteren und oberen Chirpfrequenz ω_{start} und ω_{end} gegeben.

> *Aufgabe* 1.3. Zeichnen Sie den zeitlichen Verlauf des Chirp-Signals (1.31) und den Verlauf der zugehörigen diskreten Fourier-Transformierten für $N = 256$, $U = 1$, $\omega_{start} = 0.1$ sowie $\omega_{end} = 5/2$, $T_a = 0.5\,\mathrm{s}$ in MATLAB.

Sogenannte PRBS (Pseudo Random Binary Signal) Verläufe werden häufig in Identifikationsaufgaben verwendet, da sie ähnliche Eigenschaften wie weißes Rauschen aufweisen. Ein PRBS-Signal der Ordnung O_p kann durch die Differenzengleichung

$$p_k = \mathrm{mod}\Big(a_1 p_{k-1} + a_2 p_{k-2} + \ldots + a_{O_p} p_{k-O_p}, 2\Big) \tag{1.33}$$

mit den Koeffizienten $a_j \in \{0, 1\}$, $j = 1, \ldots, O_p$, welche für unterschiedliche Ordnungen in Tabelle 1.1 dargestellt sind, ermittelt werden. Man beachte, dass sich das PRBS-Signal alle $2^{O_p} - 1$ Abtastschritte wiederholt.

Ordnung O_p	$a_j \neq 0$ für folgende j
2	$1, 2$
3	$2, 3$
4	$1, 4$
5	$2, 5$
6	$1, 6$
7	$3, 7$
8	$1, 2, 7, 8$
9	$4, 9$
10	$7, 10$
11	$9, 11$

Tabelle 1.1: Koeffizienten des PRBS-Signals.

Um die Frequenzeigenschaften des PRBS-Signals zu beeinflussen, ist häufig eine Überabtastung des Signals p_k mit dem Faktor $P_p \geq 1$ sinnvoll. Ist eine Abtastzeit T_a gegeben, dann werden die Werte von p_k entsprechend P_p-mal aufgeschaltet, d. h. es gilt

$$u((P_p k + j)T_a) = U p_k, \quad j = 0, \ldots, P_p - 1, \quad k = 0, \ldots, 2^{O_p} - 2 \tag{1.34}$$

mit der Amplitude U des Signals. Es kann gezeigt werden, dass damit im Frequenzbereich eine Tiefpassfilterung des PRBS-Signals erfolgt. In der Literatur wird typischerweise eine Überabtastung von $P_p = 4$ empfohlen. Bei einer Ordnung O_p und einer Überabtastung P_p steht damit eine Anzahl von $N = (2^{O_p} - 1)P_p$ Abtastwerten zur Verfügung.

Aufgabe 1.4. Zeichnen Sie den zeitlichen Verlauf des PRBS-Signals (1.34) und den Verlauf der zugehörigen diskreten Fourier-Transformierten für $O_p = 10$, $U = 1$ und einer Abtastzeit $T_a = 0.5$s in MATLAB. Untersuchen Sie dabei den Einfluss unterschiedlicher Werte für die Überabtastung P_p und initialisieren Sie die Differenzengleichung (1.34) mit $p_0 = 1$.

Aufgabe 1.5. Angenommen, die Folgen (u_k) und (y_k) bestehen aus N äquidistanten Abtastpunkten mit der Abtastzeit T_a. Zeigen Sie, dass die minimal auflösbare Frequenz durch $\omega_{\min} = \frac{2\pi}{NT_a}$ und die maximal auflösbare Frequenz durch $\omega_{\max} = \frac{\pi}{T_a}$ gegeben sind.

Aufgabe 1.6. Schreiben Sie unter Verwendung der `fft`-Routine von MATLAB ein Programm zur Berechnung des Frequenzganges eines linearen, zeitinvarianten Abtastsystems mit der Übertragungsfunktion $G(z)$ gemäß (1.30). Als Eingabeparameter sollen dabei die Abtastfolgen der Eingangs- und Ausgangsgröße (u_k) bzw. (y_k), die Abtastzeit T_a und die Anzahl der Messpunkte N dienen und als Ergebnis sollen die diskreten Frequenzen $\omega = \frac{2\pi}{NT_a}n$, $\left|G\left(\mathrm{e}^{\mathrm{I}\omega T_a}\right)\right|$ und $\arg\left(G\left(\mathrm{e}^{\mathrm{I}\omega T_a}\right)\right)$ für $n = 1, \ldots, \frac{N}{2} - 1$ geliefert werden. Testen Sie das Programm anhand der Strecke

$$G(s) = \frac{1}{s^2 + 0.25s + 1}$$

mithilfe eines Chirp-Signals nach (1.31), (1.32).

Die Ergebnisse der ETFE sind im Allgemeinen sehr gut für deterministische Eingangssignale, insbesondere für jene Frequenzen, die durch das Eingangssignal hinreichend gut angeregt werden. Für stochastische Eingangssignale müssen meist zusätzliche Maßnahmen, wie beispielsweise das Glätten der ETFE durch geeignete Fensterfunktionen (Hamming, Bartlett, Kaiser etc.), getroffen werden, um brauchbare Ergebnisse zu erhalten. Im Rahmen dieser Vorlesung wird auf diese Details nicht eingegegangen, der (die) interessierte Student(in) sei auf die am Ende angegebene Literatur, insbesondere das Buch von L. Ljung [1.1], verwiesen.

Bei praktischen Anwendungen hat es sich als sehr hilfreich erwiesen, *a priori Wissen* über den Prozess im Rahmen der Identifikation zu berücksichtigen. Dies kann beispielsweise dadurch geschehen, dass man das Eingangssignal so wählt, dass es das System ganz gezielt in dem interessierenden Frequenzbereich anregt. Eine weitere Möglichkeit besteht darin, bekannte Teile der zu identifizierenden Übertragungsfunktion mit zu berücksichtigen. Dazu wird die Übertragungsfunktion $G(z)$ in einen bekannten und einen unbekannten

Anteil in der Form

$$G(z) = \underbrace{\frac{z_b(z)}{n_b(z)}}_{\text{bekannt}} \underbrace{G_1(z)}_{\text{unbekannt}} \tag{1.35}$$

faktorisiert und anschließend die Identifikationsaufgabe für $G_1(z)$ gemäß Abbildung 1.4 gelöst.

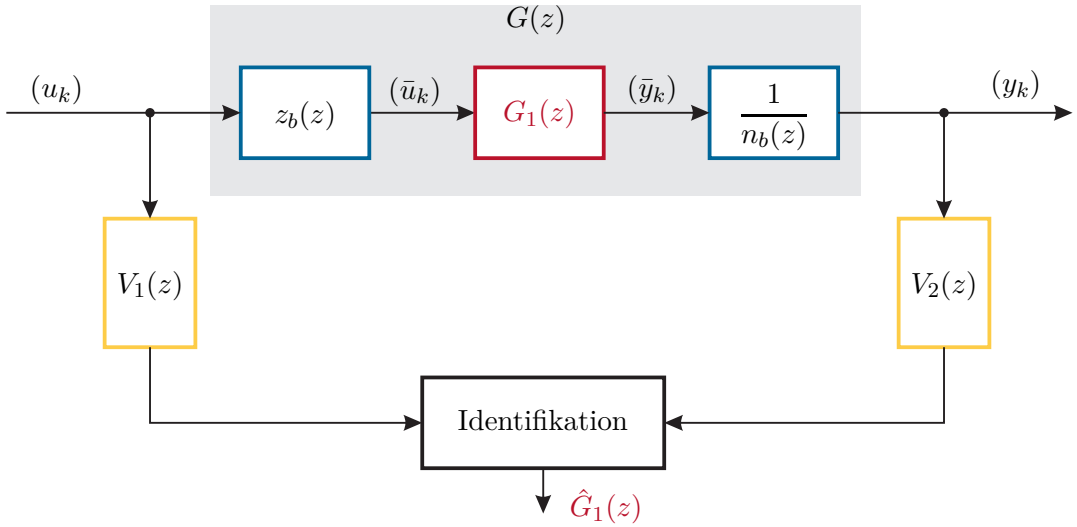


Abbildung 1.4: Identifikation unter Berücksichtigung bekannter Teile der Übertragungsfunktion (Verfahren nach Clary).

Diese Methode, auch *Verfahren nach Clary* genannt, hat sich bei der Identifikation gewisser mechanischer Systeme als sehr zielführend erwiesen. Bei diesen Systemen weiß man beispielsweise, dass die s-Übertragungsfunktion $G(s)$ eine doppelte Polstelle bei $s_i = 0$ aufweist. Damit liegt aber bereits fest, dass die zu identifizierende z-Übertragungsfunktion gemäß der Beziehung $z = \exp(sT_a)$ mit der Abtastzeit T_a eine doppelte Polstelle bei $z_i = 1$ besitzt. Diese direkte Zuordnung der Polstellen von s- und z-Übertragungsfunktionen lässt sich bekanntermaßen nicht auf deren Nullstellen übertragen. Insbesondere die Nullstellen außerhalb des Einheitskreises haben meist nur einen geringen Einfluss auf den Frequenzgang im interessierenden Frequenzbereich, sind aber extrem schwierig zu identifizieren. Eine nicht exakte, aber wirksame Vorgangsweise ist, diesen Nullstellen z_j außerhalb des Einheitskreises den Wert $z_j = -1$ zuzuordnen. Damit die Folgen (u_k) und (y_k) gemäß Abbildung 1.4 für die Identifikation aufbereitet werden können, müssen die einzelnen Vorfilter $V_1(z)$ und $V_2(z)$ proper sein, d. h. der Zählergrad muss kleiner gleich dem Nennergrad sein. Um dies zu gewährleisten, wird der bekannte Teil der Übertragungsfunktion von (1.35) in der Form

$$\frac{z_b(z)}{n_b(z)} = \frac{z_b(z)}{z^n}\frac{z^n}{n_b(z)} \quad \text{mit} \quad n = \max(\operatorname{grad}(z_b(z)), \operatorname{grad}(n_b(z))) \tag{1.36}$$

angeschrieben. Die Vorfilter lauten dann

$$V_1(z) = \frac{z_b(z)}{z^n} \quad \text{und} \quad V_2(z) = \frac{n_b(z)}{z^n}\,. \tag{1.37}$$

Aufgabe 1.7. Identifizieren Sie die Streckenübertragungsfunktion

$$G(s) = \frac{1}{s^2}\frac{1}{s^2 + 0.25s + 1} \tag{1.38}$$

mit dem Clary-Verfahren für die Abtastzeit $T_a = 0.5\,\mathrm{s}$. Nehmen Sie dabei an, dass Sie vom System wissen, dass es einen Doppelintegrator besitzt. Verwenden Sie dabei als Eingangssignal das Chirp-Signal von Aufgabe 1.3.

1.3 Parametrische Verfahren

Wie der Name impliziert, werden bei diesen Verfahren Modelle mit einer finiten Anzahl von Parametern zugrunde gelegt. Diese Parameter werden dann im Rahmen der Identifikationsaufgabe so bestimmt, dass das Modell im Sinne eines Gütekriteriums bestmöglich mit der zu identifizierenden Strecke übereinstimmt. In der Literatur existiert eine Unmenge von unterschiedlichen Modellstrukturen, weswegen im ersten Schritt eine Klassifizierung dieser Modelle vorgenommen wird. Die nachfolgende Klassifizierung orientiert sich dabei im Wesentlichen an der MATLAB System–Identification–Toolbox.

1.3.1 Modellstrukturen

Ausgangspunkt der Betrachtungen ist wiederum das System von Abbildung 1.3. Für die weiteren Schritte ist es erforderlich, die stochastische Störung (das Rauschen) (v_k) in einer geeigneten Form zu charakterisieren. Ein relativ einfacher Zugang besteht darin, (v_k) als Ausgangsfolge eines linearen zeitinvarianten Systems mit der Übertragungsfunktion $H(z)$ und dem *weißen Rauschen* (w_k) (Folge von unabhängigen Zufallsvariablen) mit einer vorgegebenen Wahrscheinlichkeitsdichtefunktion zu modellieren, d. h.

$$v_k = \sum_{i=-\infty}^{k} h_{k-i} w_i = \sum_{i=0}^{\infty} h_i w_{k-i} \quad \text{mit} \quad (h_k) = \mathcal{Z}^{-1}\{H(z)\}\ . \tag{1.39}$$

Dieser Ansatz ist für die meisten praktischen Anwendungen ausreichenden, es kann jedoch nicht jede mögliche stochastische Störung (v_k) damit charakterisiert werden. Man beachte, dass durch die Wahl der Wahrscheinlichkeitsdichtefunktion von (w_k) der Charakter des stochastischen Störsignals (v_k) gezielt beeinflusst werden kann. Es ist beispielsweise naheliegend, dass für ein und dasselbe System mit der Übertragungsfunktion $H(z)$ ein weißes Rauschen $(w_{2,k})$ mit der Wahrscheinlichkeitsdichtefunktion

$$\begin{aligned} w_{2,k} &= 0 \quad \text{mit der Wahrscheinlichkeit } 1-\mu \\ w_{2,k} &= r \quad \text{mit der Wahrscheinlichkeit } \mu \end{aligned} \tag{1.40}$$

für ein sehr kleines μ und einer gleichverteilten Zufallsvariable $r \in (-1, 1)$ ein vollkommen anderes Bild für $(v_{2,k})$ liefert als ein weißes Rauschen $(w_{1,k})$ mit der Wahrscheinlichkeitsdichtefunktion $w_{1,k} = r$.

Man beachte dazu die Verläufe von $(v_{1,k})$ und $(v_{2,k})$ für ein $H(z) = \mathbf{Z}\{H(s)\}$, mit

$$H(s) = \frac{1}{\frac{s^2}{(2\pi 10)^2} + 2\frac{0.2s}{2\pi 10} + 1}, \tag{1.41}$$

in Abbildung 1.5, wobei $T_a = 10\text{ms}$ und $\mu = 0.07$ verwendet wurde.

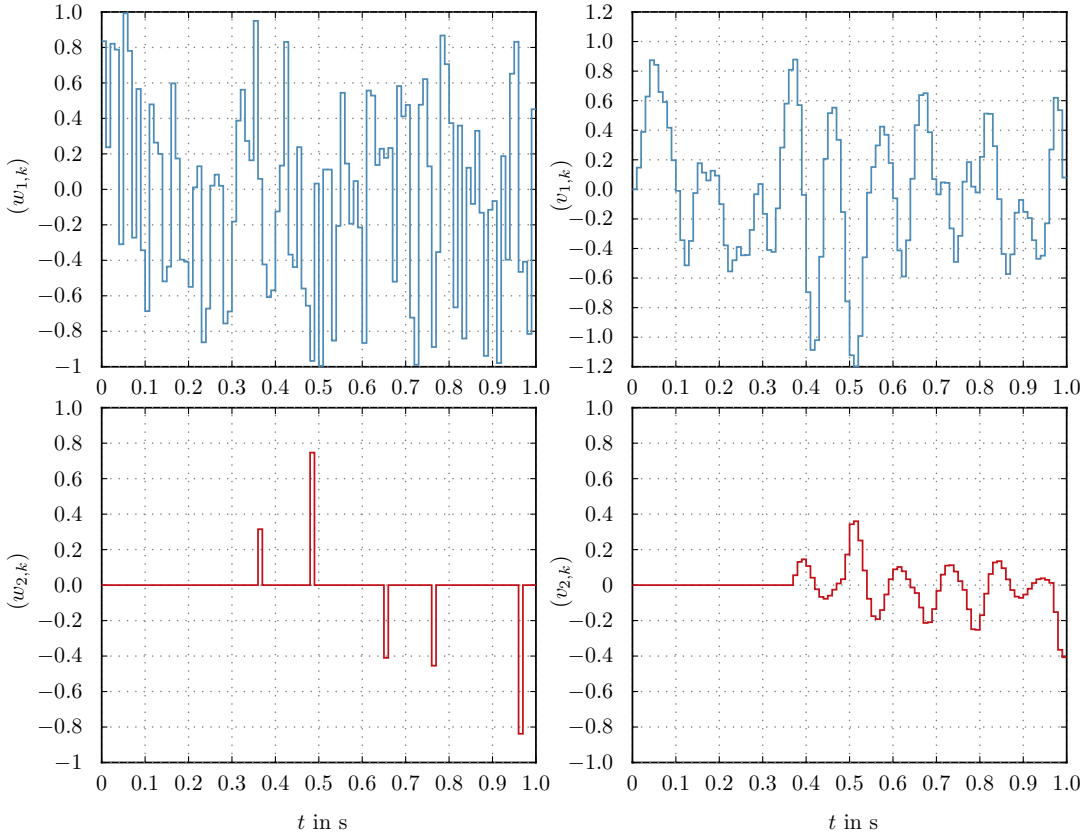


Abbildung 1.5: Antworten $(v_{1,k})$ und $(v_{2,k})$ des linearen, zeitinvarianten Abtastsystems zweiter Ordnung (1.41) auf weißes Rauschen $(w_{1,k})$ bzw. $(w_{2,k})$ mit unterschiedlicher Wahrscheinlichkeitsdichtefunktion.

Kombiniert man nun (1.1) mit (1.39), dann lautet die Ausgangsgröße y_k von Abbildung 1.3

$$y_k = \sum_{i=0}^{\infty} g_i u_{k-i} + \sum_{i=0}^{\infty} h_i w_{k-i} \tag{1.42}$$

bzw. durch Einführung des Shift-Operators δ mit $u_{k+1} = \delta u_k$ bzw. $u_{k-1} = \delta^{-1} u_k$ folgt

$$y_k = \sum_{i=0}^{\infty} g_i \delta^{-i} u_k + \sum_{i=0}^{\infty} h_i \delta^{-i} w_k \ . \tag{1.43}$$

Um sich in weiterer Folge Schreibarbeit zu ersparen, schreibt man für (1.43) auch

$$y_k = G(\delta) u_k + H(\delta) w_k \tag{1.44}$$

mit den Übertragungsoperatoren

$$G(\delta) = \sum_{i=0}^{\infty} g_i \delta^{-i} \quad \text{und} \quad H(\delta) = \sum_{i=0}^{\infty} h_i \delta^{-i} \ . \tag{1.45}$$

Man beachte, dass die Ausdrücke der Übertragungsoperatoren $G(\delta)$ und $H(\delta)$ gleich den z-Übertragungsfunktionen von linearen, zeitinvarianten Abtastsystemen mit den Impulsantworten (g_k) bzw. (h_k) sind. Würde man aber $G(\delta)$ und $H(\delta)$ als z-Übertragungsfunktionen interpretieren, wäre die Schreibweise von (1.44) nicht zulässig.

Durch Darstellung von $G(\delta)$ und $H(\delta)$ in Form von rationalen Übertragungsoperatoren mit den zugehörigen Zähler- und Nennerpolynomen sowie Zusammenfassung aller gemeinsamen Pole von $G(\delta)$ und $H(\delta)$ im Polynom $A(\delta)$ ergibt sich (1.44) zu (siehe Abbildung 1.6)

$$A(\delta) y_k = \frac{B(\delta)}{F(\delta)} u_k + \frac{C(\delta)}{D(\delta)} w_k \ . \tag{1.46}$$

Anhand von (1.46) lässt sich nun eine Klassifizierung der verschiedenen Modellstrukturen durchführen.

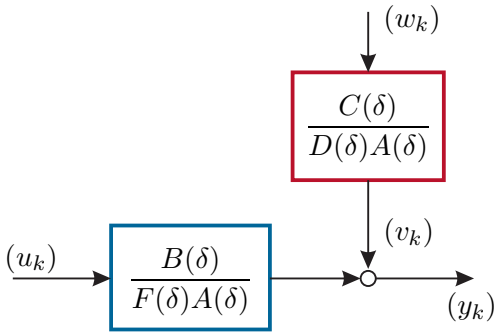


Abbildung 1.6: Zu den Modellstrukturen der parametrischen Identifikation.

ARMA-Modell

Bei Identifikationsanwendungen in den Wirtschaftswissenschaften hat man es häufig mit Modellen zu tun, die keine deterministische Eingangsgröße besitzen, bei denen also $(u_k) = (0)$ ist. Bei diesen so genannten *Zeitreihenmodellen* unterscheidet man zwischen folgenden Spezialfällen:

Das Modell

$$y_k = \frac{1}{D(\delta)} w_k \tag{1.47}$$

bzw. für $D(\delta) = d_0 + d_1\delta^{-1} + \ldots + d_n\delta^{-n}$ mit $d_0 = 1$

$$y_k = -d_1 y_{k-1} - d_2 y_{k-2} - \ldots - d_n y_{k-n} + w_k \tag{1.48}$$

wird als *Autoregressionsmodell* oder *AR (autoregressive) Modell* bezeichnet. Gleichung (1.48) kann auch in der Form

$$y_k = \mathbf{s}_k^{\mathrm{T}} \mathbf{p} \tag{1.49}$$

mit dem *Parametervektor* $\mathbf{p}$ und dem *Datenvektor* $\mathbf{s}_k$ in der Form

$$\mathbf{p}^{\mathrm{T}} = \begin{bmatrix} -d_1 & -d_2 & \ldots & -d_n & 1 \end{bmatrix} \tag{1.50a}$$

$$\mathbf{s}_k^{\mathrm{T}} = \begin{bmatrix} y_{k-1} & y_{k-2} & \ldots & y_{k-n} & w_k \end{bmatrix} \tag{1.50b}$$

angeschrieben werden. In der Statistik werden Modelle, die wie (1.48) linear in den Parametern $\mathbf{p}$ sind, auch als *lineare Regressionsmodelle* bezeichnet. Beinhaltet der Datenvektor $\mathbf{s}_k$ nur alte Werte y_j, $j < k$, der zu berechnenden Größe y_k, dann nennt man dieses Modell *autoregressiv*.

Ein Modell der Form

$$y_k = C(\delta) w_k \tag{1.51}$$

bzw. für $C(\delta) = c_0 + c_1\delta^{-1} + \ldots + c_m\delta^{-m}$ mit $c_0 = 1$

$$y_k = w_k + c_1 w_{k-1} + c_2 w_{k-2} + \ldots + c_m w_{k-m} \tag{1.52}$$

wird als *Modell des gleitenden Mittelwertes* oder *MA (moving average) Modell* bezeichnet.

Aufgabe 1.8. Zeigen Sie, dass sich die Impulsantwortfolge (g_k) der z–Übertragungsfunktion

$$G(z) = c_0 + c_1 z^{-1} + \ldots + c_m z^{-m} \tag{1.53}$$

wie folgt

$$(g_k) = (c_0, c_1, \ldots, c_m, 0, 0, \ldots) \tag{1.54}$$

berechnet. Aus diesem Grund bezeichnet man auch eine Übertragungsfunktion der Form (1.53) als FIR (finite impulse response) Modell. Analog dazu wird eine z–Übertragungsfunktion der Form

$$G(z) = \frac{1}{d_0 + d_1 z^{-1} + \ldots + d_n z^{-n}} \tag{1.55}$$

IIR (infinite impulse response) Modell genannt.

Kombiniert man (1.47) und (1.51), dann erhält man das so genannte *ARMA (autoregressive moving average) Modell*

$$y_k = \frac{C(\delta)}{D(\delta)} w_k \ . \tag{1.56}$$

Für $C(\delta) = c_0 + c_1\delta^{-1} + \ldots + c_m\delta^{-m}$ und $D(\delta) = d_0 + d_1\delta^{-1} + \ldots + d_n\delta^{-n}$, $d_0 = 1$ folgt die Ausgangsgröße y_k zu

$$y_k = c_0 w_k + c_1 w_{k-1} + c_2 w_{k-2} + \ldots + c_m w_{k-m} - d_1 y_{k-1} - d_2 y_{k-2} - \ldots - d_n y_{k-n} \ . \tag{1.57}$$

ARX-Modell

Erweitert man das AR-Modell (1.47) um den Einfluss der deterministischen Eingangsgröße u_k (*exogenous input*), dann erhält man das *ARX (autoregressive with exogenous input) Modell*

$$y_k = \frac{B(\delta)}{A(\delta)} u_k + \frac{1}{A(\delta)} w_k \ . \tag{1.58}$$

Wie man aus (1.58) erkennt, sind in diesem Fall die Nennerpolynome der Übertragungsoperatoren $G(\delta)$ und $H(\delta)$ nach (1.44) identisch. Eine mathematische Begründung für diese Wahl der Struktur wird in weiterer Folge im Rahmen der Least-Squares-Identifikation gegeben. Wie noch gezeigt wird, besteht der Hauptvorteil dieser Modellstruktur in der Tatsache, dass die Parameter linear in den Schätzfehler eingehen und daher sehr einfach über lineare Least-Squares-Methoden geschätzt werden können.

ARMAX-Modell

Analog zum ARX-Modell ergibt sich das *ARMAX (autoregressive moving average with exogenous input) Modell* zu

$$y_k = \frac{B(\delta)}{A(\delta)} u_k + \frac{C(\delta)}{A(\delta)} w_k, \tag{1.59}$$

wobei auch hier wieder die Nennerpolynome der Übertragungsoperatoren vom deterministischen und vom stochastischen Eingang u_k und w_k auf den Ausgang y_k gleich sind. Es ist an dieser Stelle zu erwähnen, dass es in der Literatur noch eine Vielzahl von weiteren Modellen gibt, die sich aber alle auf ähnliche Art und Weise wie bisher besprochen zusammensetzen und Spezialfälle von (1.46) sind. Man beachte dazu nachfolgende Tabelle 1.2.

Modellstruktur	Modellgleichung
MA	$y_k = C(\delta) w_k$
AR	$y_k = \frac{1}{D(\delta)} w_k$
ARMA	$y_k = \frac{C(\delta)}{D(\delta)} w_k$
ARX	$y_k = \frac{B(\delta)}{A(\delta)} u_k + \frac{1}{A(\delta)} w_k$
ARMAX	$y_k = \frac{B(\delta)}{A(\delta)} u_k + \frac{C(\delta)}{A(\delta)} w_k$
ARARX	$y_k = \frac{B(\delta)}{A(\delta)} u_k + \frac{1}{D(\delta)A(\delta)} w_k$
ARARMAX	$y_k = \frac{B(\delta)}{A(\delta)} u_k + \frac{C(\delta)}{D(\delta)A(\delta)} w_k$
OE (output error)	$y_k = \frac{B(\delta)}{F(\delta)} u_k + w_k$
BJ (Box-Jenkins)	$y_k = \frac{B(\delta)}{F(\delta)} u_k + \frac{C(\delta)}{D(\delta)} w_k$

Tabelle 1.2: Modellstrukturen und Modellgleichungen.

1.3.2 Methode der kleinsten Fehlerquadrate (Least Squares)

Gegeben ist das *überbestimmte* lineare Gleichungssystem

$$\mathbf{y} = \mathbf{S}\mathbf{p} \tag{1.60}$$

mit der $(m \times n)$-Matrix $\mathbf{S} \in \mathbb{R}^{m \times n}$, dem m-dimensionalen Vektor $\mathbf{y} \in \mathbb{R}^m$ sowie dem n-dimensionalen Vektor der Unbekannten $\mathbf{p} \in \mathbb{R}^n$. Für $m > n$ und $\operatorname{rang}(\mathbf{S}) = n \neq \operatorname{rang}([\mathbf{S}, \mathbf{y}])$ besitzt das Gleichungssystem (1.60) keine Lösung für $\mathbf{p}$. Es wird nun jene Lösung $\mathbf{p}_0$ gesucht, die den quadratischen Fehler

$$\min_{\mathbf{p}} \|\mathbf{e}\|_2^2 \quad \text{mit} \quad \mathbf{e} = \mathbf{y} - \mathbf{S}\mathbf{p} \tag{1.61}$$

minimiert. Setzt man die Ableitung $\|\mathbf{e}\|_2^2$ bezüglich $\mathbf{p}$ gleich Null

$$\frac{\partial}{\partial \mathbf{p}} \mathbf{e}^{\mathrm{T}} \mathbf{e} = \frac{\partial}{\partial \mathbf{p}} \underbrace{(\mathbf{y} - \mathbf{S}\mathbf{p})^{\mathrm{T}} (\mathbf{y} - \mathbf{S}\mathbf{p})}_{\mathbf{y}^{\mathrm{T}}\mathbf{y} - 2\mathbf{y}^{\mathrm{T}}\mathbf{S}\mathbf{p} + \mathbf{p}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}}\mathbf{S}\mathbf{p}} = -2\mathbf{y}^{\mathrm{T}}\mathbf{S} + 2\mathbf{p}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}}\mathbf{S} = \mathbf{0}^{\mathrm{T}}, \tag{1.62}$$

dann folgt die optimale Lösung $\mathbf{p} = \mathbf{p}_0$ im Sinne von (1.61) zu

$$\mathbf{p}_0 = \left(\mathbf{S}^{\mathrm{T}}\mathbf{S}\right)^{-1}\mathbf{S}^{\mathrm{T}}\mathbf{y}\,. \tag{1.63}$$

Der Ausdruck $\mathbf{S}^{\dagger} = \left(\mathbf{S}^{\mathrm{T}}\mathbf{S}\right)^{-1}\mathbf{S}^{\mathrm{T}}$ wird auch als *Pseudoinverse* der Matrix $\mathbf{S}$ bezeichnet (siehe die MATLAB-Befehle \ und `pinv`). Man erkennt, dass für die Regularität von $\mathbf{S}^{\mathrm{T}}\mathbf{S}$ die Matrix $\mathbf{S}$ spaltenregulär, also $\mathrm{rang}(\mathbf{S}) = n$, sein muss.

Aufgabe 1.9 (Polynomapproximation im Sinne der kleinsten Fehlerquadrate). Gegeben sind N Messpunkte, die den Zusammenhang $y_j = f(x_j)$, $j = 1, \ldots, N$ beschreiben. Setzen Sie für die Funktion $f(x)$ ein Polynom n-ter Ordnung mit $n+1 < N$ der Form

$$f(x) = p_0 + p_1 x + p_2 x^2 + \ldots + p_{n-1}x^{n-1} + p_n x^n$$

an und bestimmen Sie die Polynomkoeffizienten $p_0, p_1, \ldots, p_n$ mit Hilfe der Methode der kleinsten Fehlerquadrate. Testen Sie Ihren Ansatz anhand der Funktion $g(x) = \tanh(x/10)$ für $x_j = -20 + j$, $j = 0, \ldots, 40$, sowie unterschiedliche Ordnungen n des Polynoms.

Aufgabe 1.10. Das mathematische Modell einer permanenterregten Gleichstrommaschine lautet

$$L_a \frac{\mathrm{d}}{\mathrm{d}t} i_a = u_a - R_a i_a - k_a \omega$$
$$J_r \frac{\mathrm{d}}{\mathrm{d}t} \omega = k_a i_a - d_v \omega - d_c \mathrm{sign}(\omega)$$

mit dem Ankerstrom i_a, der Drehwinkelgeschwindigkeit ω, der Ankerspannung u_a, der Ankerinduktivität L_a, dem Ankerwiderstand R_a, der Ankerkreiskonstanten k_a, dem Trägheitsmoment J_r und den Reibparametern d_v und d_c. Bestimmen Sie aus stationären Messungen von u_a, i_a und ω die Parameter k_a, R_a, d_v und d_c mit Hilfe der Methode der kleinsten Fehlerquadrate. Testen Sie ihren Algorithmus anhand der Messdaten `data.mat` und `data_rausch.mat`, in denen die Messwerte von u_a, i_a und ω mit und ohne Messrauschen abgelegt sind. Vergleichen Sie ihre Identifikationsergebnisse mit den nominellen Parametern $R_a = 1.373\,\Omega$, $k_a = 0.0652\,\mathrm{V\,s/rad}$, $d_c = 0.0188\,\mathrm{N\,m}$ und $d_v = 43.3 \cdot 10^{-6}\,\mathrm{N\,m\,s/rad}$.

Daten zur Aufgabe 1.10:
`https://www.acin.tuwien.ac.at/file/teaching/master/Regelungssysteme-1/Daten_Aufgabe_1_10.zip`.

Die optimale Lösung $\mathbf{p}_0$ nach (1.63) der Optimierungsaufgabe (1.61) lässt folgende geometrische Deutung zu: Die n linear unabhängigen Spaltenvektoren der Matrix $\mathbf{S}$ spannen im $\mathbb{R}^m$ einen n-dimensionalen Unterraum $\mathcal{U}$ auf. Die Lösbarkeit des Gleichungssystems (1.60) ist äquivalent zur Frage, ob eine Linearkombination (beschrieben durch die Einträge im Vektor $\mathbf{p}$) der Spaltenvektoren von $\mathbf{S}$ so exisitiert, dass damit der Vektor $\mathbf{y}$ dargestellt werden kann. Damit ist das Gleichungssystem (1.60) eindeutig lösbar, wenn $\mathbf{y}$ in diesem Unterraum $\mathcal{U}$ liegt, also $\mathrm{rang}(\mathbf{S}) = \mathrm{rang}([\mathbf{S}, \mathbf{y}])$ gilt. Liegt $\mathbf{y}$ nicht im Unterraum $\mathcal{U}$, dann

entsteht zwischen der (mit Hilfe der im Sinne der kleinsten Quadrate geschätzten optimalen Lösung $\mathbf{p}_0$ berechneten) Größe $\mathbf{y}_0 = \mathbf{S}\mathbf{p}_0 = \mathbf{S}\left(\mathbf{S}^{\mathrm{T}}\mathbf{S}\right)^{-1}\mathbf{S}^{\mathrm{T}}\mathbf{y}$ und der tatsächlichen Größe $\mathbf{y}$ der Fehler

$$\mathbf{e}_0 = \mathbf{y} - \mathbf{y}_0 = \mathbf{y} - \mathbf{S}\left(\mathbf{S}^{\mathrm{T}}\mathbf{S}\right)^{-1}\mathbf{S}^{\mathrm{T}}\mathbf{y} \ . \tag{1.64}$$

Man erkennt nun, dass dieser Fehler $\mathbf{e}_0$ orthogonal auf den Unterraum $\mathcal{U}$ steht, da gilt

$$\mathbf{S}^{\mathrm{T}}\mathbf{e}_0 = \mathbf{S}^{\mathrm{T}}\mathbf{y} - \mathbf{S}^{\mathrm{T}}\mathbf{S}\left(\mathbf{S}^{\mathrm{T}}\mathbf{S}\right)^{-1}\mathbf{S}^{\mathrm{T}}\mathbf{y} = \mathbf{0} \ . \tag{1.65}$$

Abbildung 1.7 veranschaulicht diesen Sachverhalt geometrisch für $m = 3$ und $n = 2$.

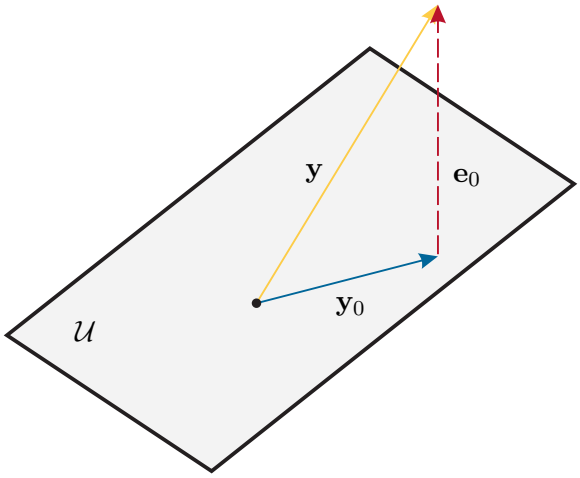


Abbildung 1.7: Zur Methode der kleinsten Fehlerquadrate.

Die Erkenntnis, dass der Fehler $\mathbf{e}_0$ orthogonal auf den Unterraum $\mathcal{U}$ steht, ist ein Spezialfall des so genannten *Projektionstheorems in einem Hilbertraum*.

Projektionstheorem in einem Hilbertraum

Zur Erläuterung des Projektionstheorems werden im Folgenden kurz die Begriffe Vektorraum, normierter Vektorraum und Hilbertraum wiederholt.

Definition 1.1 (Linearer Vektorraum)**.** Man nennt eine nichtleere Menge $\mathcal{X}$ einen linearen Vektorraum über einem (skalaren) Körper K mit den binären Operationen $+ : \mathcal{X} \times \mathcal{X} \to \mathcal{X}$ (Addition) und $\cdot : K \times \mathcal{X} \to \mathcal{X}$ (Multiplikation mit einem Skalar aus K), wenn folgende Vektorraumaxiome erfüllt sind:

(1) Die Menge $\mathcal{X}$ mit der Verknüpfung $+$ ist eine kommutative Gruppe, d. h. für $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathcal{X}$ gilt:

(1)	$\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$	Kommutativität
(2)	$\mathbf{x} + (\mathbf{y} + \mathbf{z}) = (\mathbf{x} + \mathbf{y}) + \mathbf{z}$	Assoziativität
(3)	$\mathbf{0} + \mathbf{x} = \mathbf{x}$	neutrales Element
(4)	$\mathbf{x} + (-\mathbf{x}) = \mathbf{0}$	inverses Element

(2) Die Multiplikation $\cdot$ mit einem Skalar a, $b \in K$ genügt den Gesetzen:

(1)	$a(\mathbf{x} + \mathbf{y}) = a\mathbf{x} + a\mathbf{y}$	Distributivität
(2)	$(a + b)\mathbf{x} = a\mathbf{x} + b\mathbf{x}$	Distributivität
(3)	$(ab)\mathbf{x} = a(b\mathbf{x})$	Assoziativität
(4)	$1\mathbf{x} = \mathbf{x} \quad , \quad 0\mathbf{x} = \mathbf{0}$	

Definition 1.2 (Normierter linearer Vektorraum)**.** Ein normierter linearer Vektorraum ist ein Vektorraum $\mathcal{X}$ über einem Skalarkörper K mit einer reellwertigen Funktion $\|\mathbf{x}\| : \mathcal{X} \to \mathbb{R}_+$, die jedem $\mathbf{x} \in \mathcal{X}$ eine reellwertige Zahl $\|\mathbf{x}\|$, die so genannte *Norm* von $\mathbf{x}$, zuordnet und folgende Normaxiome erfüllt:

(1)	$\|\mathbf{x}\| \geq 0$ für alle $\mathbf{x} \in \mathcal{X}$	Nichtnegativität
(2)	$\|\mathbf{x}\| = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$	
(3)	$\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$	Dreiecksungleichung
(4)	$\|\alpha\mathbf{x}\| = \|\alpha\|\|\mathbf{x}\|$ für alle $\mathbf{x} \in \mathcal{X}$ und alle $\alpha \in K$	

Definition 1.3 (Prä-Hilbertraum). Es sei $\mathcal{X}$ ein linearer Vektorraum mit dem Skalarkörper K. Eine Abbildung $\langle \mathbf{x}, \mathbf{y} \rangle : \mathcal{X} \times \mathcal{X} \to K$, die je zwei Elementen $\mathbf{x}, \mathbf{y} \in \mathcal{X}$ einen Skalar zuordnet, heißt *inneres Produkt*, wenn sie folgenden Bedingungen

(1) $\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle$ Bilinearität
(2) $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$
(3) $\langle a\mathbf{x}, \mathbf{y} \rangle = a \langle \mathbf{x}, \mathbf{y} \rangle$
(4) $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0 \quad \text{und} \quad \langle \mathbf{x}, \mathbf{x} \rangle = 0 \Leftrightarrow \mathbf{x} = 0$

mit $a \in K$, genügt.

Einen vollständigen Prä-Hilbertraum bezeichnet man als Hilbertraum, vgl. z. B. [1.2]. Man bezeichnet nun zwei Vektoren $\mathbf{x}$ und $\mathbf{y}$ aus einem Hilbertraum $\mathcal{H}$ als *orthogonal*, wenn $\langle \mathbf{x}, \mathbf{y} \rangle = 0$ gilt. Eine nichtleere Menge $\mathcal{U}$ wird *Unterraum* von $\mathcal{H}$ genannt, wenn für alle Linearkombinationen von Vektoren $\mathbf{x}$ und $\mathbf{y}$ aus $\mathcal{U}$ gilt $a\mathbf{x} + b\mathbf{y} \in \mathcal{U}$, mit den Skalaren a und b. Im Weiteren nennt man den Unterraum $\mathcal{U}$ *abgeschlossen*, wenn der Grenzwert jeder konvergenten Folge in $\mathcal{U}$ ebenfalls in $\mathcal{U}$ liegt.

Aufgabe 1.11. Zeigen Sie, dass durch $\|\mathbf{x}\|_2 = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$ eine Norm nach Definition 1.2 gegeben ist.

Das folgende Projektionstheorem gibt nun eine notwendige und hinreichende Bedingung zur Lösung nachfolgender Aufgabe an: Gegeben ist ein Vektor $\mathbf{y}$ in einem Hilbertraum $\mathcal{H}$. Es wird jener Vektor $\mathbf{y}_0$ aus einem Unterraum $\mathcal{U}$ von $\mathcal{H}$ gesucht, der die Norm $\|\mathbf{y} - \mathbf{y}_0\|$ minimiert.

Satz 1.1 (Projektionstheorem). *Es sei $\mathcal{H}$ ein Hilbertraum und $\mathcal{U}$ ein abgeschlossener Unterraum von $\mathcal{H}$. Zu jedem Vektor $\mathbf{y} \in \mathcal{H}$ existiert ein eindeutiger Vektor $\mathbf{y}_0 \in \mathcal{U}$ so, dass gilt*

$$\|\mathbf{y} - \mathbf{y}_0\| \leq \|\mathbf{y} - \mathbf{x}\| \tag{1.66}$$

für alle $\mathbf{x} \in \mathcal{U}$. Der Vektor $\mathbf{y}_0$ ist genau dann der eindeutige, minimierende Vektor, wenn gilt

$$\langle \mathbf{y} - \mathbf{y}_0, \mathbf{x} \rangle = 0 \tag{1.67}$$

für alle $\mathbf{x} \in \mathcal{U}$. Dies bedeutet, dass der Fehler $\mathbf{y} - \mathbf{y}_0$ orthogonal auf alle Vektoren des Unterraums $\mathcal{U}$ sein muss.

Der Beweis dieses Satzes kann in der am Ende dieses Kapitels angegebenen Literatur nachgelesen werden.

Im Folgenden soll das Projektionstheorem zur Lösung der Optimierungsaufgabe (1.61) herangezogen werden. Gegeben ist der Hilbertraum $\mathcal{H} = \mathbb{R}^m$ mit dem inneren Produkt

$$\langle \mathbf{x}, \mathbf{z} \rangle = \mathbf{x}^{\mathrm{T}} \mathbf{z} = \sum_{j=1}^{m} x_j z_j \tag{1.68}$$

mit

$$\mathbf{x}^{\mathrm{T}} = \begin{bmatrix} x_1 & x_2 & \dots & x_m \end{bmatrix} \tag{1.69a}$$

$$\mathbf{z}^{\mathrm{T}} = \begin{bmatrix} z_1 & z_2 & \dots & z_m \end{bmatrix} \tag{1.69b}$$

und dem abgeschlossenen Unterraum $\mathcal{U}$ von $\mathcal{H}$, welcher durch die Spaltenvektoren $\mathbf{s}_k$, $k = 1, \ldots, n$ der Matrix $\mathbf{S}$ aufgespannt wird, also $\mathcal{U} = \text{span}\{\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_n\}$. Im Sinne von (1.61) ist also jenes $\mathbf{S}\mathbf{p}_0 = \mathbf{y}_0 \in \mathcal{U}$ gesucht, das zu einem gegebenen $\mathbf{y} \in \mathcal{H}$ die Quadratnorm des Fehlers $\|\mathbf{e}\|_2^2 = \|\mathbf{y} - \mathbf{y}_0\|_2^2$ minimiert. Nach Satz 1.1 muss also gelten

$$\langle \mathbf{y} - \mathbf{S}\mathbf{p}_0, \mathbf{s}_j \rangle = \langle \mathbf{y} - \mathbf{s}_1 p_{0,1} - \mathbf{s}_2 p_{0,2} - \ldots - \mathbf{s}_n p_{0,n}, \mathbf{s}_j \rangle = 0 \quad \text{für alle } j = 1, \ldots, n \; . \tag{1.70}$$

In Matrixschreibweise unter Verwendung der Eigenschaften des inneren Produktes erhält man

$$\underbrace{\begin{bmatrix} \langle \mathbf{s}_1, \mathbf{s}_1 \rangle & \ldots & \langle \mathbf{s}_n, \mathbf{s}_1 \rangle \\ \vdots & \ddots & \vdots \\ \langle \mathbf{s}_1, \mathbf{s}_n \rangle & \ldots & \langle \mathbf{s}_n, \mathbf{s}_n \rangle \end{bmatrix}}_{\mathbf{G} = \mathbf{S}^{\mathrm{T}}\mathbf{S}} \underbrace{\begin{bmatrix} p_{0,1} \\ \vdots \\ p_{0,n} \end{bmatrix}}_{\mathbf{p}_0} = \underbrace{\begin{bmatrix} \langle \mathbf{y}, \mathbf{s}_1 \rangle \\ \vdots \\ \langle \mathbf{y}, \mathbf{s}_n \rangle \end{bmatrix}}_{\mathbf{S}^{\mathrm{T}}\mathbf{y}} \tag{1.71}$$

und damit unmittelbar die Lösung für $\mathbf{p}_0$ von (1.63). Die Matrix $\mathbf{G}$ von (1.71) wird auch als *Gramsche Matrix* bezeichnet.

Aufgabe 1.12. Zeigen Sie, dass die Gramsche Matrix von (1.71) genau dann regulär ist, wenn die Vektoren $\mathbf{s}_k$, $k = 1, \ldots, n$, linear unabhängig sind.

Aufgabe 1.13. Berechnen Sie die Lösung des quadratischen Minimierungsproblems von Abschnitt 1.2.2 mit Hilfe von Satz 1.1.

Hinweis: (zu Aufgabe 1.13) Als Hilbertraum $\mathcal{H}$ wird die Menge aller N-Tupel $\xi = \{\xi_0, \xi_2, \ldots, \xi_{N-1}\}$ mit dem inneren Produkt

$$\langle x, z \rangle = \frac{1}{N} \sum_{k=0}^{N-1} x_j z_j \quad \text{mit} \quad x, z \in \mathcal{H}$$

verwendet.

Die Menge aller N–Tupel $\{C\cos(\omega_0 k T_a) + S\sin(\omega_0 k T_a)\}$, $k = 0, \ldots, N-1$ mit den beliebigen aber konstanten Koeffizienten C und S sowie der festen Kreisfrequenz ω_0 bildet einen abgeschlossenen Unterraum $\mathcal{U}$ von $\mathcal{H}$. Die quadratische Minimierungsaufgabe von Abschnitt 1.2.2 kann nun so formuliert werden, dass man jenes $\hat{y} \in \mathcal{U}$ sucht, welches die Norm (man vergleiche dazu (1.13))

$$\|y - \hat{y}\|^2 = \frac{1}{N} \sum_{k=0}^{N-1} (y_k - \hat{y}_k)^2$$

mit

$$\hat{y}_k = \hat{Y}_c \cos(\omega_0 k T_a) + \hat{Y}_s \sin(\omega_0 k T_a)$$

für ein vorgegebenes $y \in \mathcal{H}$ minimiert. Dies ist aber nach Satz 1.1 genau dann der Fall, wenn gilt

$$\langle y - \hat{y}, x \rangle = 0 \quad \text{für alle } x \in \mathcal{U}$$

bzw. für die beiden Spezialfälle $x = \cos(\omega_0 k T_a)$ und $x = \sin(\omega_0 k T_a)$ erhält man die Bedingungen

$$\frac{1}{N}\sum_{k=0}^{N-1}\Big(y_k - \hat{Y}_c\cos(\omega_0 k T_a) - \hat{Y}_s\sin(\omega_0 k T_a)\Big)\cos(\omega_0 k T_a) = 0$$
$$\frac{1}{N}\sum_{k=0}^{N-1}\Big(y_k - \hat{Y}_c\cos(\omega_0 k T_a) - \hat{Y}_s\sin(\omega_0 k T_a)\Big)\sin(\omega_0 k T_a) = 0\ .$$

1.3.3 Least-Squares Identifikation

Ausgangspunkt der Betrachtungen ist das ARX-Modell (1.58) ohne stochastische Störung, also $(w_k) = (0)$, der Form

$$y_k = \frac{B(\delta)}{A(\delta)}u_k = \delta^{-d}\frac{b_0 + b_1\delta^{-1} + \ldots + b_m\delta^{-m}}{a_0 + a_1\delta^{-1} + \ldots + a_n\delta^{-n}}u_k \quad \text{mit} \quad d \geq 0,\ a_0 = 1\ . \tag{1.72}$$

Für eine Übertragungsfunktion $G(z)$ mit Zählergrad m und Nennergrad n errechnet sich d zu $d = n - m$. Man beachte, dass mit der Formulierung (1.72) auch Totzeiten der Form $z^{-\rho}$, $\rho > 0$, mitberücksichtigt werden können. Der Wert der Ausgangsfolge (y_k) zum k-ten Abtastzeitpunkt errechnet sich aus (1.72) zu

$$y_k = -a_1 y_{k-1} - \ldots - a_n y_{k-n} + b_0 u_{k-d} + b_1 u_{k-d-1} + \ldots + b_m u_{k-d-m} \tag{1.73}$$

bzw. in Vektorschreibweise erhält man

$$y_k = \underbrace{\begin{bmatrix} -y_{k-1} & -y_{k-2} & \ldots & -y_{k-n} & u_{k-d} & u_{k-d-1} & \ldots & u_{k-d-m} \end{bmatrix}}_{\mathbf{s}_k^{\mathrm{T}}} \underbrace{\begin{bmatrix} a_1 \\ \vdots \\ a_n \\ b_0 \\ \vdots \\ b_m \end{bmatrix}}_{\mathbf{p}} \tag{1.74}$$

mit dem *Datenvektor* $\mathbf{s}_k$ und dem *Parametervektor* $\mathbf{p}$. Im Rahmen der Identifikationsaufgabe wird nun der Parametervektor $\mathbf{p}$ so geschätzt, dass ein noch zu definierender Modellfehler minimiert wird.

Wählt man als Modellfehler den so genannten *verallgemeinerten Gleichungsfehler* nach Abbildung 1.8, dann erhält man die Beziehung

$$e_k = \hat{A}(\delta)y_k - \hat{B}(\delta)u_k \tag{1.75}$$

mit

$$\hat{A}(\delta) = 1 + \hat{a}_1\delta^{-1} + \ldots + \hat{a}_n\delta^{-n} \tag{1.76a}$$
$$\hat{B}(\delta) = \hat{b}_0\delta^{-d} + \hat{b}_1\delta^{-1-d} + \ldots + \hat{b}_m\delta^{-m-d} \tag{1.76b}$$

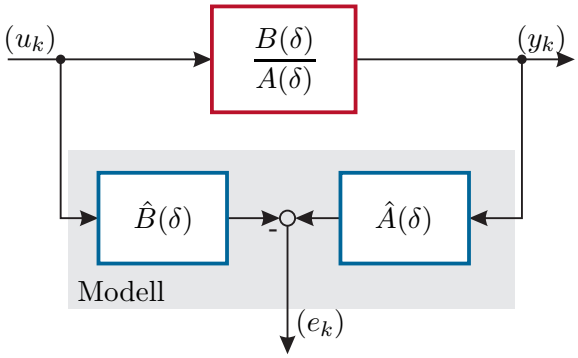


Abbildung 1.8: Zum verallgemeinerten Gleichungsfehler.

und den geschätzten Koeffizienten des Zähler- und Nennerpolynoms $\hat{b}_j$, $j = 0, \ldots, m$ sowie $\hat{a}_k$, $k = 1, \ldots, n$. Setzt man (1.76) in (1.75) ein und fasst die zu schätzenden Koeffizienten im *Parameterschätzvektor* $\hat{\mathbf{p}} = \begin{bmatrix} \hat{a}_1 & \ldots & \hat{a}_n & \hat{b}_0 & \ldots & \hat{b}_m \end{bmatrix}^{\mathrm{T}}$ zusammen, dann erhält man

$$e_k = y_k - \mathbf{s}_k^{\mathrm{T}} \hat{\mathbf{p}} \tag{1.77}$$

mit dem *Datenvektor* $\mathbf{s}_k$ gemäß (1.74). Aus (1.77) erkennt man unmittelbar, dass sich der Schätzwert $\hat{y}_k$ von y_k in der Form $\hat{y}_k = \mathbf{s}_k^{\mathrm{T}} \hat{\mathbf{p}}$ errechnet. Durch Kombination von $j = 0, \ldots, N$ Messungen lässt sich (1.77) wie folgt

$$\underbrace{\begin{bmatrix} e_0 \\ \vdots \\ e_N \end{bmatrix}}_{\mathbf{e}_N} = \underbrace{\begin{bmatrix} y_0 \\ \vdots \\ y_N \end{bmatrix}}_{\mathbf{y}_N} - \underbrace{\begin{bmatrix} \mathbf{s}_0^{\mathrm{T}} \\ \vdots \\ \mathbf{s}_N^{\mathrm{T}} \end{bmatrix}}_{\mathbf{S}_N} \hat{\mathbf{p}}_N \tag{1.78}$$

erweitern, wobei der Index N des Parameterschätzvektors $\hat{\mathbf{p}}$ andeutet, dass $N + 1$ Messungen zur Schätzung von $\mathbf{p}$ herangezogen werden. Für $N > n + m$ mit m und n gemäß (1.72) und $\operatorname{rang}(\mathbf{S}_N) \neq \operatorname{rang}([\mathbf{S}_N, \mathbf{y}_N])$ besitzt das Gleichungssystem (1.78) für $\mathbf{e}_N = \mathbf{0}$ keine Lösung. Den Überlegungen von Abschnitt 1.3.2 folgend lautet die Lösung von (1.78) im Sinne der kleinsten Fehlerquadrate (siehe (1.63))

$$\hat{\mathbf{p}}_N = \left(\mathbf{S}_N^{\mathrm{T}} \mathbf{S}_N \right)^{-1} \mathbf{S}_N^{\mathrm{T}} \mathbf{y}_N \ . \tag{1.79}$$

Hinweis: Zur Wahl der Notation sei angemerkt, dass mit $j = 0$ nicht unbedingt der Startzeitpunkt der Messung $k = 0$ gemeint ist, sondern jener Zeitpunkt k, bei dem der Eintrag u_{k-d} im Datenvektor $\mathbf{s}_k^{\mathrm{T}}$ erstmals von Null verschieden ist.

Es sei an dieser Stelle nochmals betont, dass erst die spezielle Wahl des verallgemeinerten Gleichungsfehlers nach Abbildung 1.8 ein parametrisch lineares Schätzproblem mit sich bringt.

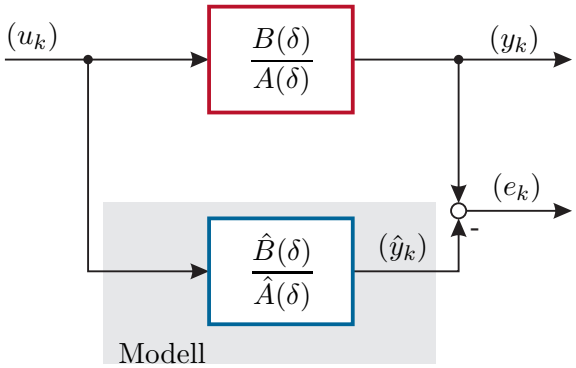


Abbildung 1.9: Zum Ausgangsfehler.

Würde man beispielsweise, wie in Abbildung 1.9 gezeigt, den so genannten *Ausgangsfehler*

$$e_k = y_k - \hat{y}_k = y_k - \frac{\hat{B}(\delta)}{\hat{A}(\delta)} u_k = y_k - \delta^{-d} \frac{\hat{b}_0 + \hat{b}_1 \delta^{-1} + \ldots + \hat{b}_m \delta^{-m}}{1 + \hat{a}_1 \delta^{-1} + \ldots + \hat{a}_n \delta^{-n}} u_k \tag{1.80}$$

heranziehen, dann sieht man aus

$$\hat{y}_k = -\hat{a}_1 \hat{y}_{k-1} - \ldots - \hat{a}_n \hat{y}_{k-n} + \hat{b}_0 u_{k-d} + \hat{b}_1 u_{k-d-1} + \ldots + \hat{b}_m u_{k-d-m} \tag{1.81}$$

bereits für $k = 0, \ldots, d+1$

$$\begin{aligned} \hat{y}_k &= 0 \quad \text{für } k = 0, 1, \ldots, d-1 \\ \hat{y}_d &= \hat{b}_0 u_0 \\ \hat{y}_{d+1} &= -\hat{a}_1 \hat{y}_d + \hat{b}_0 u_1 + \hat{b}_1 u_0 = \left(\hat{b}_1 - \hat{a}_1 \hat{b}_0\right) u_0 + \hat{b}_0 u_1, \end{aligned} \tag{1.82}$$

dass es sich in diesem Fall um ein parametrisch nichtlineares Schätzproblem handelt, welches um Größenordnungen schwieriger zu lösen ist.

Bevor angegeben wird, wie sich eine stochastische Störung (v_k) auf das Ergebnis der Parameterschätzung $\hat{\mathbf{p}}$ von (1.79) auswirkt, werden zwei grundsätzliche Eigenschaften von Parameterschätzverfahren definiert.

Definition 1.4 (Erwartungstreue Schätzung)**.** Wenn eine Schätzung für eine beliebige Anzahl N von Messungen einen systematischen Fehler

$$\mathrm{E}(\hat{\mathbf{p}}_N - \mathbf{p}) = \mathrm{E}(\hat{\mathbf{p}}_N) - \mathbf{p} = \mathbf{b} \neq \mathbf{0} \tag{1.83}$$

liefert, dann nennt man diesen Fehler *Bias*. Für eine *biasfreie* oder *erwartungstreue* Schätzung gilt daher

$$\mathrm{E}(\hat{\mathbf{p}}_N) = \mathbf{p}\ . \tag{1.84}$$

Definition 1.5 (Konsistente Schätzung)**.** Eine Schätzung wird als *konsistent* bezeichnet, wenn der Schätzwert umso genauer wird, je größer die Anzahl N der Messwerte ist, d. h. wenn gilt

$$\lim_{N\to\infty} \mathrm{E}(\hat{\mathbf{p}}_N) = \mathbf{p} \; . \tag{1.85}$$

Eine Schätzung wird als *konsistent im quadratischen Mittel* bezeichnet, wenn zusätzlich zu (1.85) die Bedingung

$$\lim_{N\to\infty} \mathrm{cov}(\hat{\mathbf{p}}_N) = \lim_{N\to\infty} \mathrm{E}\Big([\hat{\mathbf{p}}_N - \mathbf{p}][\hat{\mathbf{p}}_N - \mathbf{p}]^{\mathrm{T}}\Big) = \mathbf{0} \tag{1.86}$$

erfüllt ist.

Hinweis: Man beachte, dass eine konsistente Schätzung nichts über die Güte der Schätzung bei endlichem N aussagt. Es kann also sein, dass eine konsistente Schätzung sehr wohl bei endlichem N biasbehaftet ist.

Zur Analyse des Einflusses einer stochastischen Störung wird angenommen, dass sich die gemessene Ausgangsgröße y_k aus dem ungestörten Ausgang $\bar{y}_k$ und einer stochastischen Störung v_k in der Form $y_k = \bar{y}_k + v_k$ zusammensetzt, siehe Abbildung 1.3. Der ungestörte Ausgang errechnet sich nach (1.74) zu

$$\bar{y}_k = \underbrace{\begin{bmatrix} -\bar{y}_{k-1} & -\bar{y}_{k-2} & \dots & -\bar{y}_{k-n} & u_{k-d} & u_{k-d-1} & \dots & u_{k-d-m} \end{bmatrix}}_{\bar{\mathbf{s}}_k^{\mathrm{T}}} \underbrace{\begin{bmatrix} a_1 \\ \vdots \\ a_n \\ b_0 \\ \vdots \\ b_m \end{bmatrix}}_{\mathbf{p}} \; . \tag{1.87}$$

Ersetzt man die Einträge $\bar{y}_j$ in (1.87) durch $\bar{y}_j = y_j - v_j$, dann erhält man

$$\bar{y}_k = \mathbf{s}_k^{\mathrm{T}}\mathbf{p} + \mathbf{n}_k^{\mathrm{T}}\mathbf{p} \tag{1.88}$$

mit

$$\mathbf{s}_k^{\mathrm{T}} = \begin{bmatrix} -y_{k-1} & -y_{k-2} & \dots & -y_{k-n} & u_{k-d} & u_{k-d-1} & \dots & u_{k-d-m} \end{bmatrix} \tag{1.89a}$$

$$\mathbf{n}_k^{\mathrm{T}} = \begin{bmatrix} v_{k-1} & v_{k-2} & \dots & v_{k-n} & 0 & 0 & \dots & 0 \end{bmatrix} \; . \tag{1.89b}$$

Durch Kombination von $k = 0, \dots, N$ Messungen und unter Verwendung von $y_k = \bar{y}_k + v_k$ kann die folgende Formulierung gefunden werden

$$\underbrace{\begin{bmatrix} y_0 \\ \vdots \\ y_N \end{bmatrix}}_{\mathbf{y}_N} = \underbrace{\begin{bmatrix} \mathbf{s}_0^{\mathrm{T}} \\ \vdots \\ \mathbf{s}_N^{\mathrm{T}} \end{bmatrix}}_{\mathbf{S}_N} \mathbf{p} + \underbrace{\begin{bmatrix} \mathbf{n}_0^{\mathrm{T}} \\ \vdots \\ \mathbf{n}_N^{\mathrm{T}} \end{bmatrix}}_{\mathbf{N}_N} \mathbf{p} + \underbrace{\begin{bmatrix} v_0 \\ \vdots \\ v_N \end{bmatrix}}_{\mathbf{v}_N} \; . \tag{1.90}$$

Führt man nun eine Least-Squares Identifikation basierend auf dem bekannten Teil $\mathbf{y}_N = \mathbf{S}_N\mathbf{p}$ durch, so erhält man

$$\hat{\mathbf{p}} = \left(\mathbf{S}_N^{\mathrm{T}}\mathbf{S}_N\right)^{-1}\mathbf{S}_N^{\mathrm{T}}\mathbf{y} = \left(\mathbf{S}_N^{\mathrm{T}}\mathbf{S}_N\right)^{-1}\mathbf{S}_N^{\mathrm{T}}(\mathbf{S}_N\mathbf{p} + \mathbf{N}_N\mathbf{p} + \mathbf{v}_N) \tag{1.91}$$

bzw. nach kurzer Rechnung

$$\hat{\mathbf{p}} = \mathbf{p} + \left(\mathbf{S}_N^{\mathrm{T}}\mathbf{S}_N\right)^{-1}\mathbf{S}_N^{\mathrm{T}}(\mathbf{N}_N\mathbf{p} + \mathbf{v}_N)\ . \tag{1.92}$$

Der Erwartungswert des Schätzfehlers ergibt sich damit zu

$$\mathrm{E}(\hat{\mathbf{p}}) - \mathbf{p} = \mathrm{E}\left(\left(\mathbf{S}_N^{\mathrm{T}}\mathbf{S}_N\right)^{-1}\mathbf{S}_N^{\mathrm{T}}(\mathbf{N}_N\mathbf{p} + \mathbf{v}_N)\right) = \mathbf{b}, \tag{1.93}$$

mit dem Bias $\mathbf{b}$. Man erkennt, dass die Schätzung $\hat{\mathbf{p}}$ von $\mathbf{p}$ genau dann erwartungstreu ist, wenn

$$\mathbf{b} = \mathbf{0} \tag{1.94}$$

gilt. Weiterhin ist die Schätzung offensichtlich konsistent, wenn

$$\lim_{N\to\infty} \mathbf{b} = \mathbf{0} \tag{1.95}$$

erfüllt ist. Der nachfolgende Satz gibt nun Auskunft, welche Anforderungen an die stochastische Störung v_k bzw. die Modellstruktur gestellt werden müssen, damit diese beiden Bedingungen erfüllt sind.

Satz 1.2 (Erwartungstreue und konsistente Least-Squares Identifikation)**.** *Die Least-Squares Identifikation des Parameters* $\mathbf{p}$ *für die Identifikationsaufgabe nach* (1.90) *bzw. nach Abbildung 1.3 ist erwartungstreu und konsistent, falls die stochastische Störung* v_k *die* Yule-Walker Gleichung *eines autoregressiven Signalprozesses der Form*

$$v_k + a_1 v_{k-1} + a_2 v_{k-2} + \ldots + a_n v_{k-n} = w_k, \tag{1.96}$$

mit dem mittelwertfreien weißen Rauschen w_k *und den Koeffizienten* a_j, $j = 1, \ldots, n$, *des Nenners der zu identifizierenden Übertragungsfunktion, erfüllt.*

Das heißt, das Störsignal v_k muss durch Filterung von weißem Rauschen w_k über ein Filter mit der Übertragungsfunktion $1/A(\delta)$ erzeugt worden sein. Diese Struktur entspricht gerade der im Abschnitt 1.3.1 definierten Modellstruktur eines ARX-Modells, siehe Abbildung 1.10. Man kann zusätzlich zeigen, dass in diesem Fall die Least-Squares Identifikation auch konsistent im quadratischen Mittel ist. Für einen Beweis von Satz 1.2 wird auf die Literatur, insbesondere auf [1.3], verwiesen.

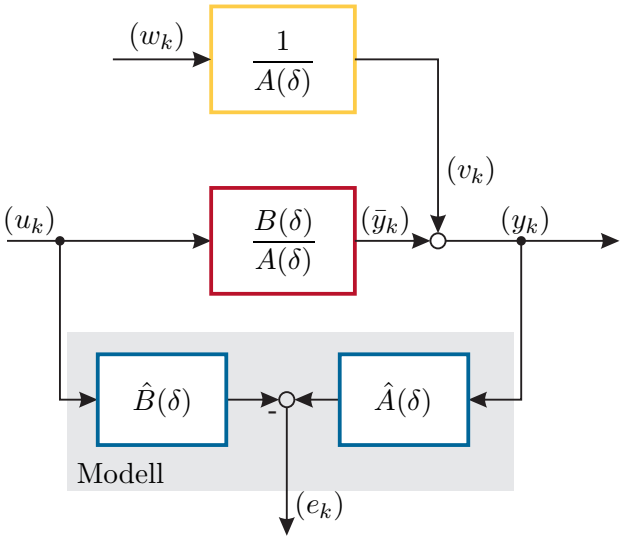


Abbildung 1.10: Zur Least-Squares Identifikation mit stochastischer Störung.

1.3.4 Rekursive Least-Squares (RLS) Identifikation

Die Parameterschätzung (1.79) eignet sich aufgrund der immer weiter anwachsenden Dimensionen des Messvektors $\mathbf{y}_N$ bzw. der Matrix $\mathbf{S}_N$ nicht für den on-line Betrieb. Im Folgenden wird auf Basis von (1.79) eine *rekursive Methode* angegeben, die mit jeder neuen Messung den Schätzwert $\hat{\mathbf{p}}$ des Parametervektors $\mathbf{p}$ verbessert. Gemäß (1.79) lautet der optimale Schätzwert für $N+1$ Messungen

$$\hat{\mathbf{p}}_N = \left(\mathbf{S}_N^{\mathrm{T}}\mathbf{S}_N\right)^{-1}\mathbf{S}_N^{\mathrm{T}}\mathbf{y}_N \tag{1.97}$$

bzw. für $N+2$ Messungen

$$\hat{\mathbf{p}}_{N+1} = \left(\mathbf{S}_{N+1}^{\mathrm{T}}\mathbf{S}_{N+1}\right)^{-1}\mathbf{S}_{N+1}^{\mathrm{T}}\mathbf{y}_{N+1}\ . \tag{1.98}$$

Partitioniert man die Datenmatrix $\mathbf{S}_{N+1}$ und den Messvektor $\mathbf{y}_{N+1}$ in der Form

$$\mathbf{S}_{N+1} = \begin{bmatrix}\mathbf{S}_N \\ \mathbf{s}_{N+1}^{\mathrm{T}}\end{bmatrix} \quad \text{und} \quad \mathbf{y}_{N+1} = \begin{bmatrix}\mathbf{y}_N \\ y_{N+1}\end{bmatrix}, \tag{1.99}$$

dann ergibt sich $\hat{\mathbf{p}}_{N+1}$ zu

$$\begin{aligned}\hat{\mathbf{p}}_{N+1} &= \left(\begin{bmatrix}\mathbf{S}_N^{\mathrm{T}} & \mathbf{s}_{N+1}\end{bmatrix}\begin{bmatrix}\mathbf{S}_N \\ \mathbf{s}_{N+1}^{\mathrm{T}}\end{bmatrix}\right)^{-1}\left(\begin{bmatrix}\mathbf{S}_N^{\mathrm{T}} & \mathbf{s}_{N+1}\end{bmatrix}\begin{bmatrix}\mathbf{y}_N \\ y_{N+1}\end{bmatrix}\right) \\ &= \left(\mathbf{S}_N^{\mathrm{T}}\mathbf{S}_N + \mathbf{s}_{N+1}\mathbf{s}_{N+1}^{\mathrm{T}}\right)^{-1}\left(\mathbf{S}_N^{\mathrm{T}}\mathbf{y}_N + \mathbf{s}_{N+1}y_{N+1}\right)\ .\end{aligned} \tag{1.100}$$

Aufgabe 1.14. Zeigen Sie die Gültigkeit von (1.100).

Zur weiteren Berechnung benötigt man nachfolgenden Hilfssatz zur *Matrizeninversion*:

Satz 1.3 (Zur Matrizeninversion). *Wenn* $\mathbf{A}$, $\mathbf{C}$ *und* $(\mathbf{A}+\mathbf{BCD})$ *reguläre quadratische Matrizen sind, dann gilt*

$$(\mathbf{A}+\mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}\Big(\mathbf{C}^{-1}+\mathbf{DA}^{-1}\mathbf{B}\Big)^{-1}\mathbf{DA}^{-1} \ . \tag{1.101}$$

Durch Anwendung des Satzes 1.3 auf $\mathbf{F} = \Big(\mathbf{S}_N^{\mathrm{T}}\mathbf{S}_N + \mathbf{s}_{N+1}\mathbf{s}_{N+1}^{\mathrm{T}}\Big)$ mit $\mathbf{A} = \mathbf{S}_N^{\mathrm{T}}\mathbf{S}_N$, $\mathbf{B} = \mathbf{s}_{N+1}$, $\mathbf{C} = 1$ und $\mathbf{D} = \mathbf{s}_{N+1}^{\mathrm{T}}$ erhält man

$$\mathbf{F}^{-1} = \Big(\mathbf{S}_N^{\mathrm{T}}\mathbf{S}_N\Big)^{-1} - \Big(\mathbf{S}_N^{\mathrm{T}}\mathbf{S}_N\Big)^{-1}\mathbf{s}_{N+1}\Big(1+\mathbf{s}_{N+1}^{\mathrm{T}}\Big(\mathbf{S}_N^{\mathrm{T}}\mathbf{S}_N\Big)^{-1}\mathbf{s}_{N+1}\Big)^{-1}\mathbf{s}_{N+1}^{\mathrm{T}}\Big(\mathbf{S}_N^{\mathrm{T}}\mathbf{S}_N\Big)^{-1} \ . \tag{1.102}$$

Mit den Abkürzungen

$$\mathbf{P}_N = \Big(\mathbf{S}_N^{\mathrm{T}}\mathbf{S}_N\Big)^{-1} \tag{1.103a}$$

$$\mathbf{P}_{N+1} = \Big(\mathbf{S}_{N+1}^{\mathrm{T}}\mathbf{S}_{N+1}\Big)^{-1} \tag{1.103b}$$

$$\mathbf{k}_{N+1} = \frac{\mathbf{P}_N\mathbf{s}_{N+1}}{\Big(1+\mathbf{s}_{N+1}^{\mathrm{T}}\mathbf{P}_N\mathbf{s}_{N+1}\Big)} \tag{1.103c}$$

gilt nach (1.102)

$$\mathbf{P}_{N+1} = \mathbf{P}_N - \mathbf{k}_{N+1}\mathbf{s}_{N+1}^{\mathrm{T}}\mathbf{P}_N \tag{1.104}$$

und

$$\mathbf{P}_{N+1}\mathbf{s}_{N+1} = \frac{\mathbf{P}_N\mathbf{s}_{N+1}\Big(1+\mathbf{s}_{N+1}^{\mathrm{T}}\mathbf{P}_N\mathbf{s}_{N+1}\Big) - \mathbf{P}_N\mathbf{s}_{N+1}\mathbf{s}_{N+1}^{\mathrm{T}}\mathbf{P}_N\mathbf{s}_{N+1}}{\Big(1+\mathbf{s}_{N+1}^{\mathrm{T}}\mathbf{P}_N\mathbf{s}_{N+1}\Big)} = \mathbf{k}_{N+1} \ . \tag{1.105}$$

Setzt man die Beziehungen (1.103)–(1.105) in (1.100) ein, dann folgt

$$\hat{\mathbf{p}}_{N+1} = \mathbf{P}_{N+1}\Big(\mathbf{S}_N^{\mathrm{T}}\mathbf{y}_N + \mathbf{s}_{N+1}y_{N+1}\Big) = \mathbf{P}_N\mathbf{S}_N^{\mathrm{T}}\mathbf{y}_N - \mathbf{k}_{N+1}\mathbf{s}_{N+1}^{\mathrm{T}}\mathbf{P}_N\mathbf{S}_N^{\mathrm{T}}\mathbf{y}_N + \mathbf{k}_{N+1}y_{N+1} \tag{1.106}$$

bzw. mit (1.97)

$$\hat{\mathbf{p}}_{N+1} = \hat{\mathbf{p}}_N + \mathbf{k}_{N+1}\Big(y_{N+1} - \mathbf{s}_{N+1}^{\mathrm{T}}\hat{\mathbf{p}}_N\Big) \ . \tag{1.107}$$

Damit lässt sich die rekursive Least-Squares Identifikation wie folgt zusammenfassen:

(1) Man wählt geeignete Startwerte $\hat{\mathbf{p}}_{-1}$ und $\mathbf{P}_{-1}$ (siehe nachfolgende Diskussion).

(2) Für die Abtastzeitpunkte $j = 0, 1, \ldots$ misst man y_j und stellt den Datenvektor gemäß (1.74)

$$\mathbf{s}_j^{\mathrm{T}} = \begin{bmatrix} -y_{j-1} & -y_{j-2} & \ldots & -y_{j-n} & u_{j-d} & u_{j-d-1} & \ldots & u_{j-d-m} \end{bmatrix} \tag{1.108}$$

auf. Anschließend kann man den Parametervektor $\mathbf{p}$ unter Zuhilfenahme von $j+1$ Messungen mit der Iterationsvorschrift

$$\mathbf{k}_j = \frac{\mathbf{P}_{j-1}\mathbf{s}_j}{\left(1 + \mathbf{s}_j^{\mathrm{T}}\mathbf{P}_{j-1}\mathbf{s}_j\right)} \tag{1.109a}$$

$$\mathbf{P}_j = \mathbf{P}_{j-1} - \mathbf{k}_j\mathbf{s}_j^{\mathrm{T}}\mathbf{P}_{j-1} \tag{1.109b}$$

$$\hat{\mathbf{p}}_j = \hat{\mathbf{p}}_{j-1} + \mathbf{k}_j\left(y_j - \mathbf{s}_j^{\mathrm{T}}\hat{\mathbf{p}}_{j-1}\right) \tag{1.109c}$$

on-line schätzen.

Im letzten Schritt muss noch die Frage nach *geeigneten Startwerten* $\hat{\mathbf{p}}_{-1}$ und $\mathbf{P}_{-1}$ geklärt werden. Dazu betrachtet man in einem ersten Schritt die Iterationsvorschrift für $\mathbf{P}_k^{-1}$ (siehe (1.100) und (1.103))

$$\mathbf{P}_{k+1}^{-1} = \mathbf{P}_k^{-1} + \mathbf{s}_{k+1}\mathbf{s}_{k+1}^{\mathrm{T}} \tag{1.110}$$

und iteriert diese für $k = -1, 0, 1, \ldots, N, \ldots$

$$\mathbf{P}_0^{-1} = \mathbf{P}_{-1}^{-1} + \mathbf{s}_0\mathbf{s}_0^{\mathrm{T}} \tag{1.111a}$$

$$\mathbf{P}_1^{-1} = \mathbf{P}_{-1}^{-1} + \mathbf{s}_0\mathbf{s}_0^{\mathrm{T}} + \mathbf{s}_1\mathbf{s}_1^{\mathrm{T}} \tag{1.111b}$$

$$\vdots \quad \vdots$$

$$\mathbf{P}_N^{-1} = \mathbf{P}_{-1}^{-1} + \sum_{j=0}^{N}\mathbf{s}_j\mathbf{s}_j^{\mathrm{T}} = \mathbf{P}_{-1}^{-1} + \mathbf{S}_N^{\mathrm{T}}\mathbf{S}_N \ . \tag{1.111c}$$

Man erkennt, dass mit der Wahl

$$\mathbf{P}_{-1} = \alpha\mathbf{E} \tag{1.112}$$

für große Werte von α $\lim_{\alpha\to\infty}\mathbf{P}_{-1}^{-1} = \mathbf{0}$ gilt, und damit stimmt (1.111) mit dem zugehörigen Ausdruck der nichtrekursiven Schätzung von (1.79) überein. Für große Werte von α in (1.112) hat demnach $\mathbf{P}_{-1}^{-1}$ einen vernachlässigbar kleinen Einfluss auf das rekursiv berechnete $\mathbf{P}_N$. Die Iterationsvorschrift für $\hat{\mathbf{p}}_k$ ergibt sich mit Hilfe von (1.105) und (1.107) zu

$$\hat{\mathbf{p}}_{k+1} = \hat{\mathbf{p}}_k + \mathbf{P}_{k+1}\mathbf{s}_{k+1}\left(y_{k+1} - \mathbf{s}_{k+1}^{\mathrm{T}}\hat{\mathbf{p}}_k\right) \ . \tag{1.113}$$

Führt man nun diese Iteration für $k = -1, 0, 1, \ldots, N, \ldots$ durch, dann erhält man unter Zuhilfenahme von (1.111)

$$\hat{\mathbf{p}}_0 = \hat{\mathbf{p}}_{-1} + \mathbf{P}_0\mathbf{s}_0\left(y_0 - \mathbf{s}_0^{\mathrm{T}}\hat{\mathbf{p}}_{-1}\right) = \mathbf{P}_0\left(\mathbf{s}_0 y_0 + \underbrace{\left(\mathbf{P}_0^{-1} - \mathbf{s}_0\mathbf{s}_0^{\mathrm{T}}\right)}_{\mathbf{P}_{-1}^{-1}}\hat{\mathbf{p}}_{-1}\right) \tag{1.114a}$$

$$\hat{\mathbf{p}}_1 = \mathbf{P}_1\left(\mathbf{s}_1 y_1 + \mathbf{P}_0^{-1}\hat{\mathbf{p}}_0\right) = \mathbf{P}_1\left(\mathbf{s}_1 y_1 + \mathbf{s}_0 y_0 + \mathbf{P}_{-1}^{-1}\hat{\mathbf{p}}_{-1}\right) \tag{1.114b}$$

$$\vdots \quad \vdots$$

$$\hat{\mathbf{p}}_N = \mathbf{P}_N\left(\sum_{j=0}^{N}\mathbf{s}_j y_j + \mathbf{P}_{-1}^{-1}\hat{\mathbf{p}}_{-1}\right) = \mathbf{P}_N\left(\mathbf{S}_N^{\mathrm{T}}\mathbf{y}_N + \mathbf{P}_{-1}^{-1}\hat{\mathbf{p}}_{-1}\right) \ . \tag{1.114c}$$

Wie man aus (1.114) erkennt, bringt die Wahl für $\mathbf{P}_{-1}^{-1}$ von (1.112) für große α mit sich, dass der Startwert $\hat{\mathbf{p}}_{-1}$ frei gewählt werden kann und trotzdem $\hat{\mathbf{p}}_N$ von (1.114) mit dem Ergebnis der nichtrekursiven Schätzung von (1.79) übereinstimmt. Sehr oft wählt man daher einfachheitshalber $\hat{\mathbf{p}}_{-1} = \mathbf{0}$.

1.3.5 Methode der gewichteten kleinsten Quadrate

Bei der Methode der gewichteten kleinsten Quadrate sucht man eine Lösung des überbestimmten linearen Gleichungssystems (siehe (1.60))

$$\mathbf{y}_N = \mathbf{S}_N \mathbf{p} \tag{1.115}$$

in der Form, dass der *gewichtete quadratische Fehler*

$$\sum_{j=0}^{N} \alpha_j e_j^2 \quad , \quad e_j = y_j - \mathbf{s}_j^{\mathrm{T}} \mathbf{p} \tag{1.116}$$

mit der Folge positiver Gewichtungskoeffizienten α_j ($\alpha_j > 0$ für alle j) bezüglich $\mathbf{p}$ minimiert wird. Man erkennt, dass sich durch die Wahl von

$$\tilde{y}_j = \sqrt{\alpha_j} y_j \quad \text{und} \quad \tilde{\mathbf{s}}_j^{\mathrm{T}} = \sqrt{\alpha_j} \mathbf{s}_j^{\mathrm{T}} \tag{1.117}$$

das Optimierungsproblem (1.116) in das klassische Least-Squares Problem gemäß (1.61)

$$\min_{\mathbf{p}} \sum_{j=0}^{N} \tilde{e}_j^2 = \min_{\mathbf{p}} \|\tilde{\mathbf{e}}_N\|_2^2 \quad \text{mit} \quad \tilde{\mathbf{e}}_N = \tilde{\mathbf{y}}_N - \tilde{\mathbf{S}}_N \mathbf{p} \tag{1.118}$$

überführen lässt. Der zugehörige rekursive Schätzer nach (1.109) lautet demnach

$$\tilde{\mathbf{k}}_j = \frac{\tilde{\mathbf{P}}_{j-1} \tilde{\mathbf{s}}_j}{\left(1 + \tilde{\mathbf{s}}_j^{\mathrm{T}} \tilde{\mathbf{P}}_{j-1} \tilde{\mathbf{s}}_j\right)} \tag{1.119a}$$

$$\tilde{\mathbf{P}}_j = \tilde{\mathbf{P}}_{j-1} - \tilde{\mathbf{k}}_j \tilde{\mathbf{s}}_j^{\mathrm{T}} \tilde{\mathbf{P}}_{j-1} \tag{1.119b}$$

$$\hat{\mathbf{p}}_j = \hat{\mathbf{p}}_{j-1} + \tilde{\mathbf{k}}_j \left(\tilde{y}_j - \tilde{\mathbf{s}}_j^{\mathrm{T}} \hat{\mathbf{p}}_{j-1}\right) . \tag{1.119c}$$

Setzt man die Beziehung (1.117) sowie die Transformation

$$\tilde{\mathbf{k}}_j = \frac{\mathbf{k}_j}{\sqrt{\alpha_j}} \quad \text{und} \quad \tilde{\mathbf{P}}_j = \mathbf{P}_j \tag{1.120}$$

in (1.119) ein, so erhält man

$$\mathbf{k}_j = \frac{\mathbf{P}_{j-1} \mathbf{s}_j}{\left(\frac{1}{\alpha_j} + \mathbf{s}_j^{\mathrm{T}} \mathbf{P}_{j-1} \mathbf{s}_j\right)} \tag{1.121a}$$

$$\mathbf{P}_j = \mathbf{P}_{j-1} - \mathbf{k}_j \mathbf{s}_j^{\mathrm{T}} \mathbf{P}_{j-1} \tag{1.121b}$$

$$\hat{\mathbf{p}}_j = \hat{\mathbf{p}}_{j-1} + \mathbf{k}_j \left(y_j - \mathbf{s}_j^{\mathrm{T}} \hat{\mathbf{p}}_{j-1}\right) . \tag{1.121c}$$

Eine spezielle Wahl für α_j ist durch

$$\alpha_j = q^{N-j} \quad \text{mit} \quad 0 < q \leq 1 \tag{1.122}$$

gegeben. Durch Einsetzen von (1.122) in das Gütekriterium von (1.116)

$$\sum_{j=0}^{N} q^{N-j} e_j^2 = \|\mathbf{e}_N\|_Q^2 = \mathbf{e}_N^{\mathrm{T}} \mathbf{Q}_N \mathbf{e}_N \quad \text{mit} \quad \mathbf{Q}_N = \begin{bmatrix} q^N & 0 & \cdots & 0 & 0 \\ 0 & q^{N-1} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & q & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \tag{1.123}$$

ist ersichtlich, dass mit dieser speziellen Wahl von α_j die Gleichungsfehler umso geringer gewichtet werden, je weiter sie zurückliegen. Man spricht dann auch von einem *exponentiell abklingenden Gedächtnis* mit dem *Gedächtnisfaktor* q. Für diesen Gedächtnisfaktor q haben sich in der praktischen Anwendung Werte im Bereich von $0.9 < q < 0.995$ als sinnvoll erwiesen. Man beachte, dass sich diese Methode auch für die Identifikation *zeitlich langsam veränderlicher Strecken* eignet, wobei die Geschwindigkeit der Parameteränderungen der Strecke natürlich die Wahl von q, also die Geschwindigkeit des Vergessens alter Messwerte, festlegt.

Aufgabe 1.15. Woher resultiert Ihrer Meinung nach der Name exponentiell abklingendes Gedächtnis?

Der rekursive Least-Squares Schätzer mit exponentiell abklingendem Gedächtnis ergibt sich unmittelbar aus (1.121) mit $\alpha_j = q^{N-j}$ und der Transformation $\mathbf{P}_j \to \mathbf{P}_j/q^{N-j}$ zu

$$\mathbf{k}_j = \frac{\mathbf{P}_{j-1}\mathbf{s}_j}{\left(q + \mathbf{s}_j^{\mathrm{T}}\mathbf{P}_{j-1}\mathbf{s}_j\right)} \tag{1.124a}$$

$$\mathbf{P}_j = \left(\mathbf{P}_{j-1} - \mathbf{k}_j \mathbf{s}_j^{\mathrm{T}} \mathbf{P}_{j-1}\right)\frac{1}{q} \tag{1.124b}$$

$$\hat{\mathbf{p}}_j = \hat{\mathbf{p}}_{j-1} + \mathbf{k}_j\left(y_j - \mathbf{s}_j^{\mathrm{T}}\hat{\mathbf{p}}_{j-1}\right) . \tag{1.124c}$$

Aufgabe 1.16. Überprüfen Sie die Richtigkeit von (1.124).

Natürlich könnte anstelle der Diagonalmatrix $\mathbf{Q}_N$ in (1.123) jede beliebige positiv definite *Gewichtungsmatrix* $\mathbf{Q}_N > 0$ gewählt werden. In diesem Fall wird nicht wie in (1.61) die Zweiernorm des Fehlers $\|\mathbf{e}_N\|_2^2$, sondern eine gewichtete Zweiernorm der Form

$$\|\mathbf{e}_N\|_Q^2 = \mathbf{e}_N^{\mathrm{T}} \mathbf{Q}_N \mathbf{e}_N \tag{1.125}$$

minimiert. Man überzeugt sich nun leicht, dass diese Aufgabe

$$\min_{\mathbf{p}} \|\mathbf{e}_N\|_Q^2 \quad \text{mit} \quad \mathbf{e}_N = \mathbf{y}_N - \mathbf{S}_N \mathbf{p} \tag{1.126}$$

durch Anwendung des Projektionstheorems von Satz 1.1 mit dem inneren Produkt

$$\langle \mathbf{x}, \mathbf{z} \rangle = \mathbf{x}^{\mathrm{T}} \mathbf{Q} \mathbf{z} \tag{1.127}$$

auf die Lösung

$$\hat{\mathbf{p}}_N = \left(\mathbf{S}_N^{\mathrm{T}} \mathbf{Q}_N \mathbf{S}_N \right)^{-1} \mathbf{S}_N^{\mathrm{T}} \mathbf{Q}_N \mathbf{y}_N \tag{1.128}$$

führt.

Aufgabe 1.17. Zeigen Sie die Gültigkeit von (1.128). Überprüfen Sie weiters, ob der Fehler $\mathbf{e}_N = \mathbf{y}_N - \mathbf{S}_N \hat{\mathbf{p}}_N$ im Sinne des inneren Produktes (1.127) auf die optimale Lösung (1.128) orthogonal steht.

Aufgabe 1.18. Zeigen Sie, dass jede positiv definite Matrix $\mathbf{Q}$ genau eine positiv definite Wurzel $\mathbf{W}$ hat, d. h. $\mathbf{Q} = \mathbf{W}^{\mathrm{T}} \mathbf{W}$.

Hinweis: (zu Aufgabe 1.18) Benutzen Sie die Tatsache, dass jede symmetrische Matrix $\mathbf{Q}$ durch eine Ähnlichkeitstransformation auf Diagonalform $\mathbf{D}$ gebracht werden kann, es gilt also

$$\mathbf{D} = \mathbf{T} \mathbf{Q} \mathbf{T}^{\mathrm{T}} \quad \text{mit} \quad \mathbf{T} \mathbf{T}^{\mathrm{T}} = \mathbf{E} \; .$$

Aufgabe 1.19. Beweisen Sie, dass (1.125) eine Norm im Sinne von Definition 1.2 ist.

Aufgabe 1.20. Gegeben ist die kontinuierliche Übertragungsfunktion

$$G(s) = \frac{K}{(sT_1 + 1)(sT_2 + 1)}$$

mit den Parametern $K, T_1, T_2 > 0$. Bestimmen Sie eine Zustandsdarstellung für $G(s)$ und implementieren Sie diese in MATLAB/SIMULINK für einen zeitveränderlichen Parameter T_2 und konstante Parameter K und T_1. Bestimmen Sie anschließend die zugehörige z-Übertragungsfunktion zu $G(s)$ für eine Abtastzeit von $T_a = 0.25\,\mathrm{s}$. Identifizieren Sie die Koeffizienten a_j und b_j des Nenner- bzw. Zählerpolynoms der diskreten Übertragungsfunktion mit Hilfe des rekursiven Least-Squares Algorithmus. Wählen Sie als Eingangssignal einen mehrfachen 3-2-1-Sprung und für die Parameter $K = 1$, $T_1 = 7.5\,\mathrm{s}$ sowie eine sprunghafte Änderung von T_2 von $5\,\mathrm{s}$ auf $2.5\,\mathrm{s}$. Untersuchen Sie den Einfluss des Vergessensfaktors q und vergleichen Sie das Ergebnis für $q = 1$ mit dem Ergebnis des off-line Least-Squares Verfahrens.

1.3.6 Least-Mean Squares (LMS) Identifikation

Eine weitere Möglichkeit, das Gleichungssystem (1.71)

$$\left(\sum_{j=0}^{N} \mathbf{s}_j \mathbf{s}_j^{\mathrm{T}}\right) \hat{\mathbf{p}}_N = \mathbf{S}_N^{\mathrm{T}} \mathbf{S}_N \hat{\mathbf{p}}_N = \mathbf{S}_N^{\mathrm{T}} \mathbf{y}_N = \sum_{j=0}^{N} \mathbf{s}_j y_j \tag{1.129}$$

on-line zu lösen, ist durch den so genannten *Least-Mean Squares (LMS) Algorithmus*, auch als *stochastisches Gradientenverfahren* bezeichnet, gegeben. Dabei wird der Parametervektor $\mathbf{p}$ in der Form

$$\hat{\mathbf{p}}_j = \hat{\mathbf{p}}_{j-1} + \mu_j \mathbf{s}_j \underbrace{\left(y_j - \mathbf{s}_j^{\mathrm{T}} \hat{\mathbf{p}}_{j-1}\right)}_{e_j} \tag{1.130}$$

mit einem geeignet gewählten Startwert $\hat{\mathbf{p}}_{-1}$ sowie dem Schätzfehler zum j-ten Zeitpunkt e_j rekursiv geschätzt. Mit Hilfe des *(zeitvarianten) Parameters* μ_j ($\mu_j \geq 0$ für alle $j \geq 0$) besteht die Möglichkeit, die Konvergenzgeschwindigkeit sowie die Empfindlichkeit gegenüber Rauschen in einem gewissen Maß einzustellen. Typischerweise wählt man für μ_j einen konstanten Wert $\mu_j = \bar{\mu}$, wobei der Parameterschätzer (1.130) für größere Werte von $\bar{\mu}$ in der Lage ist, schneller auf Änderungen des Parametervektors $\mathbf{p}$ (zeitvariante Strecke) zu reagieren, und für kleinere Werte von $\bar{\mu}$ unempfindlicher gegenüber Messrauschen ist. Um die Konvergenzgeschwindigkeit unabhängig vom Signalpegel der Einträge in $\mathbf{s}_j$ zu machen, ist es üblich, den (zeitvarianten) Parameter μ_j im LMS Algorithmus (1.130) in der Form

$$\mu_j = \frac{\bar{\mu}}{\mathbf{s}_j^{\mathrm{T}} \mathbf{s}_j} \tag{1.131}$$

bzw.

$$\mu_j = \frac{\bar{\mu}}{l_j} \quad \text{mit} \quad l_{j+1} = l_j + \bar{\mu}\left(\mathbf{s}_j^{\mathrm{T}} \mathbf{s}_j - l_j\right) \tag{1.132}$$

mit dem Anfangswert $l_{-1} = \mathbf{s}_{-1}^{\mathrm{T}} \mathbf{s}_{-1} > 0$ zu wählen. Ohne Beweis sei angemerkt, dass für hinreichend kleines $\bar{\mu}$ die Konvergenz von (1.130) gewährleistet werden kann.

Bemerkung 1.2. Der LMS Algorithmus wird sehr häufig im Rahmen von Anwendungen bei der adaptiven Filterung von Signalen (*adaptive noise cancellation*), wie beispielsweise bei der Echokompensation von übertragenen Signalen, verwendet. Man betrachte dazu die Anordnung von Abbildung 1.11, bei der sich die gemessene Signalfolge (d_k) aus einem zu identifizierenden Nutzsignalanteil (r_k) und einem nicht-messbaren Störsignalanteil $G_1(\delta)m_k$ in der Form

$$d_k = G_1(\delta) m_k + r_k \tag{1.133}$$

mit dem Störsignal (m_k) und dem unbekannten Übertragungsoperator $G_1(\delta)$ zusammensetzt. Im Weiteren sei angenommen, dass das Störsignal (m_k) und das Nutzsignal (r_k) unkorreliert sind. Das Störsignal (m_k) unterliegt zwar keiner direkten Messung, kann aber indirekt über das Signal (n_k) mit

$$n_k = G_2(\delta) m_k \tag{1.134}$$

und dem unbekannten Übertragungsoperator $G_2(\delta)$ erfasst werden. Würde man die beiden Übertragungsoperatoren $G_1(\delta)$ und $G_2(\delta)$ kennen, dann könnte man aus Kenntnis von (d_k) und (n_k) das Nutzsignal in der Form

$$r_k = d_k - \underbrace{G_1(\delta)G_2^{-1}(\delta)}_{\tilde{G}(\delta)} n_k \tag{1.135}$$

rekonstruieren. Man beachte, dass $\tilde{G}(\delta) = G_1(\delta)G_2^{-1}(\delta)$ im Allgemeinen ein nichtkausaler Übertragungsoperator ist.

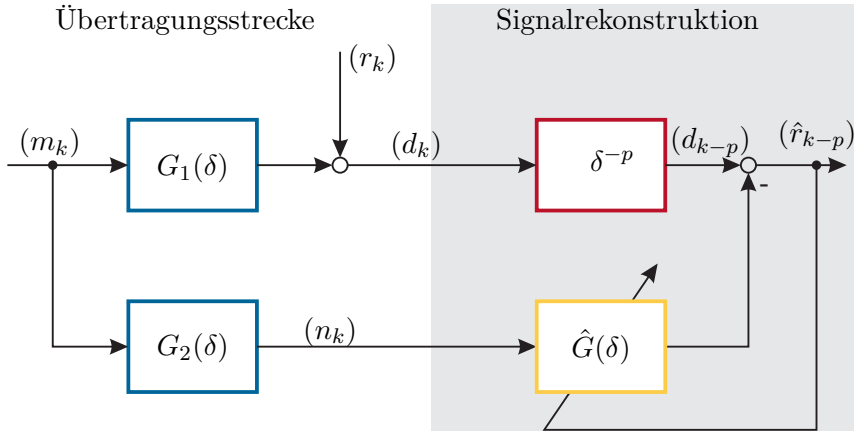


Abbildung 1.11: Zur adaptiven Signalfilterung mit dem LMS-Algorithmus.

Aufgabe 1.21. Unter welchen Voraussetzungen an $G_1(\delta)$ und $G_2(\delta)$ ist $\tilde{G}(\delta)$ ein kausaler Übertragungsoperator?

Aus diesem Grund wählt man eine positive natürliche Zahl $p > 0$ so, dass der Übertragungsoperator

$$G(\delta) = \tilde{G}(\delta)\delta^{-p} \tag{1.136}$$

auf alle Fälle kausal ist. Ersetzt man in (1.135) $\tilde{G}(\delta)$ durch $G(\delta)$ gemäß (1.136), dann erhält man

$$r_k = d_k - G(\delta)\delta^p n_k \tag{1.137a}$$
$$\delta^{-p} r_k = \delta^{-p} d_k - G(\delta) n_k \tag{1.137b}$$
$$r_{k-p} = d_{k-p} - G(\delta) n_k \ . \tag{1.137c}$$

Da aber der Übertragungsoperator $G(\delta)$ in (1.137) unbekannt ist, setzt man anstelle von $G(\delta)$ einen zu identifizierenden Übertragungsoperator $\hat{G}(\delta)$ ein. In den meisten Anwendungsfällen wird hierfür in Kombination mit dem LMS-Algorithmus ein MA-Modell (FIR-Filter) nach (1.47), (1.48) der Form

$$\hat{G}(\delta) = \hat{C}(\delta) = \hat{c}_0 + \hat{c}_1\delta^{-1} + \ldots + \hat{c}_q\delta^{-q} \tag{1.138}$$

angesetzt. Natürlich ist der Übertragungsoperator $G(\delta)$ nur in Ausnahmefällen durch ein MA-Modell beschreibbar, weshalb für eine gute Rekonstruktion des Nutzsignales (r_k) eine sehr hohe Modellordnung q des MA-Modells erforderlich ist. Setzt man für $G(\delta)$ in (1.137) den Ausdruck $\hat{G}(\delta)$ von (1.138) ein, dann errechnet sich der Schätzwert des Nutzsignales $\hat{r}_{k-p}$ zu

$$\hat{r}_{k-p} = d_{k-p} - \underbrace{\begin{bmatrix} n_k & n_{k-1} & n_{k-2} & \dots & n_{k-q} \end{bmatrix}}_{\mathbf{s}_k^{\mathrm{T}}} \underbrace{\begin{bmatrix} \hat{c}_0 \\ \hat{c}_1 \\ \vdots \\ \hat{c}_q \end{bmatrix}}_{\hat{\mathbf{p}}} \tag{1.139}$$

mit dem Schätzwert $\hat{\mathbf{p}}$ des Parametervektors und dem Datenvektor $\mathbf{s}_k^{\mathrm{T}}$. Durch Anwendung des LMS-Algorithmus (1.130) folgt der Parameterschätzer zu

$$\hat{\mathbf{p}}_j = \hat{\mathbf{p}}_{j-1} + \mu_j \mathbf{s}_j \underbrace{\left(d_{j-p} - \mathbf{s}_j^{\mathrm{T}} \hat{\mathbf{p}}_{j-1}\right)}_{\hat{\mathbf{r}}_{j-p}} \tag{1.140}$$

mit dem Parameter μ_j gemäß (1.131) oder (1.132) und einem geeignet zu wählenden Startwert $\hat{\mathbf{p}}_{-1}$. Abbildung 1.11 gibt eine grafische Veranschaulichung des Algorithmus.

Aufgabe 1.22. Gegeben ist eine Übertragungsstrecke nach Abbildung 1.11 mit der gemessenen Signalfolge (d_k), in der ein periodisches Nutzsignal $(s_k) = (20\sin(3t - \pi/4))$ verborgen ist. Das Störsignal (m_k) ist weißes Rauschen, die Abtastzeit beträgt 0.1 s und die beiden Übertragungsoperatoren lauten

$$G_1(\delta) = 10\frac{\delta + 2}{0.8 + \delta + \delta^2} \quad \text{sowie} \quad G_2(\delta) = \frac{1 - 2\delta}{(\delta + 0.2)^2(\delta - 0.8)} \ .$$

Entwerfen Sie einen on-line Algorithmus nach der LMS-Methode zur Extraktion des Nutzsignals mit Hilfe eines MA-Modells. Führen Sie diese Berechnungen in MATLAB durch und kontrollieren Sie die Ergebnisse auch im Frequenzbereich mit Hilfe der FFT. Geben Sie verschiedene Ordnungen der MA-Modelle vor und ändern Sie die Signalverzögerung p nach (1.136). Wie wirken sich diese Vorgaben und Änderungen auf das Ergebnis aus?

1.4 Literatur

[1.1] L. Ljung, *System Identification*. New Jersey, USA: Prentice Hall, 1999.

[1.2] A. Kugi, „Regelungssysteme," in *Skriptum zur Vorlesung, Institut für Automatisierungstechnik, TU Wien*, Bd. 1, Vienna, Austria, Sep. 2011.

[1.3] R. Isermann, *Identifikation dynamischer Systeme 1 und 2*, 2. Aufl. Berlin, Deutschland: Springer, 1992.

[1.4] G. Franklin, J. Powell und M. Workman, *Digital Control of Dynamic Systems*, 3. Aufl. Menlo Park, USA: Addison–Weseley, 1998.

[1.5] M. Gevers und G. Li, *Parametrization in Control, Estimation and Filtering Problems*. London, UK: Springer, 1993.

[1.6] E. Kreyszig, *Statistische Methoden und ihre Anwendungen*, 7. Aufl. Göttingen, Deutschland: Vandenhoeck & Ruprecht, 1998.

[1.7] L. Ljung und T. Söderström, *Theory and Practice of Recursive Identification*. Cambridge, USA: MIT Press, 1987.

[1.8] D. Luenberger, *Optimization by Vector Space Methods*. New York, USA: John Wiley & Sons, 1969.

[1.9] O. Nelles, *Nonlinear System Identification*. Berlin, Deutschland: Springer, 2001.

[1.10] A. Papoulis und S. Pillai, *Probability, Random Variables and Stochastic Processes*, 4. Aufl. New York, USA: McGraw-Hill, 2002.

2 Optimale Schätzer

Dieses Kapitel beschäftigt sich mit der Frage, wie aus einer Eingangsfolge (u_k) und einer gemessenen Ausgangsfolge (y_k) der Zustand $\mathbf{x}_{k+m}$ eines dynamischen Systems geschätzt werden kann. Je nach Wert von m wird der Vorgang des Schätzens als

(1) Glätten (smoothing) für $m < 0$

(2) Filtern (filtering) für $m = 0$ oder

(3) Vorhersagen (prediction) für $m > 0$

bezeichnet. Abbildung 2.1 gibt eine grafische Veranschaulichung dieser drei Fälle.

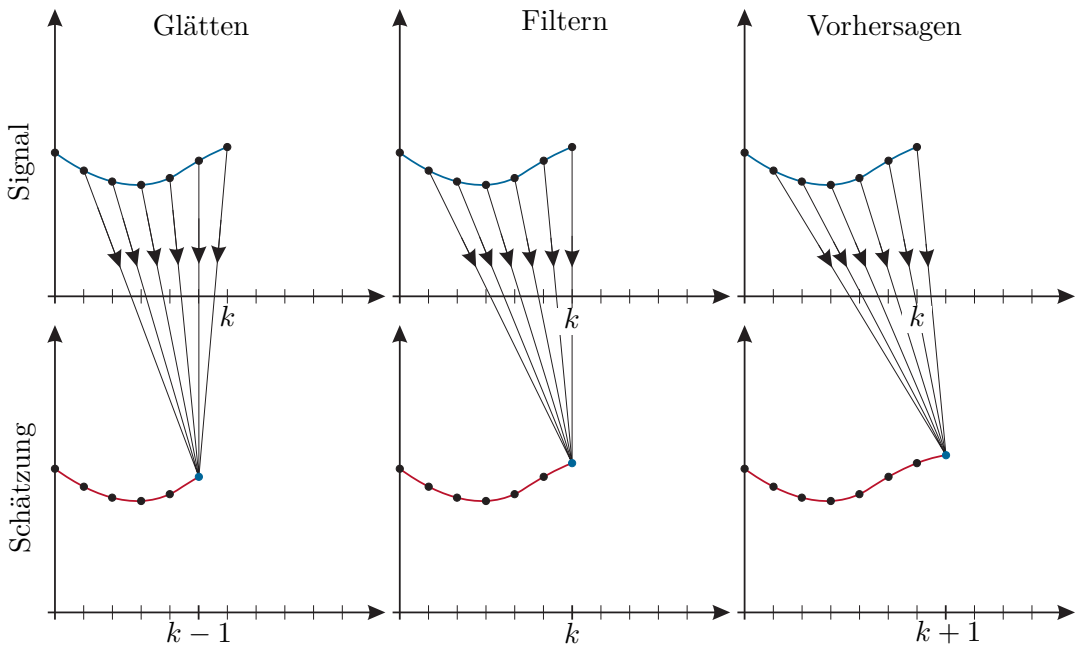


Abbildung 2.1: Zu den Begriffen Glätten, Filtern und Vorhersagen.

Die weiteren Betrachtungen werden sich auf den letzten Fall (3) beschränken, da dieser für regelungstechnische Anwendungen am interessantesten ist. Als Ergebnis der nachfolgenden Betrachtungen wird ein optimaler Zustandsbeobachter, das so genannte *Kalman-Filter*, ermittelt werden, der ein quadratisches Gütekriterium minimiert. Dazu müssen jedoch in einem Zwischenschritt die Ergebnisse der Least-Squares Schätzung des vorigen Kapitels erweitert werden.

Da in diesem Kapitel immer wieder der Erwartungswert und die Kovarianz von Zufallszahlen verwendet werden, sollen diese beiden Begriffe für normalverteilte und gleichverteilte Zufallszahlen erläutert werden.

Bemerkung 2.1 (Normalverteilte Zufallsvariablen). Eine skalare normalverteilte Zufallsvariable x ist durch die Verteilungsdichtefunktion (Gauß'sche Verteilung)

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \mathrm{e}^{-\frac{(x-m)^2}{2\sigma^2}}, \tag{2.1}$$

mit dem Mittelwert (Erwartungswert) m und der Varianz σ, definiert. Der Mittelwert (Erwartungswert, erstes Moment) und die Varianz (zweites zentrales Moment) sind, wie im Anhang A erläutert, durch

$$\mathrm{E}(x) = m = \int_{-\infty}^{\infty} x f(x)\,\mathrm{d}x \tag{2.2a}$$

$$\mathrm{E}\Big((x - \mathrm{E}(x))^2\Big) = \sigma^2 = \int_{-\infty}^{\infty} (x - \mathrm{E}(x))^2 f(x)\,\mathrm{d}x \tag{2.2b}$$

definiert. In Abbildung 2.2 ist die Wahrscheinlichkeitsdichtefunktion für unterschiedliche Parametrierungen einer normalverteilten Zufallszahl dargestellt.

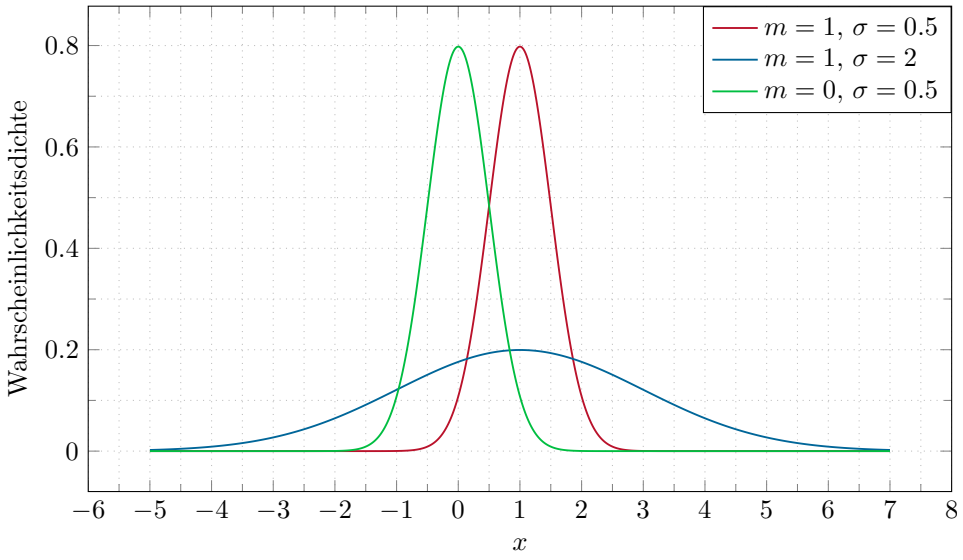


Abbildung 2.2: Wahrscheinlichkeitsdichtefunktion für unterschiedliche normalverteilte Zufallsvariablen.

Die Wahrscheinlichkeit, dass eine Zufallszahl x im Intervall δ um den Erwartungswert $\mathrm{E}(x) = m$ liegt errechnet sich zu

$$\mathrm{P}(m-\delta < x \leq m+\delta) = \int_{m-\delta}^{m+\delta} f(x)\,\mathrm{d}x = \mathrm{erf}\left(\frac{\delta}{\sqrt{2}\sigma}\right), \tag{2.3}$$

vgl. Aufgabe A.1 im Anhang A. Wählt man z.B. $\delta = \sigma$, so liegt eine Zufallszahl x mit einer Wahrscheinlichkeit von 0.68 in diesem Intervall. Weiterhin liegt eine Zufallszahl x mit einer Wahrscheinlichkeit von 0.95 im Intervall $\delta = 2\sigma$. Die Varianz stellt damit ein Maß für die Streuung der Zufallszahlen um den Erwartungswert dar.

Die Verbundwahrscheinlichkeitsdichtefunktion $f(\mathbf{x})$ eines n-dimensionalen normalverteilten Zufallsvektors $\mathbf{x}$ errechnet sich in der Form

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{n/2}\sqrt{\det(\mathbf{Q})}} \mathrm{e}^{-\frac{1}{2}(\mathbf{x}-\mathbf{m})^{\mathrm{T}}\mathbf{Q}^{-1}(\mathbf{x}-\mathbf{m})}, \tag{2.4}$$

mit dem Erwartungswert $\mathbf{m}$ und der Kovarianzmatrix $\mathbf{Q}$,

$$\mathbf{m} = \mathrm{E}(\mathbf{x}) \tag{2.5a}$$

$$\mathbf{Q} = \mathrm{E}\Big((\mathbf{x}-\mathrm{E}(\mathbf{x}))(\mathbf{x}-\mathrm{E}(\mathbf{x}))^{\mathrm{T}}\Big)\,. \tag{2.5b}$$

Für den Sonderfall $n = 2$ ist die Wahrscheinlichkeit, dass ein Zufallsvektor $\mathbf{x} = [x_1, x_2]^{\mathrm{T}}$ in einer Ellipse der Form

$$(\mathbf{x}-\mathbf{m})^{\mathrm{T}}\mathbf{Q}^{-1}(\mathbf{x}-\mathbf{m}) = C^2 \tag{2.6}$$

liegt, durch

$$\mathrm{P} = 1 - \mathrm{e}^{-\frac{C^2}{2}} \tag{2.7}$$

gegeben. Der Beweis für diese Aussage kann z.B. in [2.1] gefunden werden.

Damit kann mit Hilfe der Kovarianzmatrix $\mathbf{Q}$ das Gebiet (d.h. die Ellipsen für $n = 2$) ermittelt werden, in dem ein Zufallsvektor $\mathbf{x}$ mit einer Wahrscheinlichkeit P liegt. In der Abbildung 2.3 ist die Verteilung von 3000 Zufallsvektoren $\mathbf{x} = [x_1, x_2]^{\mathrm{T}}$, mit den normalverteilten, nicht korrelierten Zufallsvariablen x_1 ($\mathrm{E}(x_1) = m_1 = 1$, $\sigma_1 = 2$) und x_2 ($\mathrm{E}(x_2) = m_2 = 2$, $\sigma_2 = 1$) dargestellt. Weiterhin sind die Ellipsen für $\mathrm{P} = 0.5$ und $\mathrm{P} = 0.95$, und der Erwartungswert $\mathrm{E}(\mathbf{x}) = \mathbf{m}$ eingezeichnet.

Diese Betrachtung für $n = 2$ kann für n-dimensionale normalverteilte Zufallsvariablen verallgemeinert werden. In diesem Fall beschreibt das n-dimensionale Ellipsoid $(\mathbf{x}-\mathbf{m})^{\mathrm{T}}\mathbf{Q}^{-1}(\mathbf{x}-\mathbf{m}) = C^2$ ein Maß für die Verteilung des Zufallsvektors. Die Wahrscheinlichkeit P dafür, dass ein Zufallsvektor $\mathbf{x}$ in diesem Ellipsoid liegt, ist durch

$$1 - \mathrm{P} = \frac{n}{2^{n/2}\Gamma\left(\frac{n}{2}+1\right)} \int_C^\infty \xi^{n-1}\mathrm{e}^{-\frac{\xi^2}{2}}\,\mathrm{d}\xi, \tag{2.8}$$

mit der Gamma-Funktion Γ, gegeben [2.1].

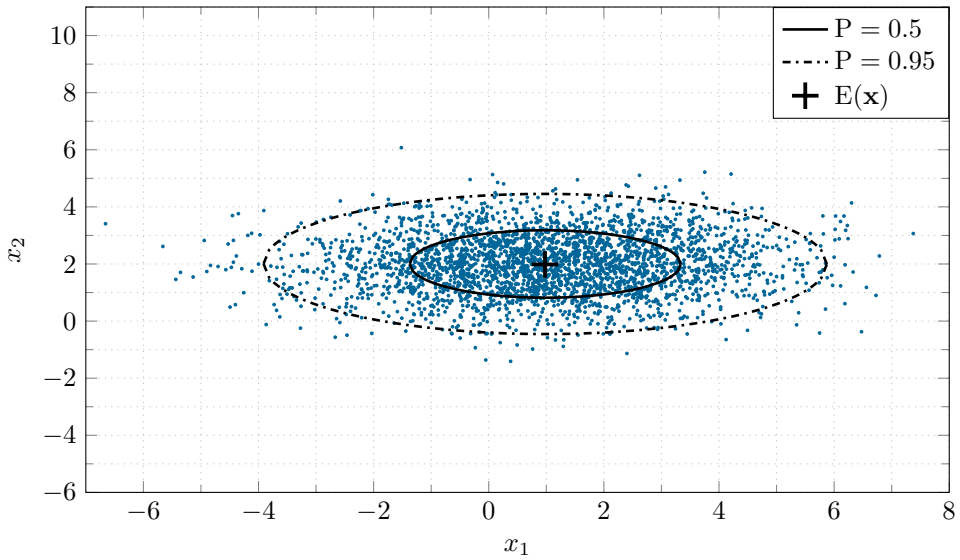


Abbildung 2.3: Grafische Darstellung des Erwartungswertes $\mathrm{E}(\mathbf{x})$ und der Kovarianzmatrix $\mathbf{Q}$ für normalverteilte Zufallsvektoren $\mathbf{x}$.

Ein exakter Zusammenhang zwischen den durch die Kovarianzmatrix $\mathbf{Q}$ definierten Ellipsoiden und der Wahrscheinlichkeit, dass ein Zufallsvektor $\mathbf{x}$ in diesem Gebiet liegt, ist nur für normalverteilte Zufallsvariablen definiert. Für andere Verteilungen stellen diese Ellipsoide nur eine mehr oder weniger genaue Approximation dar. Trotzdem ist die Kovarianzmatrix $\mathbf{Q}$ auch für diese Verteilungen ein sinnvolles Maß zur Abschätzung der Verteilung der Zufallsvektoren, weswegen auch hier häufig die zugehörigen Ellipsen bzw. Ellipsoide dargestellt werden.

Bemerkung 2.2. Eine im Intervall $[a, b]$ gleichverteilte skalare Zufallsvariable x wird durch die Wahrscheinlichkeitsdichtefunktion

$$f(x) = \begin{cases} \frac{1}{b-a} & a < x \leq b \\ 0 & \text{sonst} \end{cases} \tag{2.9}$$

definiert. Der zugehörige Erwartungwert m und die Varianz σ errechnen sich zu

$$\mathrm{E}(x) = m = \frac{a+b}{2} \tag{2.10a}$$

$$\mathrm{E}\Big((x - \mathrm{E}(x))^2\Big) = \sigma^2 = \frac{1}{12}(b-a)^2 \ . \tag{2.10b}$$

In der Abbildung 2.4 sind die Wahrscheinlichkeitsdichtefunktionen für unterschiedliche gleichverteilte Zufallsvariablen dargestellt.

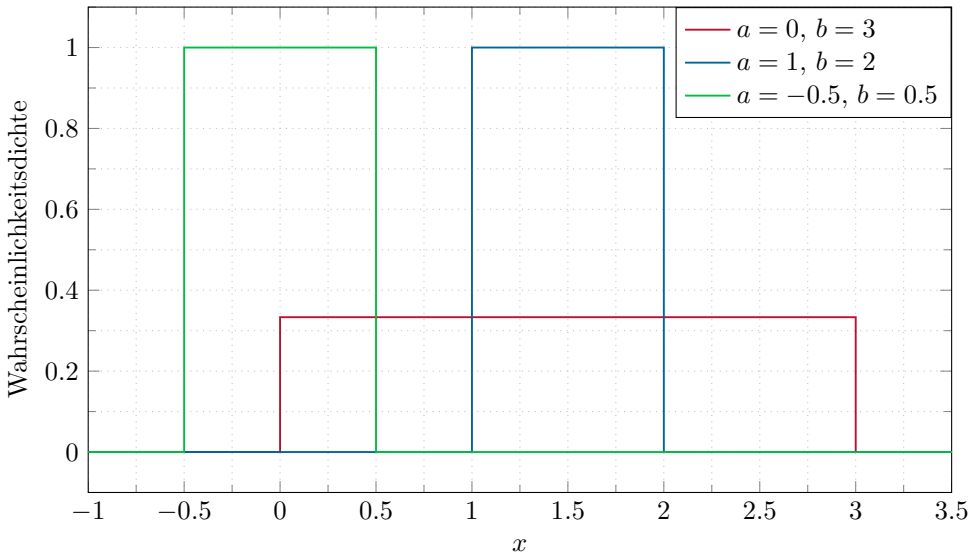


Abbildung 2.4: Wahrscheinlichkeitsdichtefunktion für unterschiedliche gleichverteilte Zufallsvariablen.

2.1 Gauß-Markov-Schätzung

Man betrachte das um die stochastische Störung $\mathbf{v}$ erweiterte, überbestimmte lineare Gleichungssystem von (1.60) (vergleiche dazu auch (1.87)-(1.90))

$$\mathbf{y} = \mathbf{S}\mathbf{p} + \mathbf{v} \tag{2.11}$$

mit der bekannten $(m \times n)$-Matrix $\mathbf{S} \in \mathbb{R}^{m \times n}$, dem n-dimensionalen Vektor der Unbekannten $\mathbf{p} \in \mathbb{R}^n$ sowie dem m-dimensionalen Messvektor $\mathbf{y} \in \mathbb{R}^m$. Man nimmt nun an, dass die stochastische Störung $\mathbf{v}$ folgende stochastische Eigenschaften besitzt

$$\mathrm{E}(\mathbf{v}) = \mathbf{0} \quad \text{und} \quad \mathrm{cov}(\mathbf{v}) = \mathrm{E}\Big(\mathbf{v}\mathbf{v}^{\mathrm{T}}\Big) = \mathbf{Q} \quad \text{mit} \quad \mathbf{Q} > 0\ . \tag{2.12}$$

Es sei an dieser Stelle angemerkt, dass $\mathbf{v}$ auch als Messfehler interpretiert werden kann. Gesucht ist nun ein *linearer Schätzer* der Form

$$\hat{\mathbf{p}} = \mathbf{K}\mathbf{y} \tag{2.13}$$

mit einer konstanten $(n \times m)$-Matrix $\mathbf{K} \in \mathbb{R}^{n \times m}$. Da $\mathbf{y}$ die Summe eines konstanten Vektors $\mathbf{S}\mathbf{p}$ und eines stochastischen Vektors (Vektor mit stochastischen Einträgen) $\mathbf{v}$ ist, sind $\mathbf{y}$, der geschätzte Parameter $\hat{\mathbf{p}}$ nach (2.13) und der Parameterfehler $\mathbf{e} = \mathbf{p} - \hat{\mathbf{p}}$ stochastische Größen. Es macht deshalb keinen Sinn, die Matrix $\mathbf{K}$ so zu bestimmen, dass $\|\mathbf{e}\|_2^2$ minimiert wird, sondern es muss die Aufgabe

$$\min_{\mathbf{K}} \mathrm{E}\Big(\|\mathbf{e}\|_2^2\Big) = \min_{\mathbf{K}} \mathrm{E}\Big([\mathbf{p} - \mathbf{K}\mathbf{y}]^{\mathrm{T}}[\mathbf{p} - \mathbf{K}\mathbf{y}]\Big) \tag{2.14}$$

gelöst werden.

Setzt man nun für $\mathbf{y}$ die Beziehung (2.11) in (2.14) ein, dann erhält man unter Berücksichtigung von (2.12) und den Ergebnissen von Anhang A (siehe Aufgabe A.3)

$$\begin{aligned}
&\min_{\mathbf{K}} \mathrm{E}\Big([\mathbf{p}-\mathbf{KSp}-\mathbf{Kv}]^{\mathrm{T}}[\mathbf{p}-\mathbf{KSp}-\mathbf{Kv}]\Big) = \\
&\min_{\mathbf{K}} \left\{ \mathrm{E}\Big([\mathbf{p}-\mathbf{KSp}]^{\mathrm{T}}[\mathbf{p}-\mathbf{KSp}]\Big) - 2\underbrace{\mathrm{E}\Big([\mathbf{p}-\mathbf{KSp}]^{\mathrm{T}}\mathbf{Kv}\Big)}_{=0} + \underbrace{\mathrm{E}\Big(\mathbf{v}^{\mathrm{T}}\mathbf{K}^{\mathrm{T}}\mathbf{Kv}\Big)}_{=\mathrm{spur}(\mathrm{E}(\mathbf{Kvv}^{\mathrm{T}}\mathbf{K}^{\mathrm{T}}))} \right\} = \\
&\min_{\mathbf{K}} \Big\{ \|\mathbf{p}-\mathbf{KSp}\|_2^2 + \mathrm{spur}\Big(\mathbf{KQK}^{\mathrm{T}}\Big) \Big\} \,. \qquad (2.15)
\end{aligned}$$

Die Matrix $\mathbf{K}$, die den Ausdruck (2.15) minimiert, ist offensichtlich eine Funktion des unbekannten Parametervektors $\mathbf{p}$. Daher ist die Lösung dieser Minimierungsaufgabe auch nicht geeignet, einen Schätzer für $\mathbf{p}$ gemäß (2.13) zu liefern. Um dieses Problem zu umgehen, wird im Folgenden eine Ersatzaufgabe gelöst: Man erkennt, dass mit der Wahl

$$\mathbf{KS} = \mathbf{E} \qquad (2.16)$$

und $\mathbf{E}$ als Einheitsmatrix die Lösung der Minimierungsaufgabe (2.15) unabhängig vom Parameter $\mathbf{p}$ ist. Diese *Nebenbedingung* (2.16) scheint zwar im ersten Augenblick willkürlich zu sein, doch berechnet man den Erwartungswert von $\hat{\mathbf{p}}$ des linearen Schätzers (2.13), dann folgt

$$\mathrm{E}(\hat{\mathbf{p}}) = \mathrm{E}(\mathbf{Ky}) = \mathrm{E}(\mathbf{KSp}+\mathbf{Kv}) = \mathbf{KS}\underbrace{\mathrm{E}(\mathbf{p})}_{=\mathbf{p}} + \mathbf{K}\underbrace{\mathrm{E}(\mathbf{v})}_{=\mathbf{0}} \,. \qquad (2.17)$$

D.h. die Nebenbedingung (2.16) bringt mit sich, dass $\mathrm{E}(\hat{\mathbf{p}}) = \mathbf{p}$ ist und somit der Schätzer *erwartungstreu* ist.

Die Aufgabe

$$\min_{\mathbf{K}} \Big\{ \mathrm{spur}\Big(\mathbf{KQK}^{\mathrm{T}}\Big) \Big\} \quad \text{unter der NB} \quad \mathbf{KS} = \mathbf{E} \qquad (2.18)$$

ist nun äquivalent zur Lösung n separater Optimierungsaufgaben der Form

$$\min_{\mathbf{k}_j} \mathbf{k}_j^{\mathrm{T}}\mathbf{Q}\mathbf{k}_j \quad \text{unter den NBen} \quad \mathbf{k}_j^{\mathrm{T}}\mathbf{S} = \mathbf{e}_j^{\mathrm{T}} \quad \text{für} \quad j = 1,\ldots,n \,. \qquad (2.19)$$

Um dies zu zeigen, schreibe man die Matrix $\mathbf{K}$ und die Einheitsmatrix $\mathbf{E}$ in der Form

$$\mathbf{K} = \begin{bmatrix} \mathbf{k}_1^{\mathrm{T}} \\ \mathbf{k}_2^{\mathrm{T}} \\ \vdots \\ \mathbf{k}_n^{\mathrm{T}} \end{bmatrix} \quad \text{und} \quad \mathbf{E} = \begin{bmatrix} \mathbf{e}_1^{\mathrm{T}} \\ \mathbf{e}_2^{\mathrm{T}} \\ \vdots \\ \mathbf{e}_n^{\mathrm{T}} \end{bmatrix} \qquad (2.20)$$

an und setzt dies in (2.18) ein

$$\min_{\mathbf{K}} \Big\{ \mathrm{spur}\Big(\mathbf{KQK}^{\mathrm{T}}\Big) \Big\} = \min_{\mathbf{K}} \left\{ \sum_{j=1}^{n} \mathbf{k}_j^{\mathrm{T}}\mathbf{Q}\mathbf{k}_j \right\} \quad \text{unter den NBen} \quad \mathbf{k}_j^{\mathrm{T}}\mathbf{S} = \mathbf{e}_j^{\mathrm{T}} \qquad (2.21)$$

für $j = 1, \ldots, n$. Da der j-te Summand in (2.21) lediglich von $\mathbf{k}_j$ abhängt, kann die Minimierungsaufgabe von (2.21) durch n Minimierungsprobleme gemäß (2.19) ersetzt werden. Für das Weitere ist also die Lösung einer Aufgabe vom Typ (2.19) für ein festes j von Interesse. Betrachtet man nun den Hilbertraum $\mathcal{H} = \mathbb{R}^m$ mit dem inneren Produkt $\langle \mathbf{x}, \mathbf{z} \rangle = \mathbf{x}^{\mathrm{T}}\mathbf{z} = \sum_{j=1}^{m} x_j z_j$, dann erkennt man, dass die Spaltenvektoren $\mathbf{k}_k$, $k = 1, \ldots, n$, der Matrix $\mathbf{K}$, also $\operatorname{span}\{\mathbf{k}_1, \mathbf{k}_2, \ldots, \mathbf{k}_n\}$, die die Nebenbedingungen von (2.19) erfüllen, keinen Unterraum des Hilbertraums $\mathcal{H}$ bilden und damit das Projektionstheorem von Satz 1.1 nicht direkt anwendbar ist.

Aufgabe 2.1. Zeigen Sie, dass die Summe $\mathbf{k}_j + \mathbf{k}_k$ die Nebenbedingung von (2.21) nicht erfüllt, selbst wenn $\mathbf{k}_j$ und $\mathbf{k}_k$ jeweils für sich dieser Nebenbedingung genügen.

2.1.1 Quadratische Minimierung mit affinen Nebenbedingungen

Zur Lösung der obigen Aufgabe wird eine Erweiterung des Projektionstheorems von Satz 1.1 benötigt:

Satz 2.1 (Erweiterung des Projektionstheorems)**.** *Es sei $\mathcal{H}$ ein Hilbertraum und $\mathcal{U}$ ein abgeschlossener Unterraum von $\mathcal{H}$. Die translatorische Verschiebung von $\mathcal{U}$ in der Form $\mathcal{A} = \mathbf{x} + \mathcal{U}$ für ein festes $\mathbf{x} \in \mathcal{H}$ wird als* lineare Varietät *oder* affiner Unterraum *bezeichnet. Dann existiert ein eindeutiger Vektor $\mathbf{x}_0 \in \mathcal{A}$ minimaler Norm und dieser ist orthogonal auf $\mathcal{U}$ (siehe Abbildung 2.5).*

Beweis von Satz 2.1. Man verschiebt den affinen Unterraum $\mathcal{A}$ durch $-\mathbf{x}$ so, dass er ein abgeschlossener Unterraum wird und wendet dann das Projektionstheorem von Satz 1.1 an. Man beachte, dass die optimale Lösung $\mathbf{x}_0$ *nicht* orthogonal auf den affinen Unterraum $\mathcal{A}$ sondern orthogonal auf $\mathcal{U}$ ist. □

Bevor nun die Aufgabe von (2.19) gelöst werden kann, sollen noch einige theoretische Grundlagen erläutert werden:

Definition 2.1 (Orthogonales Komplement)**.** Ist $\mathcal{U}$ ein Unterraum eines Hilbertraumes $\mathcal{H}$ mit dem inneren Produkt $\langle \cdot, \cdot \rangle$, dann bezeichnet man die Menge aller Vektoren, die orthogonal zu $\mathcal{U}$ sind, als das *orthogonale Komplement von $\mathcal{U}$* und schreibt dafür $\mathcal{U}^{\perp}$.

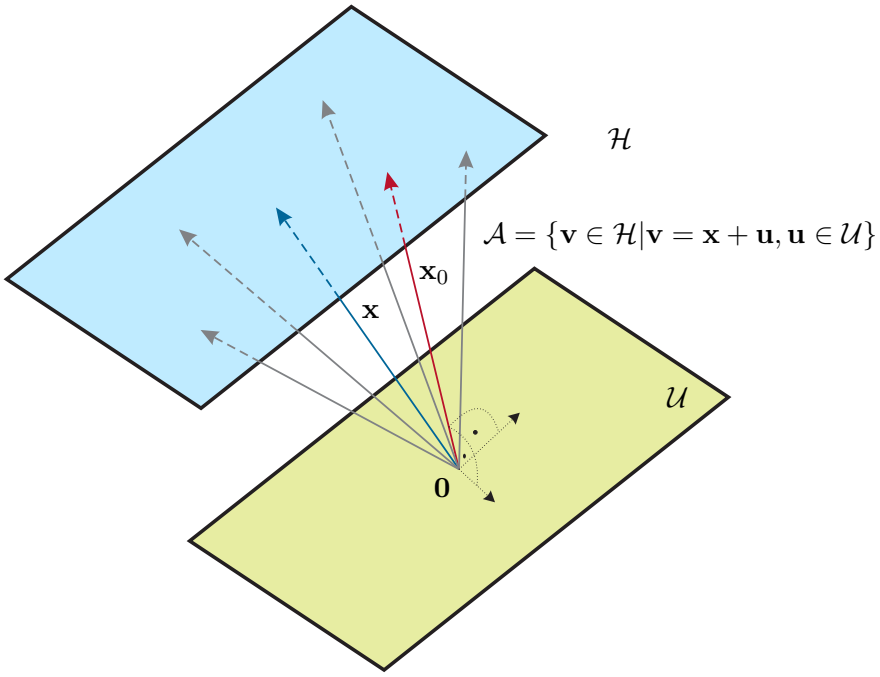


Abbildung 2.5: Zum Projektionstheorem für affine Unterräume.

Für die Unterräume $\mathcal{U}$ und $\mathcal{V}$ eines Hilbertraums gelten nun folgende Eigenschaften:

(1) Das orthogonale Komplement $\mathcal{U}^\perp$ ist ein abgeschlossener Unterraum,

(2) $\mathcal{U} \subseteq \mathcal{U}^{\perp\perp}$,

(3) wenn $\mathcal{U} \subset \mathcal{V}$ ist, dann folgt $\mathcal{V}^\perp \subset \mathcal{U}^\perp$,

(4) $\mathcal{U}^{\perp\perp\perp} = \mathcal{U}^\perp$ und

(5) $\mathcal{U}^{\perp\perp}$ ist der kleinste geschlossene Unterraum der $\mathcal{U}$ beinhaltet.

Beweis zu (1). Da die Linearkombination orthogonaler Vektoren wieder orthogonal ist, folgt unmittelbar, dass $\mathcal{U}^\perp$ ein Unterraum ist. Die Abgeschlossenheit von $\mathcal{U}^\perp$ folgt aus der Tatsache, dass wegen der Stetigkeit des inneren Produktes $\langle \cdot, \cdot \rangle$ für den Grenzwert $\mathbf{x}$ einer konvergenten Folge $(\mathbf{x}_k)$ in $\mathcal{U}^\perp$ gilt

$$\mathbf{0} = \langle \mathbf{y}, \mathbf{x}_k \rangle = \langle \mathbf{y}, \mathbf{x} \rangle \tag{2.22}$$

für alle $\mathbf{y} \in \mathcal{U}$ und damit auch $\mathbf{x} \in \mathcal{U}^\perp$. □

Aufgabe 2.2. Beweisen Sie die obigen Eigenschaften (2) bis (4).

Definition 2.2 (Direkte Summe)**.** Man bezeichnet einen Vektorraum $\mathcal{X}$ als direkte Summe zweier Unterräume $\mathcal{U}$ und $\mathcal{V}$ und schreibt dann $\mathcal{X} = \mathcal{U} \oplus \mathcal{V}$, wenn sich jeder Vektor $\mathbf{x} \in \mathcal{X}$ *eindeutig* als Summe $\mathbf{x} = \mathbf{u} + \mathbf{v}$ mit $\mathbf{u} \in \mathcal{U}$ und $\mathbf{v} \in \mathcal{V}$ darstellen lässt.

Ohne Beweis gilt nun als Folgerung des Projektionstheorems von Satz 1.1 nachfolgender Satz:

Satz 2.2. *Wenn $\mathcal{U}$ ein abgeschlossener linearer Unterraum eines Hilbertraums $\mathcal{H}$ ist, dann gilt $\mathcal{H} = \mathcal{U} \oplus \mathcal{U}^\perp$ und $\mathcal{U} = \mathcal{U}^{\perp\perp}$.*

Damit lässt sich nachfolgender Satz zur Lösung der Minimierungsaufgabe mit affinen Nebenbedingungen von (2.19) angeben:

Satz 2.3 (Minimierung mit affinen Nebenbedingung)**.** *Es sei $\mathcal{H}$ ein Hilbertraum mit den linear unabhängigen Vektoren $\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_n$. Unter allen möglichen Vektoren $\mathbf{x} \in \mathcal{H}$, die das affine Gleichungssystem*

$$\begin{aligned} \langle \mathbf{x}, \mathbf{s}_1 \rangle &= c_1 \\ \langle \mathbf{x}, \mathbf{s}_2 \rangle &= c_2 \\ \vdots\!\vdots\!\vdots &\quad \vdots \\ \langle \mathbf{x}, \mathbf{s}_n \rangle &= c_n \end{aligned} \tag{2.23}$$

mit den konstanten Koeffizienten $c_1, c_2, \ldots, c_n$ erfüllen, habe der Vektor $\mathbf{x}_0$ die minimale Norm. Dann lässt sich $\mathbf{x}_0$ in der Form

$$\mathbf{x}_0 = \sum_{j=1}^{n} p_{0,j} \mathbf{s}_j = \mathbf{S}\mathbf{p}_0 \tag{2.24}$$

anschreiben, wobei sich $\mathbf{p}_0^T = \begin{bmatrix} p_{0,1} & p_{0,2} & \ldots & p_{0,n} \end{bmatrix}$ aus der Beziehung

$$\underbrace{\begin{bmatrix} \langle \mathbf{s}_1, \mathbf{s}_1 \rangle & \cdots & \langle \mathbf{s}_n, \mathbf{s}_1 \rangle \\ \vdots & \ddots & \vdots \\ \langle \mathbf{s}_1, \mathbf{s}_n \rangle & \cdots & \langle \mathbf{s}_n, \mathbf{s}_n \rangle \end{bmatrix}}_{\mathbf{G}=\mathbf{S}^\mathrm{T}\mathbf{S}} \underbrace{\begin{bmatrix} p_{0,1} \\ \vdots \\ p_{0,n} \end{bmatrix}}_{\mathbf{p}_0} = \underbrace{\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}}_{\mathbf{c}} \tag{2.25}$$

mit der Gramschen Matrix $\mathbf{G}$ errechnet.

Beweis zu Satz 2.3. Es sei $\mathcal{U} = \operatorname{span}\{\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_n\}$ ein abgeschlossener linearer Unterraum des Hilbertraums $\mathcal{H}$. Die Menge aller möglichen Vektoren $\mathbf{x} \in \mathcal{H}$, die das affine Gleichungssystem (2.23) erfüllen, bilden einen affinen Unterraum von $\mathcal{H}$, nämlich die translatorische Verschiebung von $\mathcal{U}^\perp$ um einen Vektor $\boldsymbol{\gamma}$. Da das orthogonale Komplement $\mathcal{U}^\perp$ ein abgeschlossener Unterraum ist, kann man Satz 2.1 anwenden und weiß damit, dass die optimale Lösung $\mathbf{x}_0$ existiert und eindeutig ist sowie orthogonal auf $\mathcal{U}^\perp$ steht. Damit folgt aber $\mathbf{x}_0 \in \mathcal{U}^{\perp\perp}$ und wegen der Abgeschlossenheit von

$\mathcal{U}$ und Satz 2.2 erhält man $\mathcal{U}^{\perp\perp} = \mathcal{U}$. Da nun aber $\mathbf{x}_0 \in \mathcal{U}$ gilt, muss sich $\mathbf{x}_0$ als Linearkombination der $\mathbf{s}_j$, $j = 1, \ldots, n$, gemäß (2.24) darstellen lassen. Setzt man nun noch (2.24) in die affine Nebenbedingung (2.23) ein, erhält man das Ergebnis (2.25). □

Wendet man Satz 2.3 auf die Minimierungsaufgabe (2.19) mit $\mathbf{S} = \begin{bmatrix} \mathbf{s}_1 & \mathbf{s}_2 & \ldots & \mathbf{s}_n \end{bmatrix}$ und dem inneren Produkt $\langle \mathbf{x}, \mathbf{z} \rangle_Q = \mathbf{x}^{\mathrm{T}} \mathbf{Q} \mathbf{z}$ im Hilbertraum $\mathcal{H} = \mathbb{R}^m$ an, also

$$\min_{\mathbf{k}_j} \langle \mathbf{k}_j, \mathbf{k}_j \rangle_Q \quad \text{unter den NBen} \quad \left\langle \mathbf{k}_j, \mathbf{Q}^{-1} \mathbf{s}_l \right\rangle_Q = \delta_{jl} = \begin{cases} 1 & \text{für } j = l \\ 0 & \text{sonst} \end{cases} \tag{2.26}$$

für $j = 1, \ldots, n$, dann erhält man für $\mathbf{s}_l$ ersetzt durch $\mathbf{Q}^{-1}\mathbf{s}_l$ in (2.24) und (2.25) das Ergebnis

$$\mathbf{k}_{j,0} = \mathbf{Q}^{-1}\mathbf{S}\mathbf{p}_0 \quad \text{und} \quad \left(\mathbf{Q}^{-1}\mathbf{S}\right)^{\mathrm{T}} \mathbf{Q}\mathbf{Q}^{-1}\mathbf{S}\mathbf{p}_0 = \mathbf{e}_j \tag{2.27}$$

bzw.

$$\mathbf{k}_{j,0} = \mathbf{Q}^{-1}\mathbf{S}\left(\mathbf{S}^{\mathrm{T}}\mathbf{Q}^{-1}\mathbf{S}\right)^{-1}\mathbf{e}_j \ . \tag{2.28}$$

Nach (2.20) errechnet sich die optimale Lösung $\mathbf{K}_0$ für die Matrix $\mathbf{K}$ zu

$$\mathbf{K}_0^{\mathrm{T}} = \mathbf{Q}^{-1}\mathbf{S}\left(\mathbf{S}^{\mathrm{T}}\mathbf{Q}^{-1}\mathbf{S}\right)^{-1} \tag{2.29}$$

und damit lautet der gesuchte lineare *Gauß-Markov-Schätzer* nach (2.13)

$$\hat{\mathbf{p}} = \left(\mathbf{S}^{\mathrm{T}}\mathbf{Q}^{-1}\mathbf{S}\right)^{-1}\mathbf{S}^{\mathrm{T}}\mathbf{Q}^{-1}\mathbf{y} \ . \tag{2.30}$$

Vergleicht man nun (2.30) mit (1.128), dann erkennt man, dass das Ergebnis identisch zum Ergebnis der Least-Squares Identifikation mit gewichteten kleinsten Quadraten mit der Gewichtungsmatrix $\mathbf{Q}^{-1}$ ist, wobei gilt $\mathbf{Q} = \mathrm{cov}(\mathbf{v})$.

Für den Erwartungswert des Schätzfehlers $\mathbf{e} = \mathbf{p} - \hat{\mathbf{p}}$ erhält man gemäß (2.17) $\mathrm{E}(\mathbf{e}) = \mathbf{0}$ und die Kovarianzmatrix des Schätzfehlers lautet mit (2.16)

$$\begin{aligned} \mathrm{E}\left(\mathbf{e}\mathbf{e}^{\mathrm{T}}\right) &= \mathrm{E}\left([\mathbf{p} - \mathbf{K}\mathbf{y}][\mathbf{p} - \mathbf{K}\mathbf{y}]^{\mathrm{T}}\right) = \mathrm{E}\left([\mathbf{p} - \mathbf{K}(\mathbf{S}\mathbf{p} + \mathbf{v})][\mathbf{p} - \mathbf{K}(\mathbf{S}\mathbf{p} + \mathbf{v})]^{\mathrm{T}}\right) = \\ &= \mathrm{E}\left([\mathbf{K}\mathbf{v}][\mathbf{K}\mathbf{v}]^{\mathrm{T}}\right) = \mathbf{K}\mathbf{Q}\mathbf{K}^{\mathrm{T}} \end{aligned} \tag{2.31}$$

bzw. mit (2.29)

$$\mathrm{E}\left(\mathbf{e}\mathbf{e}^{\mathrm{T}}\right) = \left(\mathbf{S}^{\mathrm{T}}\mathbf{Q}^{-1}\mathbf{S}\right)^{-1}\mathbf{S}^{\mathrm{T}}\mathbf{Q}^{-1}\mathbf{Q}\mathbf{Q}^{-1}\mathbf{S}\left(\mathbf{S}^{\mathrm{T}}\mathbf{Q}^{-1}\mathbf{S}\right)^{-1} = \left(\mathbf{S}^{\mathrm{T}}\mathbf{Q}^{-1}\mathbf{S}\right)^{-1} \ . \tag{2.32}$$

In der Literatur wird der lineare Schätzer (2.30) oft auch als *BLUE (best linear unbiased estimate)* bezeichnet.

Ist die Kovarianzmatrix $\mathrm{cov}(\mathbf{v}) = \mathrm{E}\left(\mathbf{v}\mathbf{v}^{\mathrm{T}}\right) = \mathbf{Q}$ eine Diagonalmatrix der Form $\mathbf{Q} = \mathrm{diag}(q_0, q_1, \ldots)$ mit $q_j > 0$ für alle $j \geq 0$, dann kann für den Gauß-Markov-Schätzer (2.30) gemäß der rekursiven Methode der gewichteten kleinsten Quadrate (1.121) unmittelbar mit $\alpha_j = 1/q_j$ eine rekursive Version angegeben werden. Dies klärt auch die Frage, wie die optimale Wahl der Folge positiver Gewichtungskoeffizienten α_j ($\alpha_j > 0$ für alle j) in (1.121) aussieht.

2.2 Minimum-Varianz-Schätzung

Bisher wurde angenommen, dass über den n-dimensionalen Vektor der Unbekannten $\mathbf{p}$ in (2.11) keine Information vorhanden ist. Nun ist aber in manchen Fällen sehr wohl a priori Information über $\mathbf{p}$ in Form von stochastischen Kenngrößen (Erwartungswert, Kovarianzmatrix) bekannt. Daher wird angenommen, dass für das Gleichungssystem (vergleiche dazu auch (2.11))

$$\mathbf{y} = \mathbf{S}\mathbf{p} + \mathbf{v} \tag{2.33}$$

mit der stochastischen Störung $\mathbf{v}$, der bekannten $(m \times n)$-Matrix $\mathbf{S} \in \mathbb{R}^{m\times n}$, dem n-dimensionalen Zufallsvektor $\mathbf{p} \in \mathbb{R}^n$ sowie dem m-dimensionalen Messvektor $\mathbf{y} \in \mathbb{R}^m$ gilt

$$\begin{aligned} \mathrm{E}(\mathbf{v}) &= \mathbf{0}, \qquad \mathrm{cov}(\mathbf{v}) = \mathrm{E}\Big(\mathbf{v}\mathbf{v}^{\mathrm{T}}\Big) = \mathbf{Q} \qquad \text{mit} \quad \mathbf{Q} \geq 0 \\ \mathrm{E}(\mathbf{p}) &= \mathbf{0}, \qquad \mathrm{cov}(\mathbf{p}) = \mathrm{E}\Big(\mathbf{p}\mathbf{p}^{\mathrm{T}}\Big) = \mathbf{R} \qquad \text{mit} \quad \mathbf{R} \geq 0 \\ &\qquad\qquad\quad\; \mathrm{E}\Big(\mathbf{p}\mathbf{v}^{\mathrm{T}}\Big) = \mathbf{N}\ . \end{aligned} \tag{2.34}$$

Im Weiteren sei vorausgesetzt, dass die Matrix $\Big(\mathbf{S}\mathbf{R}\mathbf{S}^{\mathrm{T}} + \mathbf{Q} + \mathbf{S}\mathbf{N} + \mathbf{N}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}}\Big)$ regulär ist. Gesucht ist nun wiederum ein linearer Schätzer

$$\hat{\mathbf{p}} = \mathbf{K}\mathbf{y} \tag{2.35}$$

mit einer konstanten $(n \times m)$-Matrix $\mathbf{K} \in \mathbb{R}^{n\times m}$ so, dass nachfolgende Minimierungsaufgabe

$$\min_{\mathbf{K}} \mathrm{E}\Big(\|\mathbf{e}\|_2^2\Big) = \min_{\mathbf{K}} \mathrm{E}\Big([\mathbf{p} - \mathbf{K}\mathbf{y}]^{\mathrm{T}}[\mathbf{p} - \mathbf{K}\mathbf{y}]\Big) \tag{2.36}$$

gelöst wird. Durch Ausmultiplizieren von (2.36) und mit Hilfe der Beziehung $\mathrm{spur}(\mathbf{K}\mathbf{S}\mathbf{R}) = \mathrm{spur}\Big(\mathbf{R}(\mathbf{K}\mathbf{S})^{\mathrm{T}}\Big)$ erhält man

$$\begin{aligned} &\min_{\mathbf{K}} \mathrm{E}\Big([\mathbf{p} - \mathbf{K}\mathbf{y}]^{\mathrm{T}}[\mathbf{p} - \mathbf{K}\mathbf{y}]\Big) = \\ &\min_{\mathbf{K}} \left\{ \underbrace{\mathrm{E}\Big([\mathbf{p} - \mathbf{K}\mathbf{S}\mathbf{p}]^{\mathrm{T}}[\mathbf{p} - \mathbf{K}\mathbf{S}\mathbf{p}]\Big)}_{\mathrm{spur}\left(\mathrm{E}\left([\mathbf{E}-\mathbf{K}\mathbf{S}]\mathbf{p}\mathbf{p}^{\mathrm{T}}[\mathbf{E}-\mathbf{K}\mathbf{S}]^{\mathrm{T}}\right)\right)} - 2\underbrace{\mathrm{E}\Big([\mathbf{p} - \mathbf{K}\mathbf{S}\mathbf{p}]^{\mathrm{T}}\mathbf{K}\mathbf{v}\Big)}_{\mathrm{spur}(\mathrm{E}([\mathbf{E}-\mathbf{K}\mathbf{S}]\mathbf{p}\mathbf{v}^{\mathrm{T}}\mathbf{K}^{\mathrm{T}}))} + \underbrace{\mathrm{E}\Big(\mathbf{v}^{\mathrm{T}}\mathbf{K}^{\mathrm{T}}\mathbf{K}\mathbf{v}\Big)}_{\mathrm{spur}(\mathrm{E}(\mathbf{K}\mathbf{v}\mathbf{v}^{\mathrm{T}}\mathbf{K}^{\mathrm{T}}))} \right\} = \\ &\min_{\mathbf{K}} \Big\{ \mathrm{spur}\Big([\mathbf{E} - \mathbf{K}\mathbf{S}]\mathbf{R}[\mathbf{E} - \mathbf{K}\mathbf{S}]^{\mathrm{T}} - 2\mathbf{K}\mathbf{N}^{\mathrm{T}} - \mathbf{K}\Big[\mathbf{S}\mathbf{N} + \mathbf{N}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}}\Big]\mathbf{K}^{\mathrm{T}} + \mathbf{K}\mathbf{Q}\mathbf{K}^{\mathrm{T}}\Big) \Big\} = \\ &\min_{\mathbf{K}} \Big\{ \mathrm{spur}\Big(\mathbf{K}\Big(\mathbf{S}\mathbf{R}\mathbf{S}^{\mathrm{T}} + \mathbf{Q} + \mathbf{S}\mathbf{N} + \mathbf{N}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}}\Big)\mathbf{K}^{\mathrm{T}} - 2\mathbf{K}\Big(\mathbf{S}\mathbf{R} + \mathbf{N}^{\mathrm{T}}\Big)\Big)\Big\}\ . \end{aligned} \tag{2.37}$$

Schreibt man die Matrix $\mathbf{K}$ und die Einheitsmatrix $\mathbf{E}$ wie in (2.20)

$$\mathbf{K} = \begin{bmatrix} \mathbf{k}_1^{\mathrm{T}} \\ \mathbf{k}_2^{\mathrm{T}} \\ \vdots \\ \mathbf{k}_n^{\mathrm{T}} \end{bmatrix} \quad \text{und} \quad \mathbf{E} = \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \cdots & \mathbf{e}_n \end{bmatrix} \tag{2.38}$$

an, dann folgt (2.37) zu

$$\min_{\mathbf{K}} \left\{ \sum_{j=1}^{n} \left(\mathbf{k}_j^{\mathrm{T}} \left(\mathbf{S}\mathbf{R}\mathbf{S}^{\mathrm{T}} + \mathbf{Q} + \mathbf{S}\mathbf{N} + \mathbf{N}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}} \right) \mathbf{k}_j - 2\mathbf{k}_j^{\mathrm{T}} \left(\mathbf{S}\mathbf{R} + \mathbf{N}^{\mathrm{T}} \right) \mathbf{e}_j \right) \right\} . \tag{2.39}$$

Vergleicht man die Minimierungsaufgabe (2.39) mit der von (1.61), also

$$\min_{\mathbf{p}} (\mathbf{y} - \mathbf{S}\mathbf{p})^{\mathrm{T}} (\mathbf{y} - \mathbf{S}\mathbf{p}) = \min_{\mathbf{p}} \left(\mathbf{y}^{\mathrm{T}}\mathbf{y} - 2\mathbf{p}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}}\mathbf{y} + \mathbf{p}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}}\mathbf{S}\mathbf{p} \right), \tag{2.40}$$

dann erkennt man, dass die beiden Aufgaben äquivalent sind. Es lässt sich also die Lösung von (2.39) direkt durch die Lösung von (2.40) (vergleiche dazu (1.63))

$$\mathbf{p}_0 = \left(\mathbf{S}^{\mathrm{T}}\mathbf{S} \right)^{-1} \mathbf{S}^{\mathrm{T}}\mathbf{y} \tag{2.41}$$

angeben, indem man in (2.41) $\mathbf{S}^{\mathrm{T}}\mathbf{S} = \mathbf{S}\mathbf{R}\mathbf{S}^{\mathrm{T}} + \mathbf{Q} + \mathbf{S}\mathbf{N} + \mathbf{N}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}}$ und $\mathbf{S}^{\mathrm{T}}\mathbf{y} = \left(\mathbf{S}\mathbf{R} + \mathbf{N}^{\mathrm{T}} \right) \mathbf{e}_j$ setzt. Damit ergibt sich die optimale Lösung $\mathbf{K}_0$ für die Matrix $\mathbf{K}$ von (2.35) zu

$$\mathbf{K}_0^{\mathrm{T}} = \left(\mathbf{S}\mathbf{R}\mathbf{S}^{\mathrm{T}} + \mathbf{Q} + \mathbf{S}\mathbf{N} + \mathbf{N}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}} \right)^{-1} \left(\mathbf{S}\mathbf{R} + \mathbf{N}^{\mathrm{T}} \right) \tag{2.42}$$

und der gesuchte lineare *Minimum-Varianz-Schätzer* nach (2.35) lautet

$$\hat{\mathbf{p}} = \left(\mathbf{R}\mathbf{S}^{\mathrm{T}} + \mathbf{N} \right) \left(\mathbf{S}\mathbf{R}\mathbf{S}^{\mathrm{T}} + \mathbf{Q} + \mathbf{S}\mathbf{N} + \mathbf{N}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}} \right)^{-1} \mathbf{y} \, . \tag{2.43}$$

Man beachte, dass am Beginn dieses Abschnittes vorausgesetzt wurde, dass die Matrix $\left(\mathbf{S}\mathbf{R}\mathbf{S}^{\mathrm{T}} + \mathbf{Q} + \mathbf{S}\mathbf{N} + \mathbf{N}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}} \right)$ regulär ist. Die Kovarianzmatrix des Schätzfehlers errechnet

sich zu

$$\begin{aligned}
\operatorname{cov}(\mathbf{e}) &= \mathrm{E}\Big([\mathbf{p}-\mathbf{K}(\mathbf{S}\mathbf{p}+\mathbf{v})][\mathbf{p}-\mathbf{K}(\mathbf{S}\mathbf{p}+\mathbf{v})]^{\mathrm{T}}\Big) = \mathrm{E}\Big([\mathbf{E}-\mathbf{K}\mathbf{S}]\mathbf{p}\mathbf{p}^{\mathrm{T}}[\mathbf{E}-\mathbf{K}\mathbf{S}]^{\mathrm{T}}\Big) - \\
&\quad \mathrm{E}\Big(\mathbf{K}\mathbf{v}\mathbf{p}^{\mathrm{T}}\Big(\mathbf{E}-\mathbf{S}^{\mathrm{T}}\mathbf{K}^{\mathrm{T}}\Big) + (\mathbf{E}-\mathbf{K}\mathbf{S})\mathbf{p}\mathbf{v}^{\mathrm{T}}\mathbf{K}^{\mathrm{T}}\Big) + \mathrm{E}\Big(\mathbf{K}\mathbf{v}\mathbf{v}^{\mathrm{T}}\mathbf{K}^{\mathrm{T}}\Big) \\
&= (\mathbf{E}-\mathbf{K}\mathbf{S})\mathbf{R}(\mathbf{E}-\mathbf{K}\mathbf{S})^{\mathrm{T}} - \mathbf{K}\mathbf{N}^{\mathrm{T}}\Big(\mathbf{E}-\mathbf{S}^{\mathrm{T}}\mathbf{K}^{\mathrm{T}}\Big) - (\mathbf{E}-\mathbf{K}\mathbf{S})\mathbf{N}\mathbf{K}^{\mathrm{T}} + \mathbf{K}\mathbf{Q}\mathbf{K}^{\mathrm{T}} \\
&= \mathbf{R} - \mathbf{K}\Big(\mathbf{S}\mathbf{R}+\mathbf{N}^{\mathrm{T}}\Big) - \Big(\mathbf{R}\mathbf{S}^{\mathrm{T}}+\mathbf{N}\Big)\mathbf{K}^{\mathrm{T}} + \mathbf{K}\Big(\mathbf{S}\mathbf{R}\mathbf{S}^{\mathrm{T}}+\mathbf{Q}+\mathbf{S}\mathbf{N}+\mathbf{N}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}}\Big)\mathbf{K}^{\mathrm{T}} \\
&= \mathbf{R} - \Big(\mathbf{R}\mathbf{S}^{\mathrm{T}}+\mathbf{N}\Big)\Big(\mathbf{S}\mathbf{R}\mathbf{S}^{\mathrm{T}}+\mathbf{Q}+\mathbf{S}\mathbf{N}+\mathbf{N}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}}\Big)^{-1}\Big(\mathbf{S}\mathbf{R}+\mathbf{N}^{\mathrm{T}}\Big) - \\
&\quad \Big(\mathbf{R}\mathbf{S}^{\mathrm{T}}+\mathbf{N}\Big)\Big(\mathbf{S}\mathbf{R}\mathbf{S}^{\mathrm{T}}+\mathbf{Q}+\mathbf{S}\mathbf{N}+\mathbf{N}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}}\Big)^{-1}\Big(\mathbf{S}\mathbf{R}+\mathbf{N}^{\mathrm{T}}\Big) + \Big(\mathbf{R}\mathbf{S}^{\mathrm{T}}+\mathbf{N}\Big) \\
&\quad \Big(\mathbf{S}\mathbf{R}\mathbf{S}^{\mathrm{T}}+\mathbf{Q}+\mathbf{S}\mathbf{N}+\mathbf{N}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}}\Big)^{-1}\Big(\mathbf{S}\mathbf{R}\mathbf{S}^{\mathrm{T}}+\mathbf{Q}+\mathbf{S}\mathbf{N}+\mathbf{N}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}}\Big) \\
&\quad \Big(\mathbf{S}\mathbf{R}\mathbf{S}^{\mathrm{T}}+\mathbf{Q}+\mathbf{S}\mathbf{N}+\mathbf{N}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}}\Big)^{-1}\Big(\mathbf{S}\mathbf{R}+\mathbf{N}^{\mathrm{T}}\Big) \\
&= \mathbf{R} - \Big(\mathbf{R}\mathbf{S}^{\mathrm{T}}+\mathbf{N}\Big)\Big(\mathbf{S}\mathbf{R}\mathbf{S}^{\mathrm{T}}+\mathbf{Q}+\mathbf{S}\mathbf{N}+\mathbf{N}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}}\Big)^{-1}\Big(\mathbf{S}\mathbf{R}+\mathbf{N}^{\mathrm{T}}\Big)\,.
\end{aligned} \tag{2.44}$$

Aufgabe 2.3. Wenn Sie einmal Zeit und Muße haben, rechnen Sie (2.44) nach.

Im Gegensatz zum Gauß-Markov-Schätzer (2.30) liefert der Minimum-Varianz-Schätzer (2.43) auch dann bereits sinnvolle Ergebnisse, wenn weniger Messungen m als Unbekannte n zur Verfügung stehen, sofern die Matrix $\mathbf{S}\mathbf{R}\mathbf{S}^{\mathrm{T}}+\mathbf{Q}+\mathbf{S}\mathbf{N}+\mathbf{N}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}}$ regulär ist. Der Grund dieser Eigenschaft liegt offensichtlich darin, dass beim Minimum-Varianz-Schätzer stochastische Informationen über den Parametervektor $\mathbf{p}$ vorliegen und somit eine sinnvolle Schätzung auch mit weniger Messungen, ja sogar mit keiner Messung für $\mathbf{Q} > 0$, möglich ist.

Wendet man nun das Matrizeninversionslemma Satz 1.3

$$(\mathbf{A}+\mathbf{B}\mathbf{C}\mathbf{D})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}\Big(\mathbf{C}^{-1}+\mathbf{D}\mathbf{A}^{-1}\mathbf{B}\Big)^{-1}\mathbf{D}\mathbf{A}^{-1} \tag{2.45}$$

auf die Kovarianzmatrix des Schätzfehlers $\operatorname{cov}(\mathbf{e})$ von (2.44) mit $\mathbf{A} = \mathbf{R}^{-1}$, $\mathbf{B} = \mathbf{S}^{\mathrm{T}}+\mathbf{R}^{-1}\mathbf{N}$, $\mathbf{C} = \Big(\mathbf{Q}-\mathbf{N}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{N}\Big)^{-1}$ und $\mathbf{D} = \mathbf{S}+\mathbf{N}^{\mathrm{T}}\mathbf{R}^{-1}$ an, dann erhält man

$$\begin{aligned}
\operatorname{cov}(\mathbf{e}) &= \mathbf{R} - \Big(\mathbf{R}\mathbf{S}^{\mathrm{T}}+\mathbf{N}\Big)\Big(\mathbf{S}\mathbf{R}\mathbf{S}^{\mathrm{T}}+\mathbf{Q}+\mathbf{S}\mathbf{N}+\mathbf{N}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}}\Big)^{-1}\Big(\mathbf{S}\mathbf{R}+\mathbf{N}^{\mathrm{T}}\Big) \\
&= \Big(\mathbf{R}^{-1} + \Big(\mathbf{S}^{\mathrm{T}}+\mathbf{R}^{-1}\mathbf{N}\Big)\Big(\mathbf{Q}-\mathbf{N}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{N}\Big)^{-1}\Big(\mathbf{S}+\mathbf{N}^{\mathrm{T}}\mathbf{R}^{-1}\Big)\Big)^{-1}\,.
\end{aligned} \tag{2.46}$$

Im Weiteren kann man sich einfach davon überzeugen, dass sich der Minimum-Varianz-Schätzer (2.43) in der Form

$$\hat{\mathbf{p}} = \operatorname{cov}(\mathbf{e})\Big(\mathbf{S}^{\mathrm{T}}+\mathbf{R}^{-1}\mathbf{N}\Big)\Big(\mathbf{Q}-\mathbf{N}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{N}\Big)^{-1}\mathbf{y} \tag{2.47}$$

schreiben lässt.

Aufgabe 2.4. Zeigen Sie die Gültigkeit von (2.47).

Beziehung (2.47) zeigt nun, dass für $\mathbf{R}^{-1} = \mathbf{0}$, d. h. unendlich hohe Varianz des Parametervektors $\mathbf{p}$ – es ist also keine sinnvolle a priori Information für $\mathbf{p}$ vorhanden – und $\mathbf{Q} > 0$ der Minimum-Varianz-Schätzer (2.43) bzw. (2.47) in den Gauß-Markov-Schätzer (2.30) übergeht.

Das bisher Gesagte, insbesondere die Ergebnisse (2.43) und (2.44), in Kombination mit den Beziehungen

$$\mathrm{E}\Big(\mathbf{p}\mathbf{y}^{\mathrm{T}}\Big) = \mathrm{E}\Big(\mathbf{p}\mathbf{p}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}} + \mathbf{p}\mathbf{v}^{\mathrm{T}}\Big) = \Big(\mathbf{R}\mathbf{S}^{\mathrm{T}} + \mathbf{N}\Big) \tag{2.48}$$

und

$$\mathrm{E}\Big(\mathbf{y}\mathbf{y}^{\mathrm{T}}\Big) = \mathrm{E}\Big([\mathbf{S}\mathbf{p} + \mathbf{v}][\mathbf{S}\mathbf{p} + \mathbf{v}]^{\mathrm{T}}\Big) = \Big(\mathbf{S}\mathbf{R}\mathbf{S}^{\mathrm{T}} + \mathbf{Q} + \mathbf{S}\mathbf{N} + \mathbf{N}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}}\Big) \tag{2.49}$$

lässt sich im nachfolgenden Satz zusammenfassen:

Satz 2.4 (Minimum-Varianz-Schätzer)**.** *Für das Gleichungssystem* (2.33)

$$\mathbf{y} = \mathbf{S}\mathbf{p} + \mathbf{v} \tag{2.50}$$

mit den stochastischen Größen $\mathbf{p}$, $\mathbf{v}$ *und* $\mathbf{y}$ *wird angenommen, dass* $\mathrm{E}\Big(\mathbf{y}\mathbf{y}^{\mathrm{T}}\Big)$ *invertierbar ist. Die optimale lineare Schätzung* $\hat{\mathbf{p}}$ *von* $\mathbf{p}$*, die den Erwartungswert des quadratischen Fehlers* $\mathrm{E}\Big([\mathbf{p} - \hat{\mathbf{p}}]^{\mathrm{T}}[\mathbf{p} - \hat{\mathbf{p}}]\Big)$ *minimiert, ist durch*

$$\hat{\mathbf{p}} = \mathrm{E}\Big(\mathbf{p}\mathbf{y}^{\mathrm{T}}\Big)\Big[\mathrm{E}\Big(\mathbf{y}\mathbf{y}^{\mathrm{T}}\Big)\Big]^{-1}\mathbf{y} \tag{2.51}$$

mit der zugehörigen Fehlerkovarianzmatrix

$$\begin{aligned}\mathrm{cov}(\mathbf{e}) &= \mathrm{E}\Big([\mathbf{p} - \hat{\mathbf{p}}][\mathbf{p} - \hat{\mathbf{p}}]^{\mathrm{T}}\Big) = \mathrm{E}\Big(\mathbf{p}\mathbf{p}^{\mathrm{T}}\Big) - \mathrm{E}\Big(\hat{\mathbf{p}}\hat{\mathbf{p}}^{\mathrm{T}}\Big) \\ &= \mathrm{E}\Big(\mathbf{p}\mathbf{p}^{\mathrm{T}}\Big) - \mathrm{E}\Big(\mathbf{p}\mathbf{y}^{\mathrm{T}}\Big)\Big[\mathrm{E}\Big(\mathbf{y}\mathbf{y}^{\mathrm{T}}\Big)\Big]^{-1}\mathrm{E}\Big(\mathbf{y}\mathbf{p}^{\mathrm{T}}\Big)\end{aligned} \tag{2.52}$$

gegeben.

Man beachte die Ähnlichkeit von (2.51) mit der optimalen Lösung im Sinne der kleinsten Fehlerquadrate von (1.63).

Aufgabe 2.5. Zeigen Sie die Gültigkeit der Beziehung (2.52). Zeigen Sie weiters, dass gilt

$$\mathrm{E}\Big([\mathbf{p} - \hat{\mathbf{p}}][\mathbf{p} - \hat{\mathbf{p}}]^{\mathrm{T}}\Big) = \mathrm{E}\Big(\mathbf{p}[\mathbf{p} - \hat{\mathbf{p}}]^{\mathrm{T}}\Big) \quad \text{bzw.} \quad \mathrm{E}\Big(\hat{\mathbf{p}}\hat{\mathbf{p}}^{\mathrm{T}}\Big) = \mathrm{E}\Big(\mathbf{p}\hat{\mathbf{p}}^{\mathrm{T}}\Big)\ . \tag{2.53}$$

Hinweis: (zu Aufgabe 2.5) Setzen Sie einfach für $\hat{\mathbf{p}}$ den Ausdruck von (2.51) ein.

Aufgabe 2.6. Zeigen Sie, dass sich die Beziehungen (2.43) und (2.44) direkt mit Hilfe von Satz 2.4 herleiten lassen.

Aufgabe 2.7. Angenommen, die Erwartungswerte $\mathrm{E}(\mathbf{y})$ und $\mathrm{E}(\mathbf{p})$ sind nicht wie in (2.34) Null, sondern $\mathrm{E}(\mathbf{y}) = \mathbf{y}_0 \neq \mathbf{0}$ und $\mathrm{E}(\mathbf{p}) = \mathbf{p}_0 \neq \mathbf{0}$. Zeigen Sie, dass die Minimum-Varianz-Schätzung der Form

$$\hat{\mathbf{p}} = \mathbf{K}\mathbf{y} + \mathbf{b}$$

mit dem konstanten Vektor $\mathbf{b}$ durch

$$\hat{\mathbf{p}} = \mathbf{p}_0 + \mathrm{E}\Big([\mathbf{p} - \mathbf{p}_0][\mathbf{y} - \mathbf{y}_0]^{\mathrm{T}}\Big)\Big[\mathrm{E}\Big([\mathbf{y} - \mathbf{y}_0][\mathbf{y} - \mathbf{y}_0]^{\mathrm{T}}\Big)\Big]^{-1}(\mathbf{y} - \mathbf{y}_0)$$

gegeben ist.

Unter der *Minimum-Varianz-Schätzung einer linearen Funktion*

$$\mathbf{z} = \mathbf{C}\mathbf{p} \tag{2.54}$$

mit dem optimalen Schätzer

$$\hat{\mathbf{z}} = \mathbf{K}_{\mathbf{z}}\mathbf{y} \tag{2.55}$$

basierend auf den Messungen

$$\mathbf{y} = \mathbf{S}\mathbf{p} + \mathbf{v} \tag{2.56}$$

versteht man die Lösung der Minimierungsaufgabe

$$\min_{\mathbf{K}_{\mathbf{z}}} \mathrm{E}\Big([\mathbf{z} - \hat{\mathbf{z}}]^{\mathrm{T}}[\mathbf{z} - \hat{\mathbf{z}}]\Big) \ . \tag{2.57}$$

Es gilt nun folgender Satz:

Satz 2.5 (Minimum-Varianz-Schätzer einer linearen Funktion)**.** *Die lineare Minimum-Varianz-Schätzung* (2.55) *einer linearen Funktion* $\mathbf{C}\mathbf{p}$ *basierend auf den Messungen* (2.56) *ist äquivalent der linearen Funktion der Minimum-Varianz-Schätzung* $\hat{\mathbf{p}}$ *selbst, d. h. es gilt, die beste Schätzung von* $\mathbf{C}\mathbf{p}$ *ist* $\mathbf{C}\hat{\mathbf{p}}$*.*

Aufgabe 2.8. Beweisen Sie Satz 2.5.

2.2.1 Rekursive Minimum-Varianz-Schätzung

Im nächsten Schritt soll untersucht werden, wie sich der optimale Schätzwert $\hat{\mathbf{p}}$ von (2.47) durch Hinzunahme neuer Messungen verbessert. Dies ist insbesondere für On-line-Anwendungen von essentieller Bedeutung. Das Verfahren beruht nun wiederum auf den Eigenschaften des Projektionstheorems in einem Hilbertraum. Wenn $\mathcal{U}_1$ und $\mathcal{U}_2$ zwei Unterräume eines Hilbertraums bezeichnen, dann ist die Projektion eines Vektors $\mathbf{p}$ auf den Unterraum $\mathcal{U}_1 + \mathcal{U}_2$ identisch der Projektion von $\mathbf{p}$ auf $\mathcal{U}_1$ plus der Projektion auf $\mathcal{U}_2^*$, wobei $\mathcal{U}_2^*$ orthogonal zu $\mathcal{U}_1$ ist und die Beziehung $\mathcal{U}_1 \oplus \mathcal{U}_2^* = \mathcal{U}_1 + \mathcal{U}_2$ erfüllt. Wenn $\mathcal{U}_2$ durch eine endliche Anzahl von Vektoren aufgebaut ist, dann spannen die Differenzen dieser Vektoren mit ihren Projektionen auf $\mathcal{U}_1$ den Unterraum $\mathcal{U}_2^*$ auf. Die Abbildung 2.6 veranschaulicht diesen Sachverhalt.

Damit kann man folgenden Satz angeben:

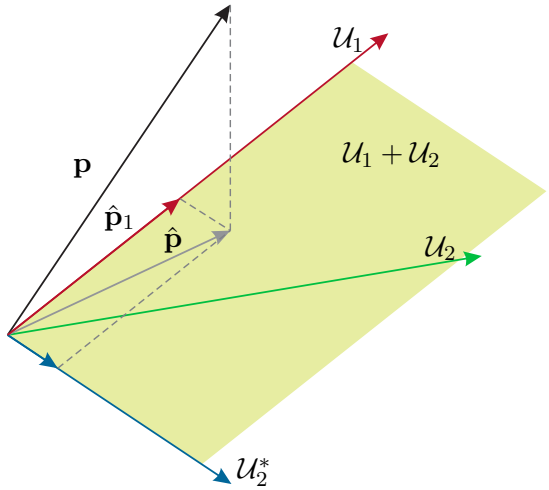


Abbildung 2.6: Zur Projektion auf die Summe orthogonale Unterräume.

Satz 2.6 (Rekursive Minimum-Varianz-Schätzung)**.** *Es sei* $\mathbf{p}$ *ein Zufallsvektor eines Hilbertraumes* $\mathcal{H}$ *von Zufallsvariablen und* $\hat{\mathbf{p}}_1$ *bezeichne die orthogonale Projektion von* $\mathbf{p}$ *auf einen geschlossenen Unterraum* $\mathcal{U}_1$ *von* $\mathcal{H}$*. Nach dem Projektionstheorem ist* $\hat{\mathbf{p}}_1$ *also die beste Schätzung von* $\mathbf{p}$ *in* $\mathcal{U}_1$*. Weiters beschreibe* $\mathbf{y}_2$ *alle jene Zufallsvektoren, die den Unterraum* $\mathcal{U}_2$ *von* $\mathcal{H}$ *aufspannen, und* $\hat{\mathbf{y}}_2$ *sei die orthogonale Projektion von* $\mathbf{y}_2$ *auf* $\mathcal{U}_1$*. Nach dem Projektionstheorem ist* $\hat{\mathbf{y}}_2$ *damit die beste Schätzung von* $\mathbf{y}_2$ *in* $\mathcal{U}_1$*. Mit* $\tilde{\mathbf{y}}_2 = \mathbf{y}_2 - \hat{\mathbf{y}}_2$ *lautet die Projektion* $\hat{\mathbf{p}}$ *von* $\mathbf{p}$ *auf* $\mathcal{U}_1 + \mathcal{U}_2$

$$\hat{\mathbf{p}} = \hat{\mathbf{p}}_1 + \mathrm{E}\left(\mathbf{p}\tilde{\mathbf{y}}_2^{\mathrm{T}}\right)\left[\mathrm{E}\left(\tilde{\mathbf{y}}_2\tilde{\mathbf{y}}_2^{\mathrm{T}}\right)\right]^{-1}\tilde{\mathbf{y}}_2 \ . \tag{2.58}$$

Damit setzt sich also die beste Schätzung $\hat{\mathbf{p}}$ *auf* $\mathcal{U}_1 + \mathcal{U}_2$ *aus der Summe der besten Schätzung von* $\mathbf{p}$ *auf* $\mathcal{U}_1$ *(*$\hat{\mathbf{p}}_1$*) und der besten Schätzung von* $\mathbf{p}$ *auf* $\mathcal{U}_2^*$ *(jener Unterraum, der durch* $\tilde{\mathbf{y}}_2$ *generiert wird) zusammen.*

Beweis zu Satz 2.6. Man überzeugt sich leicht, dass gilt $\mathcal{U}_1 + \mathcal{U}_2 = \mathcal{U}_1 \oplus \mathcal{U}_2^*$ und dass $\mathcal{U}_2^*$ orthogonal auf $\mathcal{U}_1$ ist. Die Beziehung (2.58) folgt dann aus der Tatsache, dass die Projektion auf eine Summe von Unterräumen gleich der Summe der Projektionen auf die einzelnen Unterräume ist, sofern diese orthogonal sind. □

Das Ergebnis von Satz 2.6 kann nun auch in folgender Form gedeutet werden: Wenn $\hat{\mathbf{p}}_1$ die optimale Schätzung auf Basis von Messungen, die den Unterraum $\mathcal{U}_1$ aufspannen, angibt, dann braucht man beim Erhalt neuer Messungen, die den Unterraum $\mathcal{U}_2$ aufspannen, nur jenen Teil zu berücksichtigen, der durch die Messungen in $\mathcal{U}_1$ noch nicht beschrieben wird, also jenen Teil der neuen Daten, der orthogonal auf die alten Daten steht und damit im Unterraum $\mathcal{U}_2^*$ liegt.

Als Anwendungsbeispiel betrachte man ein Gleichungssystem der Form von (2.33)

$$\mathbf{y}_1 = \mathbf{S}_1\mathbf{p} + \mathbf{v}_1 \ . \tag{2.59}$$

Im Weiteren bezeichnet $\hat{\mathbf{p}}_1 = \mathrm{E}\Big(\mathbf{p}\mathbf{y}_1^{\mathrm{T}}\Big)\Big[\mathrm{E}\Big(\mathbf{y}_1\mathbf{y}_1^{\mathrm{T}}\Big)\Big]^{-1}\mathbf{y}_1$ die optimale Minimum-Varianz-Schätzung von $\mathbf{p}$ gemäß (2.47) bzw. (2.51) auf Basis von $\dim(\mathbf{y}_1)$ Messungen mit der Fehlerkovarianzmatrix

$$\mathrm{cov}(\mathbf{p}-\hat{\mathbf{p}}_1) = \mathrm{E}\Big([\mathbf{p}-\hat{\mathbf{p}}_1][\mathbf{p}-\hat{\mathbf{p}}_1]^{\mathrm{T}}\Big) = \mathbf{P}_1 \ . \tag{2.60}$$

Es stellt sich nun die Frage, wie man die Schätzung von $\mathbf{p}$ durch Hinzunahme von neuen Messungen

$$\mathbf{y}_2 = \mathbf{S}_2\mathbf{p} + \mathbf{v}_2 \tag{2.61}$$

verbessern kann. Für die stochastische Störung $\mathbf{v}_2$ und den Zufallsparametervektor $\mathbf{p}$ gelte

$$\begin{aligned} \mathrm{E}(\mathbf{v}_2) = \mathbf{0}, \qquad \mathrm{cov}(\mathbf{v}_2) = \mathrm{E}\Big(\mathbf{v}_2\mathbf{v}_2^{\mathrm{T}}\Big) = \mathbf{Q}_2 \quad \text{mit} \quad \mathbf{Q}_2 \geq 0 \\ \mathrm{E}(\mathbf{p}) = \mathbf{0}, \qquad\qquad \mathrm{E}\Big(\mathbf{p}\mathbf{v}_2^{\mathrm{T}}\Big) = \mathbf{N}_2 \ . \end{aligned} \tag{2.62}$$

Weiters ist es natürlich sinnvoll, anzunehmen, dass die Störung $\mathbf{v}_2$ nicht mit vergangenen Messgrößen $\mathbf{y}_1$ korreliert ist und damit gilt

$$\mathrm{E}\Big(\mathbf{v}_2\mathbf{y}_1^{\mathrm{T}}\Big) = \mathbf{0} \quad \text{bzw.} \quad \mathrm{E}\Big(\mathbf{v}_2\hat{\mathbf{p}}_1^{\mathrm{T}}\Big) = \mathbf{0} \ . \tag{2.63}$$

Die beste Schätzung $\hat{\mathbf{y}}_2$ von $\mathbf{y}_2$ auf Basis der vergangenen Messwerte $\mathbf{y}_1$ lautet

$$\hat{\mathbf{y}}_2 = \mathbf{S}_2\hat{\mathbf{p}}_1 \ . \tag{2.64}$$

Damit erhält man nach Satz 2.6 mit $\tilde{\mathbf{y}}_2 = \mathbf{y}_2 - \hat{\mathbf{y}}_2$ den verbesserten Schätzwert $\hat{\mathbf{p}}_2$ zu

$$\hat{\mathbf{p}}_2 = \hat{\mathbf{p}}_1 + \mathrm{E}\Big(\mathbf{p}\tilde{\mathbf{y}}_2^{\mathrm{T}}\Big)\Big[\mathrm{E}\Big(\tilde{\mathbf{y}}_2\tilde{\mathbf{y}}_2^{\mathrm{T}}\Big)\Big]^{-1}\tilde{\mathbf{y}}_2 \tag{2.65}$$

mit

$$\mathrm{E}\Big(\mathbf{p}\tilde{\mathbf{y}}_2^{\mathrm{T}}\Big) = \mathrm{E}\Big(\mathbf{p}(\mathbf{p}-\hat{\mathbf{p}}_1)^{\mathrm{T}}\mathbf{S}_2^{\mathrm{T}} + \mathbf{p}\mathbf{v}_2^{\mathrm{T}}\Big) \overset{(2.53)}{=} \mathbf{P}_1\mathbf{S}_2^{\mathrm{T}} + \mathbf{N}_2 \tag{2.66}$$

und

$$\mathrm{E}\Big(\tilde{\mathbf{y}}_2\tilde{\mathbf{y}}_2^{\mathrm{T}}\Big) = \mathrm{E}\Big([\mathbf{S}_2(\mathbf{p}-\hat{\mathbf{p}}_1)+\mathbf{v}_2][\mathbf{S}_2(\mathbf{p}-\hat{\mathbf{p}}_1)+\mathbf{v}_2]^{\mathrm{T}}\Big) = \mathbf{S}_2\mathbf{P}_1\mathbf{S}_2^{\mathrm{T}} + \mathbf{Q}_2 + \mathbf{S}_2\mathbf{N}_2 + \mathbf{N}_2^{\mathrm{T}}\mathbf{S}_2^{\mathrm{T}} \ . \tag{2.67}$$

Aufgabe 2.9. Zeigen Sie, dass sich die Fehlerkovarianzmatrix analog zu (2.44) in der Form

$$\begin{aligned} \mathbf{P}_2 &= \mathrm{cov}(\mathbf{p}-\hat{\mathbf{p}}_2) \\ &= \mathbf{P}_1 - \Big(\mathbf{P}_1\mathbf{S}_2^{\mathrm{T}} + \mathbf{N}_2\Big)\Big(\mathbf{S}_2\mathbf{P}_1\mathbf{S}_2^{\mathrm{T}} + \mathbf{Q}_2 + \mathbf{S}_2\mathbf{N}_2 + \mathbf{N}_2^{\mathrm{T}}\mathbf{S}_2^{\mathrm{T}}\Big)^{-1}\Big(\mathbf{S}_2\mathbf{P}_1 + \mathbf{N}_2^{\mathrm{T}}\Big) \end{aligned} \tag{2.68}$$

errechnet.

Damit ergibt sich der *rekursive Minimum-Varianz-Schätzer* zu

$$\hat{\mathbf{p}}_k = \hat{\mathbf{p}}_{k-1} + \Big(\mathbf{P}_{k-1}\mathbf{S}_k^{\mathrm{T}} + \mathbf{N}_k\Big)\Big(\mathbf{S}_k\mathbf{P}_{k-1}\mathbf{S}_k^{\mathrm{T}} + \mathbf{Q}_k + \mathbf{S}_k\mathbf{N}_k + \mathbf{N}_k^{\mathrm{T}}\mathbf{S}_k^{\mathrm{T}}\Big)^{-1}(\mathbf{y}_k - \mathbf{S}_k\hat{\mathbf{p}}_{k-1}) \tag{2.69}$$

mit

$$\mathbf{P}_k = \mathbf{P}_{k-1} - \Big(\mathbf{P}_{k-1}\mathbf{S}_k^{\mathrm{T}} + \mathbf{N}_k\Big)\Big(\mathbf{S}_k\mathbf{P}_{k-1}\mathbf{S}_k^{\mathrm{T}} + \mathbf{Q}_k + \mathbf{S}_k\mathbf{N}_k + \mathbf{N}_k^{\mathrm{T}}\mathbf{S}_k^{\mathrm{T}}\Big)^{-1}\Big(\mathbf{S}_k\mathbf{P}_{k-1} + \mathbf{N}_k^{\mathrm{T}}\Big) \tag{2.70}$$

und den Anfangswerten $\mathbf{P}_{-1}$ sowie $\hat{\mathbf{p}}_{-1}$.

Nimmt man nun an, dass in jedem Iterationsschritt genau eine neue Messung hinzukommt, d. h. die Größen y_k und v_k sind Skalare, dann folgt durch Einsetzen von $\mathbf{S}_k = \mathbf{s}_k^{\mathrm{T}}$, $\mathbf{Q}_k = \mathrm{E}(v_k^2) = q_k$ und $\mathbf{N}_k = \mathrm{E}(\mathbf{p}v_k) = \mathbf{n}_k$ in (2.69), (2.70) der rekursive Minimum-Varianz-Schätzer zu

$$\mathbf{k}_k = \frac{\mathbf{P}_{k-1}\mathbf{s}_k + \mathbf{n}_k}{(q_k + 2\mathbf{s}_k^{\mathrm{T}}\mathbf{n}_k + \mathbf{s}_k^{\mathrm{T}}\mathbf{P}_{k-1}\mathbf{s}_k)} \tag{2.71a}$$

$$\mathbf{P}_k = \mathbf{P}_{k-1} - \mathbf{k}_k\Big(\mathbf{s}_k^{\mathrm{T}}\mathbf{P}_{k-1} + \mathbf{n}_k^{\mathrm{T}}\Big) \tag{2.71b}$$

$$\hat{\mathbf{p}}_k = \hat{\mathbf{p}}_{k-1} + \mathbf{k}_k\Big(y_k - \mathbf{s}_k^{\mathrm{T}}\hat{\mathbf{p}}_{k-1}\Big)\ . \tag{2.71c}$$

Man beachte auch in diesem Zusammenhang die Analogie zur rekursiven Methode der gewichteten kleinsten Quadrate (1.121) für $q_k = 1/\alpha_k$ und $\mathbf{n}_k = \mathbf{0}$.

2.3 Das Kalman-Filter

Aufbauend auf den bisherigen Überlegungen, insbesondere der rekursiven Minimum-Varianz-Schätzung, soll im nächsten Schritt das Kalman-Filter, ein im Sinne der Regelungstheorie *optimaler Beobachter*, hergeleitet werden. Bezüglich der Grundlagen der Zustandsbeobachtertheorie sei auf Kapitel 8 des Skriptums Automatisierung verwiesen. In der Literatur existieren unzählige Versionen des Kalman-Filters. Im Rahmen dieser Vorlesung wird vorerst ein lineares, zeitinvariantes, zeitdiskretes System der Form

$$\mathbf{x}_{k+1} = \mathbf{\Phi}\mathbf{x}_k + \mathbf{\Gamma}\mathbf{u}_k + \mathbf{G}\mathbf{w}_k \qquad \mathbf{x}(0) = \mathbf{x}_0 \tag{2.72a}$$

$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{D}\mathbf{u}_k + \mathbf{v}_k \tag{2.72b}$$

mit dem n-dimensionalen Zustand $\mathbf{x} \in \mathbb{R}^n$, dem p-dimensionalen deterministischen Eingang $\mathbf{u} \in \mathbb{R}^p$, dem q-dimensionalen Ausgang $\mathbf{y} \in \mathbb{R}^q$, der r-dimensionalen Störung $\mathbf{w} \in \mathbb{R}^r$, dem Messrauschen $\mathbf{v}$ sowie den Matrizen $\mathbf{\Phi} \in \mathbb{R}^{n\times n}$, $\mathbf{\Gamma} \in \mathbb{R}^{n\times p}$, $\mathbf{G} \in \mathbb{R}^{n\times r}$, $\mathbf{C} \in \mathbb{R}^{q\times n}$ und $\mathbf{D} \in \mathbb{R}^{q\times p}$ betrachtet. Es sei an dieser Stelle aber bereits angemerkt, dass das Kalman-Filter auch für lineare, zeitvariante und für zeitkontinuierliche Systeme entworfen werden kann. Es gelten nun folgende Annahmen:

(1) Von der Störung $\mathbf{w}$ und vom Messrauschen $\mathbf{v}$ wird vorausgesetzt, dass gilt

$$\mathrm{E}(\mathbf{v}_k) = \mathbf{0} \qquad \mathrm{E}\Big(\mathbf{v}_k\mathbf{v}_j^{\mathrm{T}}\Big) = \mathbf{R}\delta_{kj} \tag{2.73a}$$

$$\mathrm{E}(\mathbf{w}_k) = \mathbf{0} \qquad \mathrm{E}\Big(\mathbf{w}_k\mathbf{w}_j^{\mathrm{T}}\Big) = \mathbf{Q}\delta_{kj} \tag{2.73b}$$

$$\mathrm{E}\Big(\mathbf{w}_k\mathbf{v}_j^{\mathrm{T}}\Big) = \mathbf{0} \tag{2.73c}$$

mit $\mathbf{Q} \geq 0$ sowie $\mathbf{R} > 0$ und dem Kroneckersymbol $\delta_{kj} = 1$ für $k = j$ und $\delta_{kj} = 0$ sonst.

(2) Der Erwartungswert des Anfangswertes und die Kovarianzmatrix des Anfangsfehlers sind durch

$$\mathrm{E}(\mathbf{x}_0) = \mathbf{m}_0 \qquad \mathrm{E}\Big([\mathbf{x}_0 - \hat{\mathbf{x}}_0][\mathbf{x}_0 - \hat{\mathbf{x}}_0]^{\mathrm{T}}\Big) = \mathbf{P}_0 \geq 0 \tag{2.74}$$

mit dem Schätzwert $\hat{\mathbf{x}}_0$ des Anfangswertes $\mathbf{x}_0$ gegeben.

(3) Die Störung $\mathbf{w}_k$, $k \geq 0$, und das Messrauschen $\mathbf{v}_l$, $l \geq 0$, sind mit dem Anfangswert $\mathbf{x}_0$ nicht korreliert, d. h. es gilt

$$\mathrm{E}\Big(\mathbf{w}_k \mathbf{x}_0^{\mathrm{T}}\Big) = \mathbf{0} \tag{2.75a}$$

$$\mathrm{E}\Big(\mathbf{v}_l \mathbf{x}_0^{\mathrm{T}}\Big) = \mathbf{0}\ . \tag{2.75b}$$

Damit folgt aber wegen

$$\mathbf{x}_j = \mathbf{\Phi}^j \mathbf{x}_0 + \sum_{l=0}^{j-1} \mathbf{\Phi}^l (\mathbf{\Gamma}\mathbf{u}_{j-1-l} + \mathbf{G}\mathbf{w}_{j-1-l}) \tag{2.76}$$

und (2.73) auch die Beziehung

$$\mathrm{E}\Big(\mathbf{w}_k \mathbf{x}_j^{\mathrm{T}}\Big) = \mathbf{0} \quad \text{für } k \geq j \tag{2.77a}$$

$$\mathrm{E}\Big(\mathbf{v}_l \mathbf{x}_j^{\mathrm{T}}\Big) = \mathbf{0} \quad \text{für alle } l, j\ . \tag{2.77b}$$

Für die weiteren Betrachtungen wird nachfolgende Notation eingeführt:

Definition 2.3. Die optimale Schätzung von $\mathbf{x}_k$ unter Berücksichtigung von $0, \ldots, j$ Messungen wird mit $\hat{\mathbf{x}}(k|j)$ abgekürzt.

Satz 2.7 (Kalman-Filter). *Die optimale Schätzung $\hat{\mathbf{x}}(k+1|k)$ des Zustandes $\mathbf{x}_{k+1}$ des Systems (2.72) unter Berücksichtigung von $l = 0, \ldots, k$ Messungen errechnet sich nach der Iterationsvorschrift*

$$\begin{aligned} &\hat{\mathbf{x}}(k+1|k) = \mathbf{\Phi}\hat{\mathbf{x}}(k|k-1) + \mathbf{\Gamma}\mathbf{u}_k + \\ &\mathbf{\Phi}\mathbf{P}(k|k-1)\mathbf{C}^{\mathrm{T}}\Big(\mathbf{C}\mathbf{P}(k|k-1)\mathbf{C}^{\mathrm{T}} + \mathbf{R}\Big)^{-1}(\mathbf{y}_k - \mathbf{C}\hat{\mathbf{x}}(k|k-1) - \mathbf{D}\mathbf{u}_k) \end{aligned} \tag{2.78}$$

mit der Kovarianzmatrix des Schätzfehlers

$$\begin{aligned} &\mathbf{P}(k+1|k) = \mathbf{\Phi}\mathbf{P}(k|k-1)\mathbf{\Phi}^{\mathrm{T}} + \mathbf{G}\mathbf{Q}\mathbf{G}^{\mathrm{T}} \\ &- \mathbf{\Phi}\mathbf{P}(k|k-1)\mathbf{C}^{\mathrm{T}}\Big(\mathbf{C}\mathbf{P}(k|k-1)\mathbf{C}^{\mathrm{T}} + \mathbf{R}\Big)^{-1}\mathbf{C}\mathbf{P}(k|k-1)\mathbf{\Phi}^{\mathrm{T}} \end{aligned} \tag{2.79}$$

und den Anfangswerten $\hat{\mathbf{x}}(0|-1) = \mathbf{m}_0$ und $\mathbf{P}(0|-1) = \mathbf{P}_0$.

Beweis zu Satz 2.7. Angenommen, es wurden die Messungen $\mathbf{y}_0, \mathbf{y}_1, \ldots, \mathbf{y}_{k-1}$ für die optimale Schätzung $\hat{\mathbf{x}}(k|k-1)$ mit der Fehlerkovarianzmatrix

$$\mathbf{P}(k|k-1) = \mathrm{E}\Big([\mathbf{x}_k - \hat{\mathbf{x}}(k|k-1)][\mathbf{x}_k - \hat{\mathbf{x}}(k|k-1)]^{\mathrm{T}}\Big) \tag{2.80}$$

herangezogen. Zum Zeitpunkt k wird nun die Messung $\mathbf{y}_k$

$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{D}\mathbf{u}_k + \mathbf{v}_k \tag{2.81}$$

zur Verbesserung des Schätzwertes von $\mathbf{x}_k$ verwendet. Gemäß Satz 2.6 lautet der Schätzwert $\hat{\mathbf{x}}(k|k)$ von $\mathbf{x}_k$

$$\hat{\mathbf{x}}(k|k) = \hat{\mathbf{x}}(k|k-1) + \mathrm{E}\Big(\mathbf{x}_k\tilde{\mathbf{y}}_k^{\mathrm{T}}\Big)\Big[\mathrm{E}\Big(\tilde{\mathbf{y}}_k\tilde{\mathbf{y}}_k^{\mathrm{T}}\Big)\Big]^{-1}\tilde{\mathbf{y}}_k \tag{2.82a}$$

$$\tilde{\mathbf{y}}_k = \mathbf{y}_k - \mathbf{C}\hat{\mathbf{x}}(k|k-1) - \mathbf{D}\mathbf{u}_k \tag{2.82b}$$

bzw. mit (2.53) in

$$\mathrm{E}\Big(\mathbf{x}_k\tilde{\mathbf{y}}_k^{\mathrm{T}}\Big) = \mathrm{E}\Big(\mathbf{x}_k(\mathbf{C}\mathbf{x}_k + \mathbf{v}_k - \mathbf{C}\hat{\mathbf{x}}(k|k-1))^{\mathrm{T}}\Big) = \mathbf{P}(k|k-1)\mathbf{C}^{\mathrm{T}} \tag{2.83}$$

und

$$\mathrm{E}\Big(\tilde{\mathbf{y}}_k\tilde{\mathbf{y}}_k^{\mathrm{T}}\Big) = \mathbf{C}\mathbf{P}(k|k-1)\mathbf{C}^{\mathrm{T}} + \mathbf{R} \tag{2.84}$$

folgt

$$\begin{aligned}&\hat{\mathbf{x}}(k|k) = \hat{\mathbf{x}}(k|k-1) + \\ &\mathbf{P}(k|k-1)\mathbf{C}^{\mathrm{T}}\Big(\mathbf{C}\mathbf{P}(k|k-1)\mathbf{C}^{\mathrm{T}} + \mathbf{R}\Big)^{-1}(\mathbf{y}_k - \mathbf{C}\hat{\mathbf{x}}(k|k-1) - \mathbf{D}\mathbf{u}_k)\ .\end{aligned} \tag{2.85}$$

Damit lässt sich die Fehlerkovarianzmatrix in der Form (vergleiche dazu (2.44), (2.70))

$$\begin{aligned}&\mathbf{P}(k|k) = \mathrm{E}\Big([\mathbf{x}_k - \hat{\mathbf{x}}(k|k)][\mathbf{x}_k - \hat{\mathbf{x}}(k|k)]^{\mathrm{T}}\Big) = \\ &\mathbf{P}(k|k-1) - \mathbf{P}(k|k-1)\mathbf{C}^{\mathrm{T}}\Big(\mathbf{C}\mathbf{P}(k|k-1)\mathbf{C}^{\mathrm{T}} + \mathbf{R}\Big)^{-1}\mathbf{C}\mathbf{P}(k|k-1)\end{aligned} \tag{2.86}$$

anschreiben. Nach Satz 2.5 ist die optimale Schätzung von $\mathbf{\Phi}\mathbf{x}_k$ gleich die optimale Schätzung $\hat{\mathbf{x}}_k$ von $\mathbf{x}_k$ multipliziert mit $\mathbf{\Phi}$, es gilt also

$$\hat{\mathbf{x}}(k+1|k) = \mathbf{\Phi}\hat{\mathbf{x}}(k|k) + \mathbf{\Gamma}\mathbf{u}_k\ . \tag{2.87}$$

Für die Kovarianzmatrix des Schätzfehlers erhält man

$$\begin{aligned}\mathbf{P}(k+1|k) &= \mathrm{E}\Big([\mathbf{x}_{k+1} - \hat{\mathbf{x}}(k+1|k)][\mathbf{x}_{k+1} - \hat{\mathbf{x}}(k+1|k)]^{\mathrm{T}}\Big) \\ &= \mathrm{E}\Big([\boldsymbol{\Phi}(\mathbf{x}_k - \hat{\mathbf{x}}(k|k)) + \mathbf{G}\mathbf{w}_k][\boldsymbol{\Phi}(\mathbf{x}_k - \hat{\mathbf{x}}(k|k)) + \mathbf{G}\mathbf{w}_k]^{\mathrm{T}}\Big) \\ &= \boldsymbol{\Phi}\mathbf{P}(k|k)\boldsymbol{\Phi}^{\mathrm{T}} + \mathbf{G}\mathbf{Q}\mathbf{G}^{\mathrm{T}}\end{aligned} \tag{2.88}$$

Durch Kombination von (2.85)–(2.88) ergibt sich unmittelbar das Ergebnis von Satz 2.7. □

Es sei an dieser Stelle noch die Zusammensetzung der Kovarianzmatrix des Schätzfehlers (2.79) interpretiert: Der Term $\boldsymbol{\Phi}\mathbf{P}(k|k-1)\boldsymbol{\Phi}^{\mathrm{T}}$ beschreibt die Änderung der Kovarianzmatrix zufolge der Systemdynamik, $\mathbf{G}\mathbf{Q}\mathbf{G}^{\mathrm{T}}$ gibt die Erhöhung der Fehlervarianz zufolge der Störung $\mathbf{w}$ an und der verbleibende Ausdruck mit negativem Vorzeichen beschreibt, wie sich die Fehlervarianz durch Hinzunahme der Information neuer Messungen verringert.

2.3.1 Das Kalman-Filter als optimaler Beobachter

Führt man die Abkürzungen $\hat{\mathbf{x}}_{k+1} = \hat{\mathbf{x}}(k+1|k)$, $\hat{\mathbf{x}}_k = \hat{\mathbf{x}}(k|k-1)$, $\mathbf{P}_{k+1} = \mathbf{P}(k+1|k)$ und $\mathbf{P}_k = \mathbf{P}(k|k-1)$ ein, dann kann man (2.78) und (2.79) auch in der kompakten Form

$$\hat{\mathbf{x}}_{k+1} = \boldsymbol{\Phi}\hat{\mathbf{x}}_k + \boldsymbol{\Gamma}\mathbf{u}_k + \hat{\mathbf{K}}_k(\mathbf{y}_k - \mathbf{C}\hat{\mathbf{x}}_k - \mathbf{D}\mathbf{u}_k) \tag{2.89}$$

mit

$$\hat{\mathbf{K}}_k = \boldsymbol{\Phi}\mathbf{P}_k\mathbf{C}^{\mathrm{T}}\Big(\mathbf{C}\mathbf{P}_k\mathbf{C}^{\mathrm{T}} + \mathbf{R}\Big)^{-1} \tag{2.90}$$

und

$$\mathbf{P}_{k+1} = \boldsymbol{\Phi}\mathbf{P}_k\boldsymbol{\Phi}^{\mathrm{T}} + \mathbf{G}\mathbf{Q}\mathbf{G}^{\mathrm{T}} - \boldsymbol{\Phi}\mathbf{P}_k\mathbf{C}^{\mathrm{T}}\Big(\mathbf{C}\mathbf{P}_k\mathbf{C}^{\mathrm{T}} + \mathbf{R}\Big)^{-1}\mathbf{C}\mathbf{P}_k\boldsymbol{\Phi}^{\mathrm{T}} \tag{2.91}$$

darstellen. Gleichung (2.91) wird auch als *diskrete Riccati-Gleichung* bezeichnet. Vergleicht man (2.89) mit einem vollständigen Luenberger-Beobachter, dann erkennt man, dass das Kalman-Filter ein vollständiger Beobachter mit einer *zeitvarianten Beobachterverstärkungsmatrix* $\hat{\mathbf{K}}_k$ ist. Der Erwartungswert des Beobachtungsfehlers $\mathbf{e}_k = \mathbf{x}_k - \hat{\mathbf{x}}_k$ genügt der Iterationsvorschrift

$$\begin{aligned}\mathrm{E}(\mathbf{e}_{k+1}) &= \mathrm{E}\Big(\boldsymbol{\Phi}\mathbf{x}_k + \mathbf{G}\mathbf{w}_k - \boldsymbol{\Phi}\hat{\mathbf{x}}_k - \hat{\mathbf{K}}_k(\mathbf{C}\mathbf{x}_k + \mathbf{v}_k - \mathbf{C}\hat{\mathbf{x}}_k)\Big) \\ &= \mathrm{E}\Big(\Big(\boldsymbol{\Phi} - \hat{\mathbf{K}}_k\mathbf{C}\Big)(\mathbf{x}_k - \hat{\mathbf{x}}_k) + \mathbf{G}\mathbf{w}_k - \hat{\mathbf{K}}_k\mathbf{v}_k\Big) \\ &= \Big(\boldsymbol{\Phi} - \hat{\mathbf{K}}_k\mathbf{C}\Big)\,\mathrm{E}(\mathbf{e}_k)\ .\end{aligned} \tag{2.92}$$

Falls also $\hat{\mathbf{x}}_0 = \mathrm{E}(\mathbf{x}_0) = \mathbf{m}_0$ gesetzt wird, gilt $\mathrm{E}(\mathbf{e}_k) = \mathbf{0}$ für alle $k \geq 0$. Weiters sieht man aus (2.90) und (2.91), dass ausgehend vom Startwert $\mathbf{P}_0$ die Fehlerkovarianzmatrix $\mathbf{P}_k$ und damit auch $\hat{\mathbf{K}}_k$ ohne Kenntnis der Messgrößen $\mathbf{y}_k$ für alle $k \geq 0$ vorwegberechnet und im Rechner zwischengespeichert werden können.

Wenn keine vorherigen Messwerte über den Prozess vorliegen, setzt man typischerweise $\hat{\mathbf{x}}_0 = \mathbf{0}$ und $\mathbf{P}_0 = \alpha\mathbf{E}$ für $\alpha \gg 1$. Wenn der Beobachter eine sehr lange Zeit läuft, dann

kann das Problem mathematisch so behandelt werden, als würde er unendlich lange im Einsatz sein. Es zeigt sich nun, dass für unendlich lange Zeit die Kovarianzmatrix des Schätzfehlers auf einen stationären Wert $\mathbf{P}_\infty$ einläuft. In diesem Falle ist auch die Beobachterverstärkungsmatrix $\hat{\mathbf{K}}_\infty$ konstant und errechnet sich zu

$$\hat{\mathbf{K}}_\infty = \mathbf{\Phi}\mathbf{P}_\infty\mathbf{C}^{\mathrm{T}}\Big(\mathbf{C}\mathbf{P}_\infty\mathbf{C}^{\mathrm{T}} + \mathbf{R}\Big)^{-1} \tag{2.93}$$

mit $\mathbf{P}_\infty$ als Lösung der so genannten *diskreten algebraischen Riccati-Gleichung*

$$\mathbf{P}_\infty = \mathbf{\Phi}\mathbf{P}_\infty\mathbf{\Phi}^{\mathrm{T}} + \mathbf{G}\mathbf{Q}\mathbf{G}^{\mathrm{T}} - \mathbf{\Phi}\mathbf{P}_\infty\mathbf{C}^{\mathrm{T}}\Big(\mathbf{C}\mathbf{P}_\infty\mathbf{C}^{\mathrm{T}} + \mathbf{R}\Big)^{-1}\mathbf{C}\mathbf{P}_\infty\mathbf{\Phi}^{\mathrm{T}}\ . \tag{2.94}$$

Gleichung (2.94) hat eine eindeutige symmetrische Lösung $\mathbf{P}_\infty$ mit der Eigenschaft, dass sämtliche Eigenwerte von $\Big(\mathbf{\Phi} - \hat{\mathbf{K}}_\infty\mathbf{C}\Big)$ im offenen Inneren des Einheitskreises liegen, wenn nachfolgende Bedingungen erfüllt sind:

(1) Das Paar $(\mathbf{C}, \mathbf{\Phi})$ ist *detektierbar*, d. h. sämtliche Eigenwerte außerhalb des Einheitskreises sind beobachtbar.

(2) Das Paar $\Big(\mathbf{\Phi}, \mathbf{G}\mathbf{Q}\mathbf{G}^{\mathrm{T}}\Big)$ ist *stabilisierbar*, d. h. sämtliche Eigenwerte außerhalb des Einheitskreises sind über den Eingang $\mathbf{G}\mathbf{Q}\mathbf{G}^{\mathrm{T}}$ steuerbar.

(3) Die Matrix $\mathbf{R}$ ist positiv definit.

Eine derartige Lösung der diskreten algebraischen Riccati-Gleichung (2.94) wird auch als *stabilisierende Lösung* bezeichnet. Da für diese stabilisierende Lösung alle Eigenwerte von $\Big(\mathbf{\Phi} - \hat{\mathbf{K}}_\infty\mathbf{C}\Big)$ im offenen Inneren des Einheitskreises liegen, nimmt gemäß (2.92) der Erwartungswert des Beobachtungsfehlers ab und es gilt $\lim_{k\to\infty} \mathrm{E}(\mathbf{e}_k) = \mathbf{0}$. Die Lösung $\mathbf{P}_\infty$ der diskreten algebraischen Riccati-Gleichung (2.94) kann man einfach dadurch erhalten, dass man die diskrete Riccati-Gleichung (2.91) vom Anfangswert $\mathbf{P}_0$ beginnend so lange iteriert, bis sich $\mathbf{P}_k$ nur noch hinreichend wenig im Sinne einer Norm ändert. Obwohl die Iterationsvorschrift im Allgemeinen sehr schnell gegen einen stationären Wert konvergiert, wird in der Praxis, so auch in MATLAB, die algebraische Riccati-Gleichung (2.94) numerisch effizienter über eine Eigenvektordekomposition gelöst, siehe die MATLAB–Befehle `care` bzw. `dare`.

Aufgabe 2.10. Die Bewegung eines Satelliten um eine Achse wird in der Form

$$I\frac{\mathrm{d}^2}{\mathrm{d}t^2}\varphi = M_c - M_d$$

mit dem Trägheitsmoment I, dem Moment M_c als Stellgröße, dem Moment M_d als Störung und dem Winkel φ modelliert. Bestimmen Sie für die Abtastzeit $T_a = 1\mathrm{s}$ das zugehörige Abtastsystem der Form (2.72) für $I = 1$ und der Ausgangsgröße φ. Nehmen Sie dabei an, dass die Messung des Winkels φ mit dem Messrauschen v überlagert ist und das Störmoment M_d der auf den Prozess wirkenden Störung w

entspricht. Es gelte dabei

$$\mathrm{E}(w) = 0 \qquad \mathrm{E}\left(w^2\right) = q$$
$$\mathrm{E}(v) = 0 \qquad \mathrm{E}\left(v^2\right) = 0.1 \ .$$

Stellen Sie die Elemente von $\mathbf{P}_k$ des Kalman-Filters gemäß der Iterationsvorschrift (2.91) für den Anfangswert $\mathbf{P}_0 = \mathbf{E}$ und $q = \{0.1, 0.01, 0.001\}$ dar. Implementieren Sie das Kalman-Filter in MATLAB/SIMULINK.

Hinweis: Die Beziehung (2.93) mit der konstanten Beobachterverstärkungsmatrix $\hat{\mathbf{K}}_\infty$ gibt einen *optimalen vollständigen Beobachter* an, der *sowohl für Ein- als auch für Mehrgrößensysteme* verwendbar ist. Im Gegensatz zum Beobachterentwurf nach der Methode der Polvorgabe (man beachte dazu die Formel von Ackermann gemäß Kapitel 8 des Skriptums Automatisierung) müssen beim Kalman-Filter keine Pole des Fehlersystems gewählt werden, was insbesondere im Mehrgrößenfall sehr schwierig sein kann. Man beeinflusst das Verhalten des Fehlersystems hingegen durch die Festlegung der Kovarianzmatrizen $\mathbf{Q}$ der Störung $\mathbf{w}$ und $\mathbf{R}$ des Messrauschens $\mathbf{v}$.

Für die Wahl der Kovarianzmatrix $\mathbf{R}$ des Messrauschens $\mathbf{v}$ kann sehr oft ein interpretierbarer Ansatz, der auf den (Rausch-)Eigenschaften des Sensors beruht, gefunden werden. Im Weiteren kann über die Gewichtung der Einträge der Kovarianzmatrix $\mathbf{R}$ zwischen zuverlässigen und weniger zuverlässigen Messungen unterschieden werden. Ist eine *Messung weniger zuverlässig*, so wird der *zugehörige Eintrag* in der Hauptdiagonale *der Kovarianzmatrix sehr groß gewählt.* Diese angenommene große Varianz der Messung bedingt, dass der Beobachter diese Messung im Vergleich zu den anderen Messungen bei der Zustandsschätzung weniger gewichtet. In der praktischen Anwendung des Kalman-Filters ist es sogar üblich, während des laufenden Betriebs die Kovarianzmatrix umzuschalten, wenn einer oder mehrere Sensoren unplausible Messungen liefern oder wenn man in einen Betriebsbereich kommt, wo man bereits von vornherein weiß, dass gewisse Sensoren keine zuverlässige Information mehr zur Verfügung stellen.

Für die Prozessstörung $\mathbf{w}$ und damit die Kovarianzmatrix $\mathbf{Q}$ gelten diese Annahmen im Allgemeinen nicht. Die Annahme, dass es sich bei $\mathbf{w}$ um weißes Rauschen handelt, ist meist nicht zutreffend. Man könnte nun auf die Idee kommen, die im Allgemeinen unbekannte Matrix $\mathbf{Q}$ sehr klein oder sogar Null zu wählen. Die Wahl $\mathbf{Q} = \mathbf{0}$ entspricht aber dem Szenario einer Strecke ohne Störung $\mathbf{w}$. Wie man aus der Bedingung (2) für die Lösung der diskreten algebraischen Riccati-Gleichung entnehmen kann, führt die Wahl $\mathbf{Q} = \mathbf{0}$ (bzw. im Allgemeinen auch für $\mathbf{Q} \ll 1$) nicht zu einer stabilisierenden Lösung.

Die Wahl von $\mathbf{Q}$ (und auch $\mathbf{R}$) erfolgt in der praktischen Anwendung meist nach der Methode „Versuch und Irrtum".

Hinweis: In der MATLAB CONTROL SYSTEMS TOOLBOX wird ein etwas allgemeineres

System der Form

$$\mathbf{x}_{k+1} = \mathbf{\Phi}\mathbf{x}_k + \mathbf{\Gamma}\mathbf{u}_k + \mathbf{G}\mathbf{w}_k, \qquad \mathbf{x}(0) = \mathbf{x}_0 \tag{2.95a}$$

$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{D}\mathbf{u}_k + \mathbf{H}\mathbf{w}_k + \mathbf{v}_k, \tag{2.95b}$$

mit dem zusätzlichen Term $\mathbf{H}\mathbf{w}_k$ im gemessenen Ausgang, betrachtet. Weiterhin wird anstatt $\mathrm{E}\left(\mathbf{w}_k\mathbf{v}_k^{\mathrm{T}}\right) = \mathbf{0}$ zugelassen, dass

$$\mathrm{E}\left(\mathbf{w}_k\mathbf{v}_k^{\mathrm{T}}\right) = \mathbf{N}\delta_{kj} \neq \mathbf{0} \tag{2.96}$$

gilt. Für diesen Fall errechnet sich die optimale Zustandsschätzung in der Form

$$\hat{\mathbf{x}}_{k+1} = \mathbf{\Phi}\hat{\mathbf{x}}_k + \mathbf{\Gamma}\mathbf{u}_k + \hat{\mathbf{K}}_k(\mathbf{y}_k - \mathbf{C}\hat{\mathbf{x}}_k - \mathbf{D}\mathbf{u}_k) \tag{2.97}$$

mit

$$\hat{\mathbf{K}}_k = \left(\mathbf{\Phi}\mathbf{P}_k\mathbf{C}^{\mathrm{T}} + \mathbf{G}\mathbf{Q}\mathbf{H}^{\mathrm{T}} + \mathbf{G}\mathbf{N}\right)\left(\mathbf{C}\mathbf{P}_k\mathbf{C}^{\mathrm{T}} + \mathbf{H}\mathbf{Q}\mathbf{H}^{\mathrm{T}} + \mathbf{R} + \mathbf{H}\mathbf{N} + \mathbf{N}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\right)^{-1} \tag{2.98}$$

und

$$\mathbf{P}_{k+1} = \mathbf{\Phi}\mathbf{P}_k\mathbf{\Phi}^{\mathrm{T}} + \mathbf{G}\mathbf{Q}\mathbf{G}^{\mathrm{T}} - \hat{\mathbf{K}}_k\left(\mathbf{C}\mathbf{P}_k\mathbf{\Phi}^{\mathrm{T}} + \mathbf{H}\mathbf{Q}\mathbf{G}^{\mathrm{T}} + \mathbf{N}^{\mathrm{T}}\mathbf{G}^{\mathrm{T}}\right). \tag{2.99}$$

Der Beweis ist in der Literatur, insbesondere [2.2], zu finden.

2.3.2 Frequenzbereichseigenschaften des stationären Kalman-Filters

Es sei angenommen, dass das Messrauschen $\mathbf{v}_k$ von (2.72), d. h.

$$\mathbf{x}_{k+1} = \mathbf{\Phi}\mathbf{x}_k + \mathbf{\Gamma}\mathbf{u}_k + \mathbf{G}\mathbf{w}_k \qquad \mathbf{x}(0) = \mathbf{x}_0 \tag{2.100a}$$

$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{D}\mathbf{u}_k + \mathbf{v}_k \tag{2.100b}$$

durch ein dynamisches System der Form

$$\mathbf{z}_{k+1} = \mathbf{\Phi}_{\mathbf{z}}\mathbf{z}_k + \mathbf{n}_k \tag{2.101a}$$

$$\mathbf{v}_k = \mathbf{C}_{\mathbf{z}}\mathbf{z}_k + \mathbf{m}_k \tag{2.101b}$$

mit den stochastischen Störungen $\mathbf{n}_k$ und $\mathbf{m}_k$ erzeugt wird. Setzt man $\mathbf{v}_k$ von (2.101) in (2.100) ein, so folgt das erweiterte System in der Form

$$\underbrace{\begin{bmatrix}\mathbf{x}_{k+1}\\ \mathbf{z}_{k+1}\end{bmatrix}}_{\tilde{\mathbf{x}}_{k+1}} = \underbrace{\begin{bmatrix}\mathbf{\Phi} & \mathbf{0}\\ \mathbf{0} & \mathbf{\Phi}_{\mathbf{z}}\end{bmatrix}}_{\tilde{\mathbf{\Phi}}}\underbrace{\begin{bmatrix}\mathbf{x}_k\\ \mathbf{z}_k\end{bmatrix}}_{\tilde{\mathbf{x}}_k} + \underbrace{\begin{bmatrix}\mathbf{\Gamma}\\ \mathbf{0}\end{bmatrix}}_{\tilde{\mathbf{\Gamma}}}\mathbf{u}_k + \underbrace{\begin{bmatrix}\mathbf{G} & \mathbf{0}\\ \mathbf{0} & \mathbf{E}\end{bmatrix}}_{\tilde{\mathbf{G}}}\underbrace{\begin{bmatrix}\mathbf{w}_k\\ \mathbf{n}_k\end{bmatrix}}_{\tilde{\mathbf{w}}_k} \qquad \begin{matrix}\mathbf{x}(0) = \mathbf{x}_0\\ \mathbf{z}(0) = \mathbf{z}_0\end{matrix} \tag{2.102a}$$

$$\mathbf{y}_k = \underbrace{\begin{bmatrix}\mathbf{C} & \mathbf{C}_{\mathbf{z}}\end{bmatrix}}_{\tilde{\mathbf{C}}}\underbrace{\begin{bmatrix}\mathbf{x}_k\\ \mathbf{z}_k\end{bmatrix}}_{\tilde{\mathbf{x}}_k} + \mathbf{D}\mathbf{u}_k + \mathbf{m}_k\,. \tag{2.102b}$$

Das stationäre Kalman-Filter gemäß (2.89) und (2.93) für das erweiterte System (2.102) lautet dann

$$\begin{bmatrix}\hat{\mathbf{x}}_{k+1}\\ \hat{\mathbf{z}}_{k+1}\end{bmatrix} = \begin{bmatrix}\mathbf{\Phi} & \mathbf{0}\\ \mathbf{0} & \mathbf{\Phi_z}\end{bmatrix}\begin{bmatrix}\hat{\mathbf{x}}_k\\ \hat{\mathbf{z}}_k\end{bmatrix} + \begin{bmatrix}\mathbf{\Gamma}\\ \mathbf{0}\end{bmatrix}\mathbf{u}_k + \underbrace{\begin{bmatrix}\hat{\mathbf{K}}_\mathbf{x}\\ \hat{\mathbf{K}}_\mathbf{z}\end{bmatrix}}_{\hat{\mathbf{K}}_\infty}(\mathbf{y}_k - \mathbf{C}\hat{\mathbf{x}}_k - \mathbf{C_z}\hat{\mathbf{z}}_k - \mathbf{D}\mathbf{u}_k) \tag{2.103}$$

bzw.

$$\begin{bmatrix}\hat{\mathbf{x}}_{k+1}\\ \hat{\mathbf{z}}_{k+1}\end{bmatrix} = \begin{bmatrix}\mathbf{\Phi} - \hat{\mathbf{K}}_\mathbf{x}\mathbf{C} & -\hat{\mathbf{K}}_\mathbf{x}\mathbf{C_z}\\ -\hat{\mathbf{K}}_\mathbf{z}\mathbf{C} & \mathbf{\Phi_z} - \hat{\mathbf{K}}_\mathbf{z}\mathbf{C_z}\end{bmatrix}\begin{bmatrix}\hat{\mathbf{x}}_k\\ \hat{\mathbf{z}}_k\end{bmatrix} + \begin{bmatrix}\hat{\mathbf{K}}_\mathbf{x}\\ \hat{\mathbf{K}}_\mathbf{z}\end{bmatrix}\mathbf{y}_k + \begin{bmatrix}\mathbf{\Gamma} - \hat{\mathbf{K}}_\mathbf{x}\mathbf{D}\\ -\hat{\mathbf{K}}_\mathbf{z}\mathbf{D}\end{bmatrix}\mathbf{u}_k\ . \tag{2.104}$$

Da das System linear ist und damit das Superpositionsgesetz gilt, wird im Weiteren der Stelleingang $\mathbf{u}_k = \mathbf{0}$ gesetzt. Die z-Übertragungsmatrix vom Eingang $\mathbf{y}_k$ zum Ausgang $\hat{\mathbf{x}}_k$ lautet

$$\mathbf{G}(z) = \frac{\hat{\mathbf{x}}_z(z)}{\mathbf{y}_z(z)} = \begin{bmatrix}\mathbf{E} & \mathbf{0}\end{bmatrix}\begin{bmatrix}z\mathbf{E} - \mathbf{\Phi} + \hat{\mathbf{K}}_\mathbf{x}\mathbf{C} & \hat{\mathbf{K}}_\mathbf{x}\mathbf{C_z}\\ \hat{\mathbf{K}}_\mathbf{z}\mathbf{C} & z\mathbf{E} - \mathbf{\Phi_z} + \hat{\mathbf{K}}_\mathbf{z}\mathbf{C_z}\end{bmatrix}^{-1}\begin{bmatrix}\hat{\mathbf{K}}_\mathbf{x}\\ \hat{\mathbf{K}}_\mathbf{z}\end{bmatrix}\ . \tag{2.105}$$

Es gilt nun nachfolgender Satz:

Satz 2.8 (Übertragungsnullstellen des stationären Kalman-Filters)**.** *Die Übertragungsnullstellen von* (2.105) *sind durch die Beziehung*

$$\det(z\mathbf{E} - \mathbf{\Phi_z}) = 0 \tag{2.106}$$

gegeben.

Beweis. Bevor Satz 2.8 bewiesen wird, muss kurz auf den Begriff einer Übertragungsnullstelle eingegangen werden. Die Nullstellen der Übertragungsfunktion $G(z)$ eines Eingrößensystems werden meist als Wurzeln des Zählerpolynoms von $G(z)$ charakterisiert. Im Fall eines Mehrgrößensystems mit einer Übertragungsmatrix $\mathbf{G}(z)$ ist dies nicht mehr so einfach möglich. Eine strikte Definition ist in folgender Form gegeben: Die Nullstellen der Übertragungsmatrix $\mathbf{G}(z)$ sind die Wurzeln der Zählerpolynome der Smith-McMillan Form von $\mathbf{G}(z)$. Zur Definition der Smith-McMillan Form sei auf die am Ende angeführte Literatur verwiesen. Im Weiteren wird eine physikalische Interpretation einer Übertragungsnullstelle gegeben. Dazu betrachte man das lineare, zeitinvariante, zeitdiskrete System der Form

$$\mathbf{x}_{k+1} = \mathbf{\Phi}\mathbf{x}_k + \mathbf{\Gamma}\mathbf{u}_k \qquad \mathbf{x}(0) = \mathbf{x}_0 \tag{2.107a}$$

$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k \tag{2.107b}$$

mit dem n-dimensionalen Zustand $\mathbf{x} \in \mathbb{R}^n$, dem p-dimensionalen Eingang $\mathbf{u} \in \mathbb{R}^p$, dem q-dimensionalen Ausgang $\mathbf{y} \in \mathbb{R}^q$, sowie den Matrizen $\mathbf{\Phi} \in \mathbb{R}^{n\times n}$, $\mathbf{\Gamma} \in \mathbb{R}^{n\times p}$ und $\mathbf{C} \in \mathbb{R}^{q\times n}$. Eine komplexe Zahl z_j ist eine Übertragungsnullstelle, wenn zu einem Eingang der Form $(\mathbf{u}_k) = \mathbf{u}_0\left(z_j^k\right)$, $\mathbf{u}_0 \neq \mathbf{0}$, ein Anfangszustand $\mathbf{x}_0 \neq \mathbf{0}$ so existiert, dass der Ausgang $(\mathbf{y}_k)$ für alle Zeiten $k \geq 0$ identisch verschwindet. Diese Eigenschaft

ist in der Literatur auch unter dem Namen *transmission-blocking* bekannt.

Im z-Bereich errechnet sich die Ausgangsgröße $\mathbf{y}_z(z) = \mathcal{Z}\{(\mathbf{y}_k)\}$ von (2.107) als Antwort auf die Eingangsgröße

$$\mathbf{u}_z(z) = \mathcal{Z}\{(\mathbf{u}_k)\} = \mathcal{Z}\Big\{\mathbf{u}_0\Big(z_j^k\Big)\Big\} = \mathbf{u}_0\frac{z}{z-z_j} \tag{2.108}$$

mit dem Anfangswert $\mathbf{x}(0) = \mathbf{x}_0$ zu

$$\mathbf{y}_z(z) = \mathbf{C}(z\mathbf{E}-\mathbf{\Phi})^{-1}\left(z\mathbf{x}_0 + \mathbf{\Gamma}\mathbf{u}_0\frac{z}{z-z_j}\right). \tag{2.109}$$

Unter Zuhilfenahme der Resolventen-Identität

$$(z\mathbf{E}-\mathbf{\Phi})^{-1} - (z_j\mathbf{E}-\mathbf{\Phi})^{-1} = (z\mathbf{E}-\mathbf{\Phi})^{-1}(z_j-z)(z_j\mathbf{E}-\mathbf{\Phi})^{-1} \tag{2.110}$$

lässt sich (2.109) in der Form

$$\begin{aligned}\mathbf{y}_z(z) = \mathbf{C}\Big((z_j\mathbf{E}-\mathbf{\Phi})^{-1} + (z\mathbf{E}-\mathbf{\Phi})^{-1}(z_j-z)(z_j\mathbf{E}-\mathbf{\Phi})^{-1}\Big)\frac{\mathbf{\Gamma}\mathbf{u}_0 z}{z-z_j} + \\ \mathbf{C}(z\mathbf{E}-\mathbf{\Phi})^{-1}z\mathbf{x}_0 \\ = \mathbf{C}(z\mathbf{E}-\mathbf{\Phi})^{-1}z\Big(\mathbf{x}_0 - (z_j\mathbf{E}-\mathbf{\Phi})^{-1}\mathbf{\Gamma}\mathbf{u}_0\Big) + \mathbf{C}(z_j\mathbf{E}-\mathbf{\Phi})^{-1}\mathbf{\Gamma}\mathbf{u}_0\frac{z}{z-z_j}\end{aligned} \tag{2.111}$$

umschreiben.

Aufgabe 2.11. Beweisen Sie die Identität von (2.110).

Man erkennt aus (2.111), dass $\mathbf{y}_z(z)$ nur dann identisch verschwindet, wenn die Gleichungen

$$\mathbf{x}_0 - (z_j\mathbf{E}-\mathbf{\Phi})^{-1}\mathbf{\Gamma}\mathbf{u}_0 = \mathbf{0} \tag{2.112a}$$

$$\mathbf{C}(z_j\mathbf{E}-\mathbf{\Phi})^{-1}\mathbf{\Gamma}\mathbf{u}_0 = \mathbf{0} \tag{2.112b}$$

erfüllt sind. Zusammenfassend lässt sich also feststellen, dass für eine Übertragungsnullstelle z_j nichttriviale Vektoren $\mathbf{u}_0$ und $\mathbf{x}_0$ existieren, die dem Gleichungssystem

$$\begin{bmatrix}(z_j\mathbf{E}-\mathbf{\Phi}) & -\mathbf{\Gamma} \\ \mathbf{C} & \mathbf{0}\end{bmatrix}\begin{bmatrix}\mathbf{x}_0 \\ \mathbf{u}_0\end{bmatrix} = \begin{bmatrix}\mathbf{0} \\ \mathbf{0}\end{bmatrix} \tag{2.113}$$

genügen und die Eigenschaft besitzen, dass der Ausgang $(\mathbf{y}_k)$ von (2.107) für die Eingangsgröße $(\mathbf{u}_k) = \mathbf{u}_0\Big(z_j^k\Big)$ und den Anfangswert $\mathbf{x}_0$ für alle Zeiten $k \geq 0$ identisch verschwindet.

Wendet man nun dieses Ergebnis auf die Übertragungsmatrix $\mathbf{G}(z)$ von (2.105) mit dem Eingang $\mathbf{y}_k$ und dem Ausgang $\hat{\mathbf{x}}_k$ an, so lauten in diesem Fall die Bedingungen

(2.113)

$$\begin{bmatrix} z_j\mathbf{E} - \mathbf{\Phi} + \hat{\mathbf{K}}_\mathbf{x}\mathbf{C} & \hat{\mathbf{K}}_\mathbf{x}\mathbf{C}_\mathbf{z} \\ \hat{\mathbf{K}}_\mathbf{z}\mathbf{C} & z_j\mathbf{E} - \mathbf{\Phi}_\mathbf{z} + \hat{\mathbf{K}}_\mathbf{z}\mathbf{C}_\mathbf{z} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{x}}_0 \\ \hat{\mathbf{z}}_0 \end{bmatrix} - \begin{bmatrix} \hat{\mathbf{K}}_\mathbf{x} \\ \hat{\mathbf{K}}_\mathbf{z} \end{bmatrix} \mathbf{y}_0 = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix} \tag{2.114a}$$

$$\begin{bmatrix} \mathbf{E} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{x}}_0 \\ \hat{\mathbf{z}}_0 \end{bmatrix} = \mathbf{0} \; . \tag{2.114b}$$

Aufgabe 2.12. Leiten Sie die Beziehungen (2.114) her.

Aus (2.114b) erhält man $\hat{\mathbf{x}}_0 = \mathbf{0}$ und damit folgt

$$\begin{aligned} \hat{\mathbf{K}}_\mathbf{x}(\mathbf{C}_\mathbf{z}\hat{\mathbf{z}}_0 - \mathbf{y}_0) &= \mathbf{0} \\ \left(z_j\mathbf{E} - \mathbf{\Phi}_\mathbf{z} + \hat{\mathbf{K}}_\mathbf{z}\mathbf{C}_\mathbf{z}\right)\hat{\mathbf{z}}_0 - \hat{\mathbf{K}}_\mathbf{z}\mathbf{y}_0 &= \mathbf{0} \; . \end{aligned} \tag{2.115}$$

Unter Voraussetzung der Spaltenregularität von $\hat{\mathbf{K}}_\mathbf{x}$ folgt $\mathbf{C}_\mathbf{z}\hat{\mathbf{z}}_0 - \mathbf{y}_0 = \mathbf{0}$ und damit $(z_j\mathbf{E} - \mathbf{\Phi}_\mathbf{z})\hat{\mathbf{z}}_0 = \mathbf{0}$. Man erkennt nun, dass (2.115) genau dann nichttriviale Lösungen für $\mathbf{y}_0$ und $\hat{\mathbf{z}}_0$ besitzt, wenn gilt $\det(z_j\mathbf{E} - \mathbf{\Phi}_\mathbf{z}) = 0$, womit Satz 2.8 gezeigt ist. □

Hinweis: Dieses Ergebnis zeigt, dass das stationäre Kalman-Filter (2.103) Nullstellen an den Polen des Störmodells (2.101) aufweist. Um nun ein Kalman-Filter zu entwerfen, das bestimmte Frequenzen abblockt, wählt man einfach ein Störmodell mit Polen an diesen Frequenzen. Generell lässt sich sagen, dass je größer die Energie der Störung bei den jeweiligen Frequenzen ist, desto mehr unterdrückt das Kalman-Filter diese Frequenzen.

Aufgabe 2.13. Entwerfen Sie ein Kalman-Filter zur Schätzung der Drehwinkelgeschwindigkeit ω eines permanenterregten Gleichstrommotors auf Basis einer Messung des Drehwinkels φ. Für eine gewisse Wahl der Parameter und unter Vernachlässigung der Dynamik des elektrischen Teilsystems erhält man das Modell des Gleichstrommotors in der Form

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} \varphi \\ \omega \end{bmatrix} &= \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} \varphi \\ \omega \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u \\ y = \varphi &= \begin{bmatrix} 1 & 0 \end{bmatrix}\begin{bmatrix} \varphi \\ \omega \end{bmatrix} \; . \end{aligned}$$

Es ist nun gewünscht, dass Resonanzfrequenzen mit einer Kreisfrequenz ω_0, welche durch die mechanische Last resultieren, vom Kalman-Filter unterdrückt werden. Das

zugehörige Störmodell kann in der Form

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\omega_0^2 & -2\xi\omega_0 \end{bmatrix}\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} 0 \\ \omega_0^2 \end{bmatrix} n$$
$$v = \begin{bmatrix} 1 & 0 \end{bmatrix}\begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$$

mit weißem Rauschen n als Eingangssignal modelliert werden. Entwerfen Sie für diese Spezifikationen ein Kalman-Filter für eine Abtastzeit $T_a = 0.05$ s und untersuchen Sie den Einfluss verschiedener Werte für ξ im Bereich $0.01 < \xi < 0.7$. Wählen Sie dazu $\omega_0 = 3$, $\mathbf{G} = \mathbf{E}$, $\mathbf{Q} = \mathbf{E}$ sowie $\mathbf{R} = 1$. Zeichnen Sie das Bode-Diagramm des Kalman-Filters und testen Sie ihr Kalman-Filter durch Simulation.

2.4 Das Extended Kalman-Filter

Bevor das Extended Kalman-Filter als Beobachter für nichtlineare Systeme besprochen wird, soll im Folgenden das Kalman-Filter von Abschnitt 2.3 für lineare zeitvariante Abtastsysteme der Form

$$\mathbf{x}_{k+1} = \mathbf{\Phi}_k\mathbf{x}_k + \mathbf{\Gamma}_k\mathbf{u}_k + \mathbf{G}_k\mathbf{w}_k \qquad \mathbf{x}(0) = \mathbf{x}_0 \tag{2.116a}$$
$$\mathbf{y}_k = \mathbf{C}_k\mathbf{x}_k + \mathbf{D}_k\mathbf{u}_k + \mathbf{v}_k \tag{2.116b}$$

mit dem n-dimensionalen Zustand $\mathbf{x} \in \mathbb{R}^n$, dem p-dimensionalen deterministischen Eingang $\mathbf{u} \in \mathbb{R}^p$, dem q-dimensionalen Ausgang $\mathbf{y} \in \mathbb{R}^q$, der r-dimensionalen Störung $\mathbf{w} \in \mathbb{R}^r$, dem Messrauschen $\mathbf{v}$ sowie den zeitvarianten Matrizen $\mathbf{\Phi}_k \in \mathbb{R}^{n\times n}$, $\mathbf{\Gamma}_k \in \mathbb{R}^{n\times p}$, $\mathbf{G}_k \in \mathbb{R}^{n\times r}$, $\mathbf{C}_k \in \mathbb{R}^{q\times n}$ und $\mathbf{D}_k \in \mathbb{R}^{q\times p}$ angeschrieben werden. Analog zu Abschnitt 2.3 werden folgende Annahmen getroffen:

(1) Für die Störung $\mathbf{w}$ und das Messrauschen $\mathbf{v}$ wird vorausgesetzt, dass gilt

$$\mathrm{E}(\mathbf{v}_k) = \mathbf{0} \qquad \mathrm{E}\left(\mathbf{v}_k\mathbf{v}_j^{\mathrm{T}}\right) = \mathbf{R}_k\delta_{kj} \tag{2.117a}$$
$$\mathrm{E}(\mathbf{w}_k) = \mathbf{0} \qquad \mathrm{E}\left(\mathbf{w}_k\mathbf{w}_j^{\mathrm{T}}\right) = \mathbf{Q}_k\delta_{kj} \tag{2.117b}$$
$$\mathrm{E}\left(\mathbf{w}_k\mathbf{v}_j^{\mathrm{T}}\right) = \mathbf{0} \tag{2.117c}$$

mit $\mathbf{Q}_k \geq 0$ sowie $\mathbf{R}_k > 0$ und dem Kroneckersymbol $\delta_{kj} = 1$ für $k = j$ und $\delta_{kj} = 0$ sonst.

(2) Der Erwartungswert des Anfangswertes und die Kovarianzmatrix des Anfangsfehlers sind durch

$$\mathrm{E}(\mathbf{x}_0) = \mathbf{m}_0 \qquad \mathrm{E}\left([\mathbf{x}_0 - \hat{\mathbf{x}}_0][\mathbf{x}_0 - \hat{\mathbf{x}}_0]^{\mathrm{T}}\right) = \mathbf{P}_0 \geq 0 \tag{2.118}$$

mit dem Schätzwert $\hat{\mathbf{x}}_0$ des Anfangswertes $\mathbf{x}_0$ gegeben.

(3) Die Störung $\mathbf{w}_k$, $k \geq 0$, und das Messrauschen $\mathbf{v}_l$, $l \geq 0$, sind mit dem Anfangswert $\mathbf{x}_0$ nicht korreliert, d. h. es gilt

$$\mathrm{E}\Big(\mathbf{w}_k \mathbf{x}_0^{\mathrm{T}}\Big) = \mathbf{0} \tag{2.119a}$$

$$\mathrm{E}\Big(\mathbf{v}_l \mathbf{x}_0^{\mathrm{T}}\Big) = \mathbf{0}\ . \tag{2.119b}$$

Die Herleitung des Kalman-Filters für das System (2.116) erfolgt auf vollkommen analoge Art und Weise wie im Abschnitt 2.3 und lautet für $k \geq 0$, vergleiche dazu (2.89)–(2.91),

$$\hat{\mathbf{K}}_k = \mathbf{\Phi}_k \mathbf{P}_k \mathbf{C}_k^{\mathrm{T}} \Big(\mathbf{C}_k \mathbf{P}_k \mathbf{C}_k^{\mathrm{T}} + \mathbf{R}_k\Big)^{-1} \tag{2.120a}$$

$$\hat{\mathbf{x}}_{k+1} = \mathbf{\Phi}_k \hat{\mathbf{x}}_k + \mathbf{\Gamma}_k \mathbf{u}_k + \hat{\mathbf{K}}_k (\mathbf{y}_k - \mathbf{C}_k \hat{\mathbf{x}}_k - \mathbf{D}_k \mathbf{u}_k) \tag{2.120b}$$

$$\mathbf{P}_{k+1} = \mathbf{\Phi}_k \mathbf{P}_k \mathbf{\Phi}_k^{\mathrm{T}} + \mathbf{G}_k \mathbf{Q}_k \mathbf{G}_k^{\mathrm{T}} - \mathbf{\Phi}_k \mathbf{P}_k \mathbf{C}_k^{\mathrm{T}} \Big(\mathbf{C}_k \mathbf{P}_k \mathbf{C}_k^{\mathrm{T}} + \mathbf{R}_k\Big)^{-1} \mathbf{C}_k \mathbf{P}_k \mathbf{\Phi}_k^{\mathrm{T}}\ . \tag{2.120c}$$

Wenn der Anfangswert $\mathbf{x}_0$ bekannt ist, dann setzt man $\hat{\mathbf{x}}_0 = \mathbf{x}_0$ und $\mathbf{P}_0 = \mathbf{0}$ und für den Fall, dass über den Anfangswert keine Information vorliegt, wird typischerweise $\hat{\mathbf{x}}_0 = \mathbf{0}$ und $\mathbf{P}_0 = \alpha \mathbf{E}$ mit $\alpha \gg 1$ gewählt. Es sei darauf hingewiesen, dass sich auch diese Darstellung für $\mathbf{H} \neq \mathbf{0}$ und $\mathrm{E}\Big(\mathbf{w}_k \mathbf{v}_j^{\mathrm{T}}\Big) \neq \mathbf{0}$ analog zu den Diskussionen im letzten Abschnitt verallgemeinern lassen.

In der Literatur ist es oft üblich, das Kalman-Filter in einer etwas anderen Form darzustellen. Dabei wird die optimale Schätzung des Zustandes $\mathbf{x}_k$ und der Fehlerkovarianzmatrix $\mathbf{P}_k$ unter Berücksichtigung von $0, \ldots, k-1$ Messungen, vergleiche Definition 2.3,

$$\hat{\mathbf{x}}_k^- = \hat{\mathbf{x}}(k|k-1) \tag{2.121a}$$

$$\mathbf{P}_k^- = \mathbf{P}(k|k-1) = \mathrm{E}\Bigg(\Big[\mathbf{x}_k - \hat{\mathbf{x}}_k^-\Big]\Big[\mathbf{x}_k - \hat{\mathbf{x}}_k^-\Big]^{\mathrm{T}}\Bigg) \tag{2.121b}$$

als *a priori Schätzung* und die optimale Schätzung von $\mathbf{x}_k$ und $\mathbf{P}_k$ unter Berücksichtigung von $0, \ldots, k$ Messungen

$$\hat{\mathbf{x}}_k^+ = \hat{\mathbf{x}}(k|k) \tag{2.122a}$$

$$\mathbf{P}_k^+ = \mathbf{P}(k|k) = \mathrm{E}\Bigg(\Big[\mathbf{x}_k - \hat{\mathbf{x}}_k^+\Big]\Big[\mathbf{x}_k - \hat{\mathbf{x}}_k^+\Big]^{\mathrm{T}}\Bigg) \tag{2.122b}$$

als *a posteriori Schätzung* bezeichnet. (2.120) kann damit in der äquivalenten Form

Kalman Verstärkungsmatrix:	$\hat{\mathbf{L}}_k = \mathbf{P}_k^- \mathbf{C}_k^{\mathrm{T}} \Big(\mathbf{C}_k \mathbf{P}_k^- \mathbf{C}_k^{\mathrm{T}} + \mathbf{R}_k\Big)^{-1}$	(2.123a)
Zustandsschätzung Update:	$\hat{\mathbf{x}}_k^+ = \hat{\mathbf{x}}_k^- + \hat{\mathbf{L}}_k \Big(\mathbf{y}_k - \mathbf{C}_k \hat{\mathbf{x}}_k^- - \mathbf{D}_k \mathbf{u}_k\Big)$	(2.123b)
Fehlerkovarianz Update:	$\mathbf{P}_k^+ = \Big(\mathbf{E} - \hat{\mathbf{L}}_k \mathbf{C}_k\Big) \mathbf{P}_k^-$	(2.123c)
Zustandsextrapolation (2.87):	$\hat{\mathbf{x}}_{k+1}^- = \mathbf{\Phi}_k \hat{\mathbf{x}}_k^+ + \mathbf{\Gamma}_k \mathbf{u}_k$	(2.123d)
Fehlerkovarianzextrapolation (2.88):	$\mathbf{P}_{k+1}^- = \mathbf{\Phi}_k \mathbf{P}_k^+ \mathbf{\Phi}_k^{\mathrm{T}} + \mathbf{G}_k \mathbf{Q}_k \mathbf{G}_k^{\mathrm{T}}$	(2.123e)

für $k \geq 0$ und die Anfangswerte $\hat{\mathbf{x}}_0^- = \hat{\mathbf{x}}_0$ und $\mathbf{P}_0^- = \mathbf{P}_0$, angeschrieben werden.

Aufgabe 2.14. Zeigen Sie die Äquivalenz der Beziehungen (2.120) und (2.123). Führen Sie dazu in (2.123) folgende Substitutionen durch $\hat{\mathbf{x}}_{k+1}^- = \hat{\mathbf{x}}_{k+1}$, $\hat{\mathbf{x}}_k^- = \hat{\mathbf{x}}_k$, $\mathbf{P}_{k+1}^- = \mathbf{P}_{k+1}$, $\mathbf{P}_k^- = \mathbf{P}_k$ und $\mathbf{\Phi}_k \hat{\mathbf{L}}_k = \hat{\mathbf{K}}_k$.

Dem Extended Kalman-Filter (EKF) Entwurf liegt im Allgemeinen ein nichtlineares, zeitvariantes, zeitkontinuierliches Mehrgrößensystem der Form

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x} = \mathbf{f}(\mathbf{x}, \mathbf{u}, \mathbf{w}, t) \qquad \mathbf{x}(0) = \mathbf{x}_0 \tag{2.124a}$$

$$\mathbf{y} = \mathbf{h}(\mathbf{x}, \mathbf{u}, \mathbf{v}, t) \tag{2.124b}$$

mit dem n-dimensionalen Zustand $\mathbf{x} \in \mathbb{R}^n$, dem p-dimensionalen deterministischen Eingang $\mathbf{u} \in \mathbb{R}^p$, dem q-dimensionalen Ausgang $\mathbf{y} \in \mathbb{R}^q$, der r-dimensionalen Störung $\mathbf{w} \in \mathbb{R}^r$ und dem Messrauschen $\mathbf{v}$ zugrunde. Da das Kalman-Filter im Normalfall in einem Digitalrechner implementiert wird, die Stellgrößen über ein Halteglied nullter Ordnung (D/A-Wandler) mit der Abtastzeit T_a auf den Prozess aufgeschaltet und die Messgrößen mit der Abtastzeit T_a über einen A/D-Wandler abgetastet werden, muss das zu (2.124) zugehörige Abtastsystem

$$\mathbf{x}_{k+1} = \mathbf{F}_k(\mathbf{x}_k, \mathbf{u}_k, \mathbf{w}_k) \qquad \mathbf{x}(0) = \mathbf{x}_0 \tag{2.125a}$$

$$\mathbf{y}_k = \mathbf{h}_k(\mathbf{x}_k, \mathbf{u}_k, \mathbf{v}_k) \tag{2.125b}$$

berechnet werden. Aus der Vorlesung Automatisierung (Abschnitt 6.2.1) ist bekannt, dass zur Bestimmung des Abtastsystems die exakte Lösung von (2.124) notwendig ist.

Hinweis: Die Lösung eines nichtlinearen Differentialgleichungssystems der Form (2.124) ist bekanntermaßen nur in Spezialfällen möglich ist. Daher wird im Folgenden eine Näherungslösung mit Hilfe eines Integrationsverfahrens gesucht. Dazu nimmt man an, dass die Stellgröße $\mathbf{u}(t)$ und die Störung $\mathbf{w}(t)$ für das Abtastintervall $kT_a \leq t < (k+1)T_a$ konstant sind, d. h. $\mathbf{u}(t) = \mathbf{u}(kT_a) = \mathbf{u}_k$ und $\mathbf{w}(t) = \mathbf{w}(kT_a) = \mathbf{w}_k$, und integriert die Differenzialgleichung (2.124a) im Abtastintervall

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \int_{kT_a}^{(k+1)T_a} \mathbf{f}(\mathbf{x}(t), \mathbf{u}_k, \mathbf{w}_k, t)\,\mathrm{d}t \tag{2.126}$$

mit $\mathbf{x}_{k+1} = \mathbf{x}((k+1)T_a)$ und $\mathbf{x}_k = \mathbf{x}(kT_a)$. Die Approximation des Integrals in (2.126) kann auf unterschiedliche Arten erfolgen. Im Weiteren sollen lediglich zwei mögliche Lösungen angegeben werden:

(1) *Euler-Verfahren*

$$\int_{kT_a}^{(k+1)T_a} \mathbf{f}(\mathbf{x}(t), \mathbf{u}_k, \mathbf{w}_k, t)\,\mathrm{d}t = \mathbf{f}(\mathbf{x}_k, \mathbf{u}_k, \mathbf{w}_k, kT_a)T_a \tag{2.127}$$

(2) *Runge-Kutta Verfahren vierter Ordnung*

$$\int_{kT_a}^{(k+1)T_a} \mathbf{f}(\mathbf{x}(t), \mathbf{u}_k, \mathbf{w}_k, t)\,\mathrm{d}t = \frac{\Delta\mathbf{x}_1 + 2\Delta\mathbf{x}_2 + 2\Delta\mathbf{x}_3 + \Delta\mathbf{x}_4}{6} \tag{2.128}$$

mit

$$\begin{aligned}
\Delta\mathbf{x}_1 &= \mathbf{f}(\mathbf{x}_k, \mathbf{u}_k, \mathbf{w}_k, kT_a)T_a \\
\Delta\mathbf{x}_2 &= \mathbf{f}\left(\mathbf{x}_k + \frac{\Delta\mathbf{x}_1}{2}, \mathbf{u}_k, \mathbf{w}_k, kT_a + \frac{T_a}{2}\right)T_a \\
\Delta\mathbf{x}_3 &= \mathbf{f}\left(\mathbf{x}_k + \frac{\Delta\mathbf{x}_2}{2}, \mathbf{u}_k, \mathbf{w}_k, kT_a + \frac{T_a}{2}\right)T_a \\
\Delta\mathbf{x}_4 &= \mathbf{f}(\mathbf{x}_k + \Delta\mathbf{x}_3, \mathbf{u}_k, \mathbf{w}_k, (k+1)T_a)T_a\ .
\end{aligned} \tag{2.129}$$

Dabei wurde für $\Delta\mathbf{x}_4$ in (2.129) der linksseitige Grenzwert $\lim_{t\to(k+1)T_a} \mathbf{u}(t) = \mathbf{u}_k$ und $\lim_{t\to(k+1)T_a} \mathbf{w}(t) = \mathbf{w}_k$ eingesetzt.

Die Ausgangsgleichung (2.125b) des Abtastsystems erhält man für beide Fälle sehr einfach, indem man die Ausgangsgleichung (2.124b) des zeitkontinuierlichen Mehrgrößensystems für $t = kT_a$ auswertet

$$\mathbf{y}_k = \mathbf{h}(\mathbf{x}_k, \mathbf{u}_k, \mathbf{v}_k, kT_a) = \mathbf{h}_k(\mathbf{x}_k, \mathbf{u}_k, \mathbf{v}_k)\ . \tag{2.130}$$

Nimmt man nun an, dass durch die Beziehungen (2.126)–(2.130) ein Abtastsystem der Form (2.125) gegeben ist, dann beruht die *Idee des Extended Kalman-Filters* darauf, dass für die rechte Seite von (2.125a) eine Taylor-Reihenentwicklung um den Punkt $\mathbf{x}_k = \hat{\mathbf{x}}_k^+$, $\mathbf{u}_k = \mathbf{u}_k$ und $\mathbf{w}_k = \mathbf{0}$ durchgeführt und nach dem linearen Term abgebrochen wird, d. h.

$$\mathbf{x}_{k+1} = \mathbf{F}_k\left(\hat{\mathbf{x}}_k^+, \mathbf{u}_k, \mathbf{0}\right) + \frac{\partial}{\partial\mathbf{x}_k}\mathbf{F}_k\left(\hat{\mathbf{x}}_k^+, \mathbf{u}_k, \mathbf{0}\right)\left(\mathbf{x}_k - \hat{\mathbf{x}}_k^+\right) + \frac{\partial}{\partial\mathbf{w}_k}\mathbf{F}_k\left(\hat{\mathbf{x}}_k^+, \mathbf{u}_k, \mathbf{0}\right)\mathbf{w}_k\ . \tag{2.131}$$

Analog dazu wird die rechte Seite der Ausgangsgleichung (2.125b) in eine Taylor-Reihe um den Punkt $\mathbf{x}_k = \hat{\mathbf{x}}_k^-$, $\mathbf{u}_k = \mathbf{u}_k$ und $\mathbf{v}_k = \mathbf{0}$ entwickelt und nach dem linearen Term abgebrochen

$$\mathbf{y}_k = \mathbf{h}_k\left(\hat{\mathbf{x}}_k^-, \mathbf{u}_k, \mathbf{0}\right) + \frac{\partial}{\partial\mathbf{x}_k}\mathbf{h}_k\left(\hat{\mathbf{x}}_k^-, \mathbf{u}_k, \mathbf{0}\right)\left(\mathbf{x}_k - \hat{\mathbf{x}}_k^-\right) + \frac{\partial}{\partial\mathbf{v}_k}\mathbf{h}_k\left(\hat{\mathbf{x}}_k^-, \mathbf{u}_k, \mathbf{0}\right)\mathbf{v}_k\ . \tag{2.132}$$

Man beachte, dass hierbei durchgehend folgende vereinfachte Schreibweise

$$\frac{\partial}{\partial\mathbf{x}_k}\mathbf{F}_k\left(\hat{\mathbf{x}}_k^-, \mathbf{u}_k, \mathbf{0}\right) = \left.\frac{\partial}{\partial\mathbf{x}_k}\mathbf{F}_k(\mathbf{x}_k, \mathbf{u}_k, \mathbf{w}_k)\right|_{\mathbf{x}_k=\hat{\mathbf{x}}_k^-, \mathbf{u}_k=\mathbf{u}_k, \mathbf{w}_k=\mathbf{0}} \tag{2.133}$$

verwendet wurde. Die Beziehungen (2.131) und (2.132) können für die weiteren Betrachtungen kompakter in der Form

$$\mathbf{x}_{k+1} = \mathbf{\Phi}_k\mathbf{x}_k + \bar{\mathbf{u}}_k + \mathbf{G}_k\mathbf{w}_k \tag{2.134a}$$

$$\mathbf{y}_k = \mathbf{C}_k\mathbf{x}_k + \breve{\mathbf{u}}_k + \breve{\mathbf{v}}_k \tag{2.134b}$$

mit

$$\begin{aligned}
\mathbf{\Phi}_k &= \frac{\partial}{\partial \mathbf{x}_k}\mathbf{F}_k\Big(\hat{\mathbf{x}}_k^+, \mathbf{u}_k, \mathbf{0}\Big) & \bar{\mathbf{u}}_k &= \mathbf{F}_k\Big(\hat{\mathbf{x}}_k^+, \mathbf{u}_k, \mathbf{0}\Big) - \mathbf{\Phi}_k\hat{\mathbf{x}}_k^+ \\
\mathbf{G}_k &= \frac{\partial}{\partial \mathbf{w}_k}\mathbf{F}_k\Big(\hat{\mathbf{x}}_k^+, \mathbf{u}_k, \mathbf{0}\Big) & \mathbf{C}_k &= \frac{\partial}{\partial \mathbf{x}_k}\mathbf{h}_k\Big(\hat{\mathbf{x}}_k^-, \mathbf{u}_k, \mathbf{0}\Big) \\
\breve{\mathbf{u}}_k &= \mathbf{h}_k\Big(\hat{\mathbf{x}}_k^-, \mathbf{u}_k, \mathbf{0}\Big) - \mathbf{C}_k\hat{\mathbf{x}}_k^- & \breve{\mathbf{v}}_k &= \frac{\partial}{\partial \mathbf{v}_k}\mathbf{h}_k\Big(\hat{\mathbf{x}}_k^-, \mathbf{u}_k, \mathbf{0}\Big)\mathbf{v}_k
\end{aligned} \tag{2.135}$$

geschrieben werden. Es ist nun offensichtlich, dass die Struktur des Systems (2.134) unmittelbar die Anwendung des Kalman-Filters gemäß (2.123) ermöglicht. Die für die Implementierung erforderlichen Rechenschritte werden im Folgenden nochmals zusammengefasst.

(1) Für das nichtlineare, zeitvariante, zeitkontinuierliche Mehrgrößensystem (2.124) berechne man ein Abtastsystem der Form (2.125).

(2) Der geschätzte Zustand und die Kovarianzmatrix des Schätzfehlers müssen für den Anfangszeitpunkt mit $\hat{\mathbf{x}}_0^-$ und $\mathbf{P}_0^-$ initialisiert werden.

(3) Von der Störung $\mathbf{w}_k$ und vom Messrauschen $\breve{\mathbf{v}}_k$ in (2.134) wird wiederum vorausgesetzt, dass gilt

$$\mathrm{E}(\breve{\mathbf{v}}_k) = \mathbf{0} \qquad \mathrm{E}\Big(\breve{\mathbf{v}}_k\breve{\mathbf{v}}_j^{\mathrm{T}}\Big) = \mathbf{R}_k\delta_{kj} \tag{2.136a}$$

$$\mathrm{E}(\mathbf{w}_k) = \mathbf{0} \qquad \mathrm{E}\Big(\mathbf{w}_k\mathbf{w}_j^{\mathrm{T}}\Big) = \mathbf{Q}_k\delta_{kj} \tag{2.136b}$$

$$\mathrm{E}\Big(\mathbf{w}_k\breve{\mathbf{v}}_j^{\mathrm{T}}\Big) = \mathbf{0} \tag{2.136c}$$

mit $\mathbf{Q}_k \geq 0$ sowie $\mathbf{R}_k > 0$ und dem Kroneckersymbol $\delta_{kj} = 1$ für $k = j$ und $\delta_{kj} = 0$ sonst.

(4) Die Iterationsgleichungen des Extended Kalman-Filters lauten dann für $k \geq 0$

$$\mathbf{C}_k = \frac{\partial}{\partial \mathbf{x}_k}\mathbf{h}_k\Big(\hat{\mathbf{x}}_k^-, \mathbf{u}_k, \mathbf{0}\Big) \tag{2.137a}$$

$$\hat{\mathbf{L}}_k = \mathbf{P}_k^-\mathbf{C}_k^{\mathrm{T}}\Big(\mathbf{C}_k\mathbf{P}_k^-\mathbf{C}_k^{\mathrm{T}} + \mathbf{R}_k\Big)^{-1} \tag{2.137b}$$

$$\hat{\mathbf{x}}_k^+ = \hat{\mathbf{x}}_k^- + \hat{\mathbf{L}}_k\Big(\mathbf{y}_k - \mathbf{C}_k\hat{\mathbf{x}}_k^- - \breve{\mathbf{u}}_k\Big) = \hat{\mathbf{x}}_k^- + \hat{\mathbf{L}}_k\Big(\mathbf{y}_k - \mathbf{h}_k\Big(\hat{\mathbf{x}}_k^-, \mathbf{u}_k, \mathbf{0}\Big)\Big) \tag{2.137c}$$

$$\mathbf{P}_k^+ = \Big(\mathbf{E} - \hat{\mathbf{L}}_k\mathbf{C}_k\Big)\mathbf{P}_k^- \tag{2.137d}$$

$$\mathbf{\Phi}_k = \frac{\partial}{\partial \mathbf{x}_k}\mathbf{F}_k\Big(\hat{\mathbf{x}}_k^+, \mathbf{u}_k, \mathbf{0}\Big) \tag{2.137e}$$

$$\mathbf{G}_k = \frac{\partial}{\partial \mathbf{w}_k}\mathbf{F}_k\Big(\hat{\mathbf{x}}_k^+, \mathbf{u}_k, \mathbf{0}\Big) \tag{2.137f}$$

$$\hat{\mathbf{x}}_{k+1}^- = \mathbf{\Phi}_k\hat{\mathbf{x}}_k^+ + \underbrace{\mathbf{F}_k\Big(\hat{\mathbf{x}}_k^+, \mathbf{u}_k, \mathbf{0}\Big) - \mathbf{\Phi}_k\hat{\mathbf{x}}_k^+}_{\bar{\mathbf{u}}_k} = \mathbf{F}_k\Big(\hat{\mathbf{x}}_k^+, \mathbf{u}_k, \mathbf{0}\Big) \tag{2.137g}$$

$$\mathbf{P}_{k+1}^- = \mathbf{\Phi}_k\mathbf{P}_k^+\mathbf{\Phi}_k^{\mathrm{T}} + \mathbf{G}_k\mathbf{Q}_k\mathbf{G}_k^{\mathrm{T}}\ . \tag{2.137h}$$

Hinweis: Beim Extended Kalman-Filter Entwurf wird angenommen, dass die linearisierte Transformation von Mittelwert und Kovarianz in guter Genauigkeit dem Mittelwert und der Kovarianz der nichtlinearen Transformation entspricht. Diese Annahme trifft im Allgemeinen nicht zu, weshalb zur Verbesserung des Beobachterentwurfes für nichtlineare Systeme häufig das im nächsten Abschnitt betrachtete *Unscented Kalman-Filter* verwendet wird.

Aufgabe 2.15. Das mathematische Modell

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}x_1 &= x_2 + w_1\\ \frac{\mathrm{d}}{\mathrm{d}t}x_2 &= \frac{1}{2}\rho_0\exp\Big(-\frac{x_1}{k}\Big)C_w\frac{A}{m}x_2^2 - g + w_2\end{aligned}$$

beschreibt den freien Fall eines Körpers der Masse m und der Querschnittsfläche A in der Erdatmosphäre mit der Höhe x_1 und der Geschwindigkeit x_2. Der Term $\rho_0\exp(-x_1/k)$ entspricht dabei der höhenabhängigen Dichte in der Atmosphäre (ρ_0 Dichte auf Meereshöhe), womit der Term mit x_2^2 die Verzögerung durch den Luftwiderstand mit dem Luftwiderstandsbeiwert C_w beschreibt. Weiterhin stellt g die Erdbeschleunigung dar.

Das Prozessrauschen ist durch die stochastischen Größen w_1 und w_2 gegeben. Die Höhe x_1 kann über die Ausgangsgleichung

$$y = x_1 + v$$

mit dem Messrauschen v ermittelt werden.

Entwerfen Sie für die Parameter $\rho_0 = 1.2\,\mathrm{kg/m^3}$, $g = 9.81\,\mathrm{m/s^2}$, $k = 9100\,\mathrm{m}$, $A = 0.5\,\mathrm{m^2}$, $m = 100\,\mathrm{kg}$, und $C_w = 0.5$ ein Extended Kalman-Filter, das neben der Höhe x_1 und der Geschwindigkeit x_2 auch den konstanten Luftwiderstandsbeiwert

C_w schätzt. Erweitern Sie dazu das Differentialgleichungssystem um den Zustand $x_3 = C_w$ mit

$$\frac{\mathrm{d}}{\mathrm{d}t}x_3 = 0 + w_3$$

und der Prozessstörungskomponente w_3. Verwenden Sie zur Bestimmung des Abtastsystems das Euler-Verfahren. Nehmen Sie an, dass die nominellen Werte bzw. Anfangsbedingungen der Größen C_w, x_1 und x_2 normalverteilt sind. Die entsprechenden Werte der Mittelwerte und Varianzen können der nachfolgenden Tabelle 2.1 entnommen werden. Für die Simulation nehmen Sie folgende Werte an: $C_w = 0.6$, $x_1(0) = 39\,500\,\mathrm{m}$ und $x_2(0) = -10\,\mathrm{m/s}$.

Variable	Mittelwert	Varianz
C_w	0.5	1
$x_1(0)$	$39 \cdot 10^3\,\mathrm{m}$	$1 \cdot 10^4\,\mathrm{m}^2$
$x_2(0)$	$0\,\mathrm{m/s}$	$1\,\mathrm{m}^2/\mathrm{s}^2$

Tabelle 2.1: Mittelwerte und Varianzen der Parameter bzw. der Anfangsbedingungen.

Aufgabe 2.16. Zur Positionsbestimmung eines Fahrzeuges in einem zweidimensionalen Raum (o-Achse: Ost-Koordinate, n-Achse: Nord-Koordinate) soll ein Extended Kalman-Filter eingesetzt werden. Mehrere Messstationen mit den Koordinaten (O_i, N_i), $i = 1, \ldots, M$ messen den Abstand zum Fahrzeug. Die Beschleunigung des Fahrzeugs in Nord- und Ost-Richtung wird durch weißes Rauschen modelliert. Das Differenzengleichungssystem

$$\underbrace{\begin{bmatrix} o_{k+1} \\ n_{k+1} \\ o_{v,k+1} \\ n_{v,k+1} \end{bmatrix}}_{\mathbf{x}_{k+1}} = \underbrace{\begin{bmatrix} 1 & 0 & T_a & 0 \\ 0 & 1 & 0 & T_a \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{\mathbf{\Phi}} \underbrace{\begin{bmatrix} o_k \\ n_k \\ o_{v,k} \\ n_{v,k} \end{bmatrix}}_{\mathbf{x}_k} + \underbrace{\begin{bmatrix} w_{1,k} \\ w_{2,k} \\ w_{3,k} \\ w_{4,k} \end{bmatrix}}_{\mathbf{w}_k}$$

beschreibt das Fahrzeugverhalten, wobei o_k und n_k bzw. $o_{v,k}$ und $n_{v,k}$ die Koordinaten bzw. Geschwindigkeiten des Fahrzeuges bezogen auf den Ursprung eines ortsfesten Koordinatensystems in Ost- und Nordrichtung zum Zeitpunkt kT_a mit der Abtastzeit T_a bezeichnen und $w_{j,k}$, $j = 1, \ldots, 4$ die Komponenten des Prozessrauschens sind. Im Weiteren sind durch

$$y_{i,k} = \sqrt{(n_k - N_i)^2 + (o_k - O_i)^2} + v_{i,k}, \quad i = 1, \ldots, M$$

mit dem Messrauschen $v_{i,k}$, $i = 1, \ldots, M$ die Entfernungsmessungen des Fahrzeuges von den Stationen gegeben. Nehmen Sie an, dass alle stochastischen Größen $(w_{j,k})$ und $(v_{i,k})$ normalverteilt, unkorreliert und mittelwertfrei sind. Die Abtastzeit sei mit

$T_a = 0.1\,\mathrm{s}$ gegeben. Für die Kovarianzmatrix des Prozessrauschens gelte

$$\mathrm{E}\left(\mathbf{w}_k \mathbf{w}_j^{\mathrm{T}}\right) = \mathbf{Q}\delta_{kj} \quad \text{mit} \quad \mathbf{Q} = \mathrm{diag}(0, 0, 4, 4)$$

und die Kovarianz des Messrauschens sei

$$\mathrm{E}(v_{i,k} v_{i,j}) = R_i \delta_{kj} \quad \text{mit} \quad R_i = 1, \quad i = 1, \ldots, M \ .$$

Der Anfangszustand $\mathbf{x}_0^{\mathrm{T}} = \begin{bmatrix} 0 & 0 & 50 & 50 \end{bmatrix}$ ist exakt bekannt. Simulieren Sie das System für 60 Sekunden und entwerfen Sie ein Extended Kalman-Filter zur Schätzung der Zustände. Variieren Sie die Anzahl und Position der Messstationen.

2.5 Das Unscented Kalman-Filter

Das im letzten Kapitel behandelte Extended Kalman-Filter ist der (industrielle) Standard für die Zustandsschätzung von nichtlinearen dynamischen Systemen. Das Extended Kalman-Filter besitzt jedoch den Nachteil, dass es unzuverlässige Schätzungen (des Erwartungswerts und der Fehlerkovarianz) liefern kann, wenn das System ausgeprägte Nichtlinearitäten aufweist. Diese Unzuverlässigkeit resultiert aus der Linearisierung der nichtlinearen Systemdynamik, welche für die Berechnung des Erwartungswerts und der Kovarianz des Zustands verwendet wird. Um diese Problematik genauer darzustellen, wird im nächsten Abschnitt gezeigt, wie sich der Erwartungswert und die Kovarianz einer Zufallsvariable durch eine nichtlineare Transformation verändert.

2.5.1 Erwartungswert und Kovarianz von nichtlinearen Transformationen

Im Folgenden wird das Beispiel einer Transformation von Zylinderkoordinaten in kartesische Koordinaten betrachtet

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \mathbf{h}(\mathbf{x}) = \begin{bmatrix} x_1 \cos(x_2) \\ x_1 \sin(x_2) \end{bmatrix}. \tag{2.138}$$

Dabei wird angenommen, dass der Zufallsvektor $\mathbf{x} = [x_1, x_2]^{\mathrm{T}}$ durch die nicht korrelierten, gleichverteilten Zufallsvariabeln x_1 (Intervall: $a_1 = m_{x1} - \delta_{x1} = 1 - 0.01$, $b_1 = m_{x1} + \delta_{x1} = 1 + 0.01$) und x_2 (Intervall: $a_2 = m_{x2} - \delta_{x2} = \frac{\pi}{2} - 0.35$, $b_2 = m_{x2} + \delta_{x2} = \frac{\pi}{2} + 0.35$) definiert ist. In Abbildung 2.7 sind 10000 Zufallsvektoren, die nach dieser Verteilung generiert wurden, dargestellt. Der Erwartungswert (Mittelwert) $\mathbf{m}_x$ ergibt sich zu $\mathbf{m}_x = [1, \pi/2]^{\mathrm{T}}$ und die Kovarianzmatrix errechnet sich zu

$$\mathbf{Q}_x = \begin{bmatrix} \frac{1}{3}10^{-4} & 0 \\ 0 & \frac{49}{12}10^{-2} \end{bmatrix}. \tag{2.139}$$

In Abbildung 2.7 ist die Kovarianzmatrix durch die Ellipse $(\mathbf{x} - \mathbf{m}_x)^{\mathrm{T}} \mathbf{Q}_x^{-1} (\mathbf{x} - \mathbf{m}_x) = 1$ dargestellt.

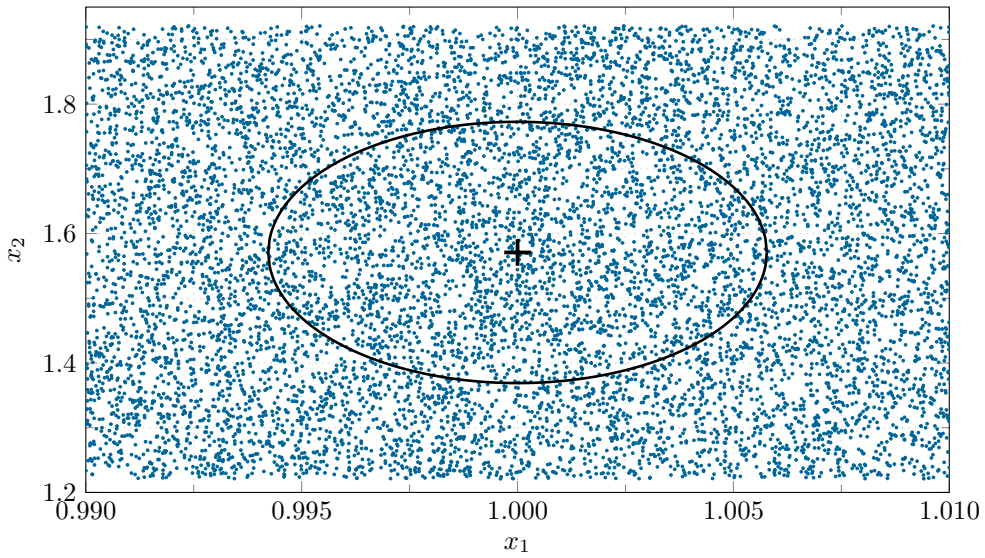


Abbildung 2.7: Verteilung der Zufallsvektoren $\mathbf{x}$ mit zugehörigem Erwartungswert $\mathbf{m}_x$ und Kovarianzmatrix $\mathbf{Q}_x$.

Der Erwartungswert $\mathrm{E}(\mathbf{y}) = \mathbf{m}_y$ der mittels der nichtlinearen Transformation berechneten Größe $\mathbf{y} = \mathbf{h}(\mathbf{x})$ ergibt sich aus

$$\begin{bmatrix} m_{y1} \\ m_{y2} \end{bmatrix} = \begin{bmatrix} \mathrm{E}(h_1(\mathbf{x})) \\ \mathrm{E}(h_2(\mathbf{x})) \end{bmatrix} = \begin{bmatrix} \mathrm{E}(x_1 \cos(x_2)) \\ \mathrm{E}(x_1 \sin(x_2)) \end{bmatrix} . \tag{2.140}$$

Berücksichtigt man die Unabhängigkeit der Zufallsvariablen x_1 und x_2 und führt man die Aufspaltung $\mathbf{x} = \tilde{\mathbf{x}} + \mathbf{m}_x$ ein, so erhält man für die erste Zeile von (2.140)

$$\mathrm{E}(x_1 \cos(x_2)) = \mathrm{E}(x_1)\, \mathrm{E}(\cos(\tilde{x}_2 + m_{x2})) . \tag{2.141}$$

Es gilt $\mathrm{E}(x_1) = m_{x1} = 1$. Spaltet man den cos-Term in (2.141) auf und setzt den Erwartungswert $m_{x2} = \pi/2$ ein, so erhält man

$$\mathrm{E}(\cos(\tilde{x}_2 + m_{x2})) = \mathrm{E}(\cos(\tilde{x}_2)\cos(m_{x2}) - \sin(\tilde{x}_2)\sin(m_{x2})) = -\,\mathrm{E}(\sin(\tilde{x}_2)) . \tag{2.142}$$

Da der Erwartungswert einer schiefsymmetrischen Funktion einer Zufallsvariable mit symmetrischer Verteilungsdichtefunktion verschwindet, gilt $\mathrm{E}(\sin(\tilde{x}_2)) = 0$ und somit auch $m_{y1} = 0$. Um diesen letzten Schritt zu zeigen, betrachte man

$$\mathrm{E}(\sin(\tilde{x}_2)) = \int_{-\delta_{x2}}^{\delta_{x2}} \sin(\tilde{x}_2) \frac{1}{2\delta_{x2}} \,\mathrm{d}\tilde{x}_2 = \frac{1}{2\delta_{x2}} (-\cos(\delta_{x2}) + \cos(-\delta_{x2})) = 0 . \tag{2.143}$$

Auf analoge Art kann man die zweite Zeile von (2.140) zu $\mathrm{E}(x_1 \sin(x_2)) = \sin(\delta_{x2})/\delta_{x2}$ berechnen. Zusammengefasst erhält man den exakten Erwartungswert $\mathbf{m}_y$ der Zufallsvariablen $\mathbf{y}$ in der Form

$$\mathbf{m}_y = \begin{bmatrix} 0 \\ \frac{\sin(\delta_{x2})}{\delta_{x2}} \end{bmatrix} . \tag{2.144}$$

Die exakte Berechnung des Erwartungswerts $\mathbf{m}_y$ ist nur für einige wenige Sonderfälle möglich. Man beachte, dass in diesem Beispiel eine einfache Gleichverteilung und eine einfache nichtlineare Transformation angenommen wurde um eine analytische Lösung berechnen zu können. Eine naheliegende Möglichkeit den Erwartungswert näherungsweise zu berechnen besteht in der Approximation der nichtlinearen Transformation $\mathbf{h}(\mathbf{x})$ durch eine Taylorreihe, die um den Erwartungswert $\mathbf{m}_x$ des Zufallsvektors $\mathbf{x}$ ermittelt wurde.

$$\mathbf{h}(\mathbf{x}) = \mathbf{h}(\mathbf{m}_x) + \mathrm{D}_{\tilde{\mathbf{x}}}\mathbf{h} + \frac{1}{2!}\mathrm{D}^2_{\tilde{\mathbf{x}}}\mathbf{h} + \frac{1}{3!}\mathrm{D}^3_{\tilde{\mathbf{x}}}\mathbf{h} + \ldots, \tag{2.145}$$

angegeben werden, wobei die Schreibweise

$$\mathrm{D}^k_{\tilde{\mathbf{x}}}\mathbf{h} = \left(\sum_{i=1}^{n} \tilde{x}_i \frac{\partial}{\partial x_i}\right)^k \mathbf{h}(\mathbf{x})\bigg|_{\mathbf{x}=\mathbf{m}_x} \tag{2.146}$$

verwendet wird. Darin bezeichnet n die Dimension der Zufallsvariable $\mathbf{x}$.

Verwendet man zur Berechnung des Erwartungswerts eine Taylorreihenapproximation von $\mathbf{h}(\mathbf{x})$ der Ordnung 1, so erhält man

$$\mathbf{h}(\mathbf{x}) \approx \mathbf{h}_l(\mathbf{x}) = \mathbf{h}(\mathbf{m}_x) + \tilde{x}_1 \frac{\partial \mathbf{h}(\mathbf{x})}{\partial x_1}\bigg|_{\mathbf{x}=\mathbf{m}_x} + \tilde{x}_2 \frac{\partial \mathbf{h}(\mathbf{x})}{\partial x_2}\bigg|_{\mathbf{x}=\mathbf{m}_x}. \tag{2.147}$$

Der zugehörige Erwartungswert berechnet sich damit zu

$$\mathrm{E}(\mathbf{h}_l(\mathbf{x})) = \mathbf{h}(\mathbf{m}_x) + \frac{\partial \mathbf{h}(\mathbf{x})}{\partial x_1}\bigg|_{\mathbf{x}=\mathbf{m}_x} \mathrm{E}(\tilde{x}_1) + \frac{\partial \mathbf{h}(\mathbf{x})}{\partial x_2}\bigg|_{\mathbf{x}=\mathbf{m}_x} \mathrm{E}(\tilde{x}_2) = \mathbf{h}(\mathbf{m}_x) \tag{2.148}$$

und damit

$$\mathrm{E}(\mathbf{h}_l(\mathbf{x})) = \mathbf{m}_{yl} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{2.149}$$

Ein Vergleich mit dem exakten Ergebnis (2.144) zeigt, dass die Fehler mit steigendem δ_{x2} anwachsen. In Abbildung 2.8 ist ein Vergleich des exakten Erwartungswerts $\mathbf{m}_y$ mit dem Erwartungswert $\mathbf{m}_{yl}$, welcher auf Basis der Linearisierung ermittelt wurde, dargestellt. Da diese Linearisierung auch inhärent in der Berechnung eines EKF inkludiert ist, muss auch für ein EKF eine Abweichung vom exakten Ergebnis erwartet werden. Diese Abweichung wächst mit steigender Nichtlinearität der Strecke und mit größerer Kovarianz der Zufallsvariablen an.

Eine Möglichkeit den Fehler zufolge der Linearisierung zu verringern, ist die Verwendung einer Approximation $\mathbf{h}_q(\mathbf{x})$ von $\mathbf{h}(\mathbf{x})$ zweiter Ordnung. Für diesen Fall ergibt sich der approximierte Erwartungswert

$$\mathrm{E}(\mathbf{h}_q(\mathbf{x})) = \mathbf{m}_{yq} = \begin{bmatrix} 0 \\ 1 - \frac{\sigma^2_{x2}}{2} \end{bmatrix}, \tag{2.150}$$

mit der Varianz σ_{x2} von x_2. In Abbildung 2.8 erkennt man die erwartete Verbesserung der Approximation von $\mathrm{E}(\mathbf{h}(\mathbf{x}))$ durch $\mathrm{E}(\mathbf{h}_q(\mathbf{x}))$ im Vergleich zur linearen Approximation $\mathrm{E}(\mathbf{h}_l(\mathbf{x}))$.

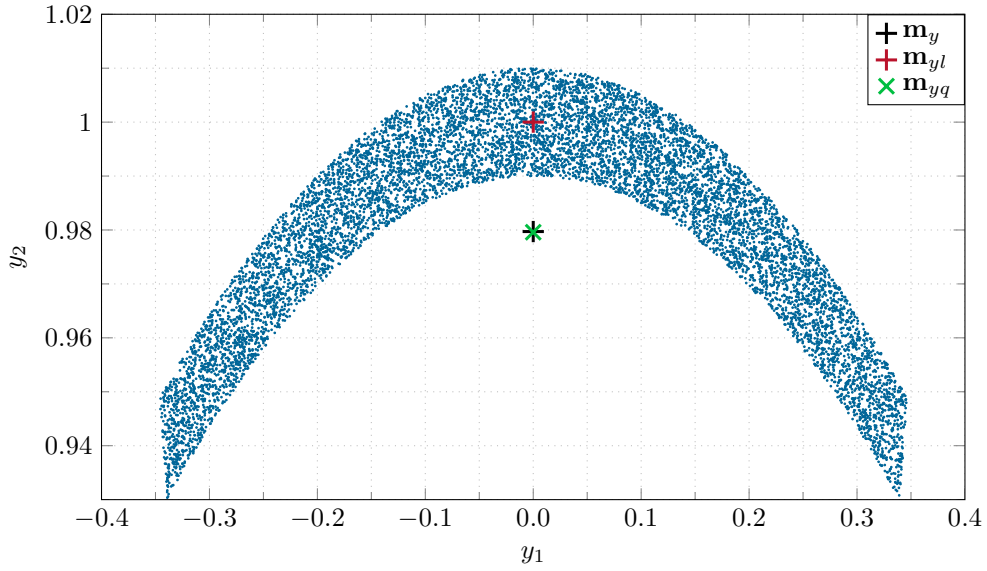


Abbildung 2.8: Verteilung der Zufallsvektoren $\mathbf{y} = \mathbf{h}(\mathbf{x})$ mit dem Erwartungswert $\mathrm{E}(\mathbf{y}) = \mathbf{m}_y$, der Approximation $\mathbf{m}_{yl}$ auf Basis einer Taylorreihe 1. Ordnung und der Approximation $\mathbf{m}_{yq}$ auf Basis einer Taylorreihe 2. Ordnung.

Aufgabe 2.17. Zeigen Sie das Ergebnis $\mathrm{E}(\mathbf{h}_q(\mathbf{x}))$ von (2.150).

Damit könnte durch hinreichend hohe Approximationsordnung von $\mathbf{h}(\mathbf{x})$ eine beliebig genaue Approximation von $\mathrm{E}(\mathbf{h}(\mathbf{x}))$ ermittelt werden. Diese Vorgehensweise hat jedoch zwei wesentliche Nachteile:

1. Zur Approximation einer allgemeinen nichtlinearen Funktion kann eine sehr große Approximationsordnung notwendig sein. Dies führt zu einer wesentlichen Erhöhung des Rechenaufwands.

2. In der Berechnung des Erwartungswerts bei einer Approximation der Ordnung k sind die zentralen Momente bis zur Ordnung k des Zufallsvektors $\mathbf{x}$ notwendig.

Deswegen ist diese Vorgehensweise für eine praktische Implementierung nur bedingt sinnvoll.

Zur Charakterisierung einer (normalverteilten) Zufallfsvariable $\mathbf{x}$ ist auch die Kovarianzmatrix $\mathrm{E}\Big((\mathbf{x}-\mathbf{m}_x)(\mathbf{x}-\mathbf{m}_x)^{\mathrm{T}}\Big) = \mathbf{Q}_x$ notwendig. Daher ist es interessant zu untersuchen, wie sich die Kovarianzmatrix zufolge der nichtlinearen Transformation ändert. Für das betrachtete Beispiel gilt

$$\begin{aligned}\mathbf{Q}_y &= \mathrm{E}\left(\begin{bmatrix} x_1\cos(x_2) \\ x_1\sin(x_2) - \frac{\sin(\delta_{x2})}{\delta_{x2}} \end{bmatrix}\begin{bmatrix} x_1\cos(x_2) & x_1\sin(x_2) - \frac{\sin(\delta_{x2})}{\delta_{x2}} \end{bmatrix}\right) \\ &= \mathrm{E}\left(\begin{bmatrix} x_1^2\cos(x_2)^2 & x_1^2\cos(x_2)\sin(x_2) - x_1\cos(x_2)\frac{\sin(\delta_{x2})}{\delta_{x2}} \\ x_1^2\cos(x_2)\sin(x_2) - x_1\cos(x_2)\frac{\sin(\delta_{x2})}{\delta_{x2}} & \left(x_1\sin(x_2) - \frac{\sin(\delta_{x2})}{\delta_{x2}}\right)^2 \end{bmatrix}\right),\end{aligned} \tag{2.151}$$

was nach kurzer Rechnung auf

$$\mathbf{Q}_y = \begin{bmatrix} \frac{1}{2}(1+\sigma_{x1}^2)\left(1-\frac{\sin(2\delta_{x2})}{2\delta_{x2}}\right) & 0 \\ 0 & \frac{1}{2}(1+\sigma_{x1}^2)\left(1+\frac{\sin(2\delta_{x2})}{2\delta_{x2}}\right) - \left(\frac{\sin(\delta_{x2})}{\delta_{x2}}\right)^2 \end{bmatrix} \tag{2.152}$$

führt.

Aufgabe 2.18. Rechnen Sie die Lösung (2.152) nach.

Zur Schätzung der transformierten Kovarianzmatrix $\mathbf{Q}_y$ mit Hilfe der Taylorreihenapproximation 1. Ordnung verwendet man den Ausdruck

$$\mathbf{y}_l - \mathrm{E}(\mathbf{y}_l) = \mathrm{D}_{\tilde{\mathbf{x}}}\mathbf{h} = \tilde{x}_1 \left.\frac{\partial \mathbf{h}}{\partial x_1}\right|_{\mathbf{x}=\mathbf{m}_x} + \tilde{x}_2 \left.\frac{\partial \mathbf{h}}{\partial x_2}\right|_{\mathbf{x}=\mathbf{m}_x}. \tag{2.153}$$

Damit ergibt sich die Approximation $\mathbf{Q}_{yl}$ der Kovarianzmatrix zu

$$\begin{aligned} \mathbf{Q}_{yl} &= \mathrm{E}\left(\left[\tilde{x}_1 \left.\frac{\partial \mathbf{h}}{\partial x_1}\right|_{\mathbf{x}=\mathbf{m}_x} + \tilde{x}_2 \left.\frac{\partial \mathbf{h}}{\partial x_2}\right|_{\mathbf{x}=\mathbf{m}_x} \right] \left[\tilde{x}_1 \left.\frac{\partial \mathbf{h}}{\partial x_1}\right|_{\mathbf{x}=\mathbf{m}_x} + \tilde{x}_2 \left.\frac{\partial \mathbf{h}}{\partial x_2}\right|_{\mathbf{x}=\mathbf{m}_x} \right]^{\mathrm{T}} \right) \\ &= \left[\left.\frac{\partial \mathbf{h}}{\partial x_1}\right|_{\mathbf{x}=\mathbf{m}_x} \quad \left.\frac{\partial \mathbf{h}}{\partial x_2}\right|_{\mathbf{x}=\mathbf{m}_x} \right] \mathrm{E}\left(\begin{bmatrix} \tilde{x}_1^2 & \tilde{x}_1\tilde{x}_2 \\ \tilde{x}_1\tilde{x}_2 & \tilde{x}_2^2 \end{bmatrix} \right) \begin{bmatrix} \left.\frac{\partial \mathbf{h}}{\partial x_1}\right|_{\mathbf{x}=\mathbf{m}_x}^{\mathrm{T}} \\ \left.\frac{\partial \mathbf{h}}{\partial x_2}\right|_{\mathbf{x}=\mathbf{m}_x}^{\mathrm{T}} \end{bmatrix} = \mathbf{H}\mathbf{Q}_x\mathbf{H}^{\mathrm{T}} \end{aligned} \tag{2.154}$$

bzw. für das betrachtete Beispiel

$$\mathbf{Q}_{yl} = \begin{bmatrix} \sigma_{x2}^2 & 0 \\ 0 & \sigma_{x1}^2 \end{bmatrix}. \tag{2.155}$$

In Abbildung 2.9 ist ein Vergleich der Kovarianzmatrix $\mathbf{Q}_y$ und der Approximation $\mathbf{Q}_{yl}$ in Form der durch sie definierten Ellipsen für $C = 1$ dargestellt.

Satz 2.9 (Approximationsordnung)**.** *Der Erwartungswert $\mathbf{m}_y$ und die Kovarianzmatrix $\mathbf{Q}_y$ einer Zufallsvariable $\mathbf{y}$, welche sich durch ein nichtlineare Transformation der Form $\mathbf{y} = \mathbf{h}(\mathbf{x})$ errechnet, ergeben sich zu*

$$\mathbf{m}_y = \mathbf{h}(\mathbf{m}_x) + \mathrm{E}\left(\mathrm{D}_{\tilde{\mathbf{x}}}\mathbf{h} + \frac{1}{2!}\mathrm{D}_{\tilde{\mathbf{x}}}^2\mathbf{h} + \frac{1}{3!}\mathrm{D}_{\tilde{\mathbf{x}}}^3\mathbf{h} + \frac{1}{4!}\mathrm{D}_{\tilde{\mathbf{x}}}^4\mathbf{h} + \dots \right) \tag{2.156}$$

und

$$\begin{aligned} \mathbf{Q}_y = \mathbf{H}\mathbf{Q}_x\mathbf{H}^{\mathrm{T}} &+ \mathrm{E}\left(\frac{1}{3!}\mathrm{D}_{\tilde{\mathbf{x}}}\mathbf{h}\left(\mathrm{D}_{\tilde{\mathbf{x}}}^3\mathbf{h}\right)^{\mathrm{T}} + \frac{1}{4}\mathrm{D}_{\tilde{\mathbf{x}}}^2\mathbf{h}\left(\mathrm{D}_{\tilde{\mathbf{x}}}^2\mathbf{h}\right)^{\mathrm{T}} + \frac{1}{3!}\mathrm{D}_{\tilde{\mathbf{x}}}^3\mathbf{h}(\mathrm{D}_{\tilde{\mathbf{x}}}\mathbf{h})^{\mathrm{T}} \right) \\ &- \mathrm{E}\left(\mathrm{D}_{\tilde{\mathbf{x}}}^2\mathbf{h}\right)\mathrm{E}\left(\mathrm{D}_{\tilde{\mathbf{x}}}^2\mathbf{h}\right)^{\mathrm{T}} + \dots, \end{aligned} \tag{2.157}$$

mit $\mathbf{H} = \partial\mathbf{h}/\partial\mathbf{x}|_{\mathbf{x}=\mathbf{m}_x}$.

Um den Erwartungswert $\mathbf{m}_y$ mit einer Approximationsgenauigkeit der Ordnung m zu berechnen, sind die partiellen Ableitungen von $\mathbf{h}$ und die zentralen Momente von $\mathbf{x}$ bis zur m.ten Ordnung notwendig. Weiterhin kann der Term der m.ten Ordnung der

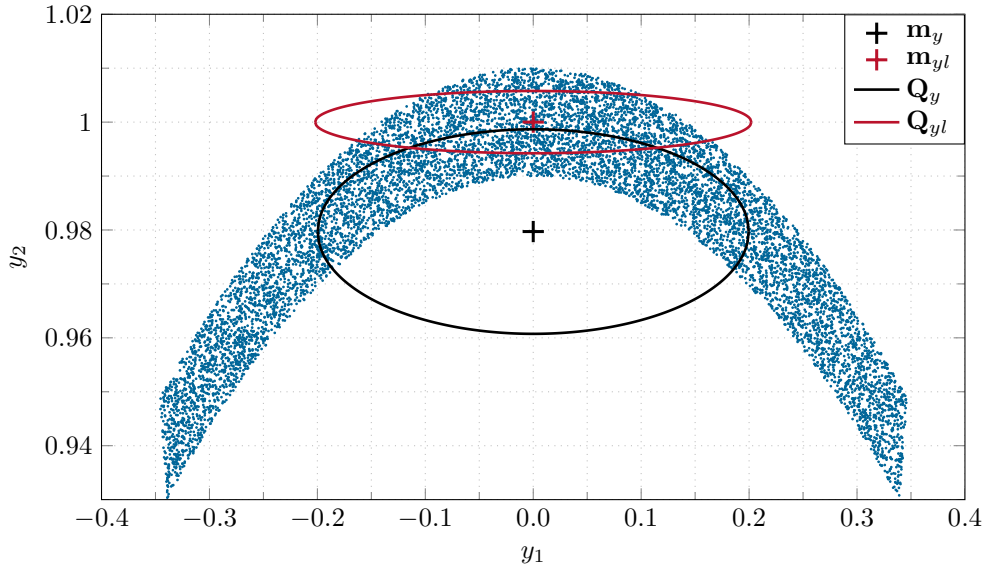


Abbildung 2.9: Verteilung der Zufallsvektoren $\mathbf{y} = \mathbf{h}(\mathbf{x})$ mit Erwartungswert $\mathrm{E}(\mathbf{y}) = \mathbf{m}_y$ bzw. Kovarianzmatrix $\mathbf{Q}_y$ und die Approximationen $\mathbf{m}_{yl}$ bzw. $\mathbf{Q}_{yl}$ auf Basis einer Taylorreihe 1. Ordnung.

Reihe der Kovarianzmatrix nur bestimmt werden, wenn die Ableitungen von $\mathbf{h}$ *und die zentralen Momente von* $\mathbf{x}$ *bis zur* $2m$*.ten Ordnung bekannt sind.*

Da das EKF maßgeblich auf dieser Linearisierung beruht, ist für Systeme mit ausgeprägter Nichtlinearität eine bessere Form der Approximation des Erwartungswertes und der Kovarianzmatrix notwendig.

2.5.2 Die Unscented-Transformation

Die Unscented-Transformation beruht darauf, dass es häufig einfacher ist die Verteilung einer Zufallsvariable zu approximieren als eine allgemeine nichtlineare Funktion oder Transformation. Bei der Unscented-Transformation wird eine Menge $\mathcal{S}$ von Sigmapunkten $\boldsymbol{\xi}_i$ so gewählt, dass deren Erwartungswert $\mathbf{m}_{xu}$ und die Kovarianzmatrix $\mathbf{Q}_{xu}$ jenen des ursprünglichen Zufallsvektors $\mathbf{x}$ entsprechen. Die Anwendung einer nichtlinearen Transformation $\mathbf{y} = \mathbf{h}(\mathbf{x})$ auf jeden dieser Sigmapunkte $\boldsymbol{\xi}_i$ ergibt die transformierten Sigmapunkte $\boldsymbol{\eta}_i = \mathbf{h}(\boldsymbol{\xi}_i)$. Die statistischen Eigenschaften der transformierten Punkte $\boldsymbol{\eta}_i$ sind dann eine Approximation der exakten statistischen Eigenschaften der nichtlinearen Transformation.

Die Menge der Sigmapunkte $\mathcal{S}$ ist durch die Vektoren $\boldsymbol{\xi}_i$ und die zugehörigen Gewichtungen W_i definiert, d. h. $\mathcal{S} = \{\boldsymbol{\xi}_i, W_i | i = 1, \ldots, p\}$. Die Wahl der Sigmapunkte ist nicht eindeutig, d. h. für einen Zufallsvektor $\mathbf{x}$ mit Erwartungswert $\mathbf{m}_x$ und Kovarianzmatrix $\mathbf{Q}_x$ gibt es mehrere mögliche Realisierungen von Sigmapunkten. Um eine erwartungstreue

Schätzung zu erhalten, müssen die Gewichtungen W_i die Zwangsbedingung

$$\sum_{i=1}^{p} W_i = 1 \tag{2.158}$$

erfüllen.

In der Literatur wird für einen Zufallsvektor $\mathbf{x} \in \mathbb{R}^n$ häufig eine Menge $\mathcal{S}$ von $2n$ Punkten, die auf dem Ellipsoid $(\mathbf{x} - \mathbf{m}_x)\mathbf{Q}_x^{-1}(\mathbf{x} - \mathbf{m}_x)^{\mathrm{T}} = n$ liegen, verwendet. Die Sigmapunkte für einen Zufallsvektor $\mathbf{x}$ sind demnach durch

$$\boldsymbol{\xi}_i = \begin{cases} \mathbf{m}_x + \left(\sqrt{n\mathbf{Q}_x}\right)_i^{\mathrm{T}} & \text{für } i = 1, \ldots, n \\ \mathbf{m}_x - \left(\sqrt{n\mathbf{Q}_x}\right)_{i-n}^{\mathrm{T}} & \text{für } i = n+1, \ldots, 2n \end{cases} \tag{2.159}$$

definiert, mit der Dimension n, dem Erwartungswert $\mathbf{m}_x$ und der Kovarianzmatrix $\mathbf{Q}_x$ des Zufallsvektors $\mathbf{x}$. Weiterhin beschreibt $\left(\sqrt{n\mathbf{Q}_x}\right)_i$ die i.te Zeile der Wurzel der Matrix $n\mathbf{Q}_x$. Die zugehörigen Gewichtungen ergeben sich zu

$$W_i = \frac{1}{2n} \; . \tag{2.160}$$

Hinweis: Die Wurzel $\mathbf{R}$ einer Matrix $\mathbf{Q}$ kann sehr effizient mit Hilfe der Cholesky-Zerlegung (MATLAB Befehl `chol`) ermittelt werden. Es gilt dann $\mathbf{R}^{\mathrm{T}}\mathbf{R} = \mathbf{Q}$.

Aufgabe 2.19. Zeigen Sie, dass der Mittelwert und die Kovarianzmatrix der durch (2.159) und (2.160) definierten Sigmapunkte dem Mittelwert $\mathbf{m}_x$ und der Kovarianzmatrix $\mathbf{Q}_x$ des Zufallsvektors $\mathbf{x}$ entsprechen.

Die Vorgehensweise zur Bestimmung des approximierten Erwartungswertes $\mathbf{m}_{yu}$ und der Kovarianzmatrix $\mathbf{Q}_{yu}$ mit Hilfe der Unscented-Transformation besteht aus folgenden Schritten:

1. Die nichtlineare Transformation $\mathbf{h}(\mathbf{x})$ wird auf jeden Sigmapunkt $\boldsymbol{\xi}_i$ angewandt und ergibt die transformierten Sigmapunkte $\boldsymbol{\eta}_i$

   $$\boldsymbol{\eta}_i = \mathbf{h}(\boldsymbol{\xi}_i) \; . \tag{2.161}$$

2. Der Erwartungswert $\mathbf{m}_{yu}$ ergibt sich aus der gewichteten Summe der transformierten Sigmapunkte

   $$\mathbf{m}_{yu} = \sum_{i=1}^{2n} W_i \boldsymbol{\eta}_i \; . \tag{2.162}$$

3. Die Kovarianzmatrix $\mathbf{Q}_{yu}$ der transformierten Sigmapunkte errechnet sich in der Form

   $$\mathbf{Q}_{yu} = \sum_{i=1}^{2n} W_i (\boldsymbol{\eta}_i - \mathbf{m}_{yu})(\boldsymbol{\eta}_i - \mathbf{m}_{yu})^{\mathrm{T}} \; . \tag{2.163}$$

Wie in der Literatur dokumentiert, würden bereits n Sigmapunkte genügen um den Erwartungswert und die Kovarianzmatrix korrekt zu approximieren. Durch die symmetrische Wahl der $2n$ Sigmapunkte nach (2.159) wird jedoch auch das dritte zentrale Moment exakt erfüllt.

Die Approximationsordnung der Unscented-Transformation ist 2, d. h. der Mittelwert und die Kovarianzmatrix werden bis zum zweiten Term korrekt wiedergegeben. Man beachte, dass zur Berechnung des Mittelwerts und der Kovarianz keine höheren Momente von $\mathbf{x}$ oder partielle Ableitungen der nichtlinearen Transformation notwendig sind. Letztere Eigenschaft erlaubt es damit, die Unscented-Transformation auch auf nicht stetig-differenzierbare (sogar nicht stetige) nichtlineare Transformationen anzuwenden.

Im Folgenden wird die Anwendung der Unscented-Transformation auf die nichtlineare Transformation $\mathbf{y} = \mathbf{h}(\mathbf{x})$ aus (2.138) in Abschnitt 2.5.1 betrachtet. Die $2n = 4$ Sigmapunkte $\boldsymbol{\xi}_i$ nach (2.159) sind in der Abbildung 2.10 zusammen mit dem exakten Mittelwert $\mathbf{m}_x$ und der durch die Ellipse für $C = \sqrt{2}$ charakterisierten Kovarianzmatrix $\mathbf{Q}_x$ dargestellt. Wie erwartet sind die Sigmapunkte $\boldsymbol{\xi}_i$ symmetrisch auf dieser Ellipse verteilt.

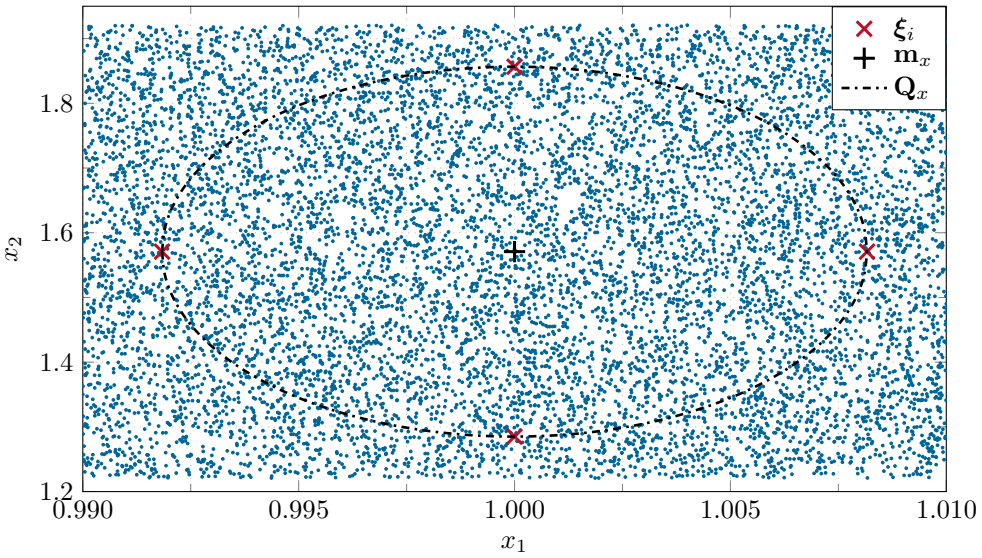


Abbildung 2.10: Gleichverteilte Zufallsvariable $\mathbf{x}$ nach (2.138) mit dem Mittelwert $\mathbf{m}_x$, der Kovarianzmatrix $\mathbf{Q}_x$ und den Sigmapunkten $\boldsymbol{\xi}_i$.

In Abbildung 2.11 sind die transformierten Sigmapunkte $\boldsymbol{\eta}_i = \mathbf{h}(\boldsymbol{\xi}_i)$ dargestellt. Weiterhin sind der nach (2.162) approximierte Mittelwert $\mathbf{m}_{yu}$ und die approximierte Kovarianzmatrix $\mathbf{Q}_{yu}$ nach (2.163) eingezeichnet. Im Vergleich zur mit Hilfe der Linearisierung berechneten Kovarianzmatrix $\mathbf{Q}_{yl}$ aus dem letzten Abschnitt 2.5.1 ergibt sich eine drastische Verbesserung der Approximation, vgl. Abbildung 2.9.

Bemerkung 2.3. Wie bereits angemerkt ist die Wahl der Sigmapunkte nicht eindeutig. Neben den in diesem Skriptum dargestellten Sigmapunkten auf Basis der Unscented-Transformation werden in der Literatur auch häufig sogenannte simplex Sigmapunkte oder spherische Sigmapunkte verwendete, siehe z. B. [2.3]. Diese zeichnen

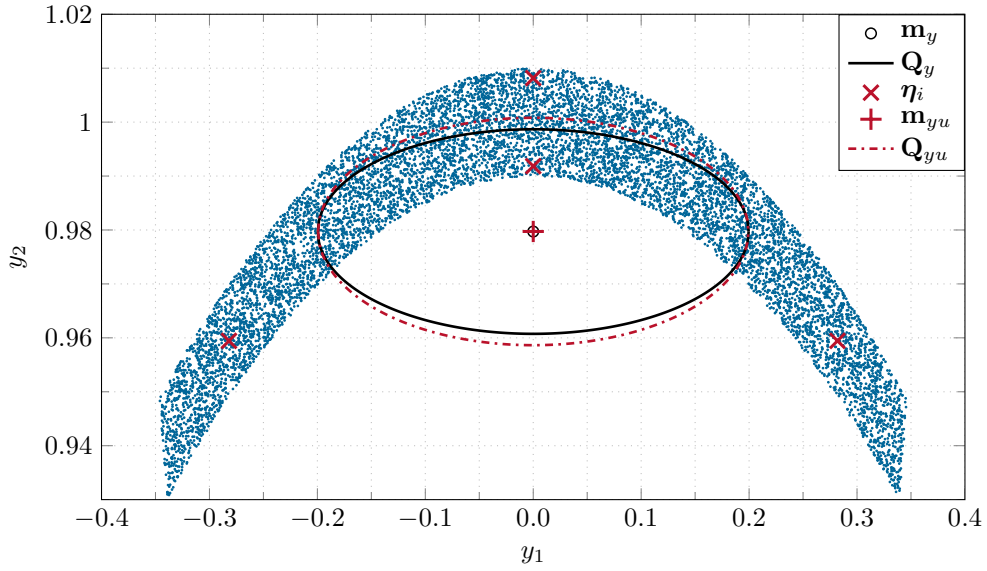


Abbildung 2.11: Transformierte Zufallsvariable $\mathbf{y} = \mathbf{h}(\mathbf{x})$ nach (2.138) mit dem exakten Mittelwert $\mathbf{m}_y$, der exakten Kovarianzmatrix $\mathbf{Q}_y$, den transformierten Sigmapunkten $\boldsymbol{\eta}_i = \mathbf{h}(\boldsymbol{\xi}_i)$, dem approximierten Mittelwert $\mathbf{m}_{yu}$ und der approximierten Kovarianzmatrix $\mathbf{Q}_{yu}$.

sich durch eine etwas einfachere Berechnung aus, weisen jedoch Nachteile bei der Approximationsgenauigkeit auf.

Wenn der Zufallsvektor $\mathbf{x}$ normalverteilt ist, dann erweist sich häufig die erweiterte Wahl der Sigmapunkte der Form

$$\boldsymbol{\xi}_0 = \mathbf{m}_x \tag{2.164a}$$

$$\boldsymbol{\xi}_i = \begin{cases} \mathbf{m}_x + \left(\sqrt{\frac{n}{1-\lambda}\mathbf{Q}_x}\right)_i^{\mathrm{T}} & \text{für } i = 1, \ldots, n \\ \mathbf{m}_x - \left(\sqrt{\frac{n}{1-\lambda}\mathbf{Q}_x}\right)_{i-n}^{\mathrm{T}} & \text{für } i = n+1, \ldots, 2n \end{cases} \tag{2.164b}$$

mit dem skalaren Parameter λ als sinnvoll. Die zugehörigen Gewichtungen sind in der Form

$$W_0 = \lambda \tag{2.165a}$$

$$W_i = \frac{1-\lambda}{2n} \,. \tag{2.165b}$$

gegeben. Es kann gezeigt werden, dass die Wahl $\lambda = 1 - n/3$ für normalverteilte Zufallsvektoren $\mathbf{x}$ optimal ist [2.4]. Die Begründung für die Wahl der erweiterten Sigmapunkte basiert auf einer Analyse des Einflusses von Momenten höherer Ordnung auf die Approximation der Kovarianzmatrix. Für eine detaillierte Analyse und die

Ermittlung des optimalen Werts von λ für normalverteilte Zufallsvektoren sei auf die Literatur, insbesondere [2.4], [2.3], [2.5], verwiesen.

Um den Einfluss der erweiterten Sigmapunkte (2.164) und der Gewichtungen (2.165) auf den mittels der Unscented-Transformation geschätzten Mittelwert und die zugehörige Kovarianzmatrix zu analysieren, betrachte man nochmals die nichtlineare Transformation nach (2.138). Es wird nun angenommen, dass der Zufallsvektor $\mathbf{x} = [x_1, x_2]^{\mathrm{T}}$ durch normalverteilte Zufallszahlen x_1 (Erwartungswert $m_{x1} = 1$, Varianz $\sigma_{x1} = 0.01$) und x_2 (Erwartungswert $m_{x2} = \pi/2$, Varianz $\sigma_{x2} = 0.35$) definiert ist. In Abbildung 2.12 ist die Verteilung von 10000 Zufallsvektoren $\mathbf{x}$, die in MATLAB mit dem Befehl `randn` erzeugt wurden, mit deren Mittelwert $\mathbf{m}_x$ und Kovarianzmatrix $\mathbf{Q}_x$ dargestellt. Weiterhin sind die 4 Sigmapunkte $\boldsymbol{\xi}_i$ nach (2.159) und die 5 erweiterten Sigmapunkte $\boldsymbol{\xi}_{ei}$ nach (2.164) eingezeichnet. Dabei wurde die für normalverteilte Zufallsvektoren optimale Wahl $\lambda = 1 - 2/3 = 1/3$ getroffen. Man erkennt, dass die erweiterten Sigmapunkte $\boldsymbol{\xi}_{ei}$ auf einer Ellipse mit vergrößertem Wert von C liegen. Äquivalent würde die Wahl $\lambda < 0$ dazu führen, dass die Sigmapunkte auf einer Ellipse mit verringertem Wert von C zu liegen kommen.

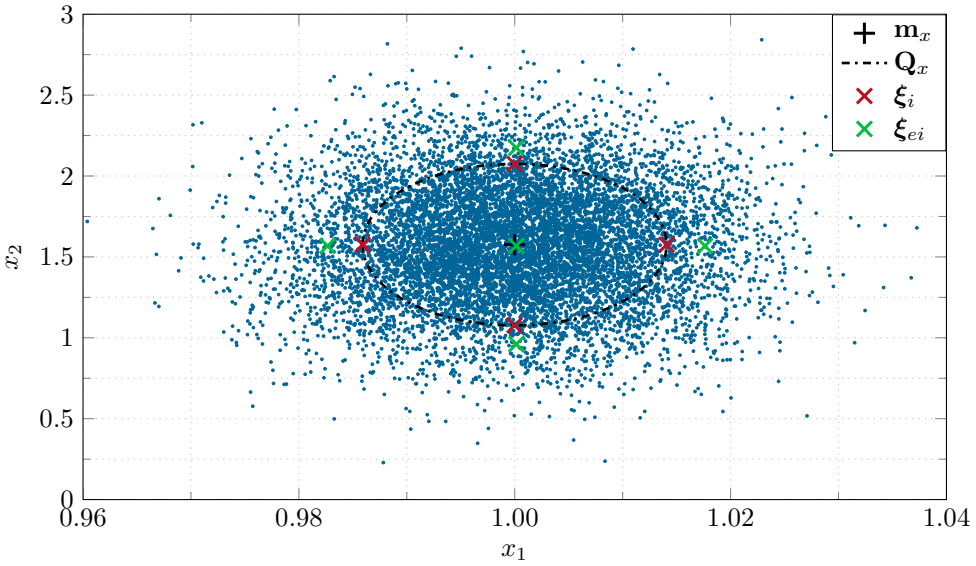


Abbildung 2.12: Normalverteilter Zufallsvektor $\mathbf{x}$ mit Mittelwert $\mathbf{m}_x$ und Kovarianzmatrix $\mathbf{Q}_x$, und die Sigmapunkte $\boldsymbol{\xi}_i$ nach (2.159) sowie die erweiterten Sigmapunkte $\boldsymbol{\xi}_{ei}$ nach (2.164) für $\lambda = 1/3$.

Wendet man wiederum die nichtlineare Transformation $\mathbf{y} = \mathbf{h}(\mathbf{x})$ aus (2.138) auf die Sigmapunkte an, so erhält man die in Abbildung 2.13 dargestellten transformierten Sigmapunkte $\boldsymbol{\eta}_i$ und die transformierten erweiterten Sigmapunkte $\boldsymbol{\eta}_{ei}$. Den Vorteil der erweiterten Sigmapunkte sieht man in der Approximation der Kovarianzmatrix $\mathbf{Q}_y$. Hier liefern die erweiterten Sigmapunkte $\boldsymbol{\xi}_{ei}$ aus (2.164) mit den zugehörigen Gewichtungen W_i nach (2.165) eine wesentliche Verbesserung der Approximationsgenauigkeit im Vergleich zu den ursprünglichen Sigmapunkten nach (2.159).

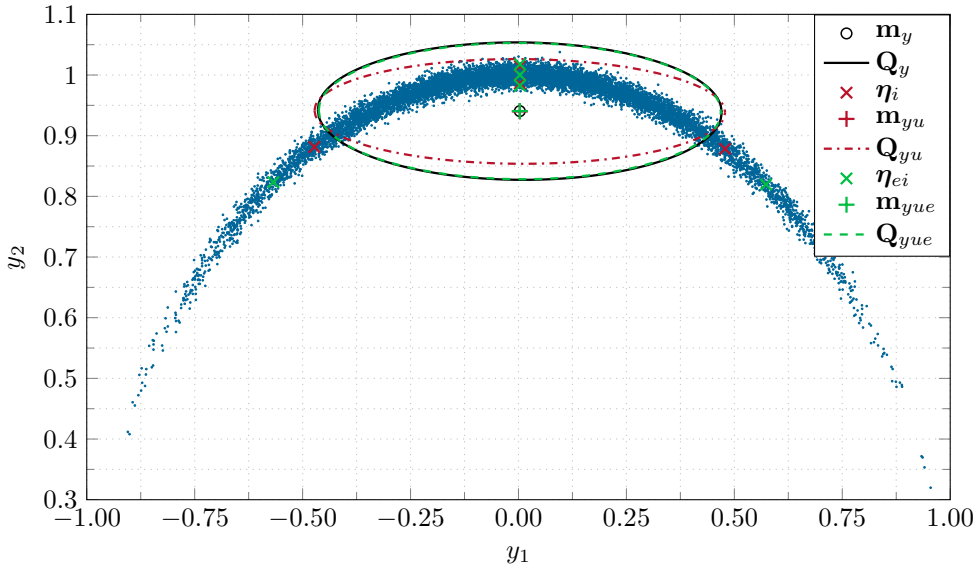


Abbildung 2.13: Transformierte normalverteilte Zufallsvariable $\mathbf{y} = \mathbf{h}(\mathbf{x})$ mit dem exakten Mittelwert $\mathbf{m}_y$ und der exakten Kovarianzmatrix $\mathbf{Q}_y$. Weiterhin ist die Approximation des Mittelwerts $\mathbf{m}_{yu}$ und der Kovarianzmatrix $\mathbf{Q}_{yu}$ auf Basis der Sigmapunkte aus (2.159) mit den Gewichtungen (2.160) und die Approximation des Mittelwerts $\mathbf{m}_{yue}$ und der Kovarianzmatrix $\mathbf{Q}_{yue}$ auf Basis der erweiterten Sigmapunkte aus (2.164) mit den Gewichtungen (2.165) dargestellt.

2.5.3 Zustandsschätzung dynamischer Systeme mit Hilfe der Unscented-Transformation

In diesem Abschnitt wird die Unscented-Transformation zur Zustandsschätzung von nichtlinearen dynamischen Systemen verwendet. In Analogie zum Extended Kalmanfilter aus Abschnitt 2.4 wird ein nichtlineares, zeitdiskretes dynamisches System der Form

$$\mathbf{x}_{k+1} = \mathbf{F}_k(\mathbf{x}_k, \mathbf{u}_k, \mathbf{w}_k) \tag{2.166a}$$
$$\mathbf{y}_k = \mathbf{h}_k(\mathbf{x}_k, \mathbf{u}_k, \mathbf{v}_k), \tag{2.166b}$$

mit dem Zustand $\mathbf{x}_k$, dem deterministischen Eingang $\mathbf{u}_k$, der Störung $\mathbf{w}_k$ und dem Messrauschen $\mathbf{v}_k$, betrachtet. Es wird vorausgesetzt, dass die statistischen Eigenschaften von $\mathbf{w}_k$ und $\mathbf{v}_k$ bekannt sind und durch

$$\mathrm{E}(\mathbf{v}_k) = \mathbf{0} \qquad \mathrm{E}\left(\mathbf{v}_k \mathbf{v}_j^{\mathrm{T}}\right) = \mathbf{R}\delta_{kj} \tag{2.167a}$$
$$\mathrm{E}(\mathbf{w}_k) = \mathbf{0} \qquad \mathrm{E}\left(\mathbf{w}_k \mathbf{w}_j^{\mathrm{T}}\right) = \mathbf{Q}\delta_{kj} \tag{2.167b}$$
$$\mathrm{E}\left(\mathbf{w}_k \mathbf{v}_j^{\mathrm{T}}\right) = \mathbf{0}, \tag{2.167c}$$

mit $\mathbf{Q} > 0$, $\mathbf{R} > 0$, gegeben sind.

Weiterhin ist eine Schätzung $\hat{\mathbf{x}}_0$ des Anfangswertes $\mathbf{x}_0$ in Form des Erwartungswertes $\mathbf{m}_0$ gegeben, d. h. $\hat{\mathbf{x}}_0 = \mathrm{E}(\mathbf{x}_0) = \mathbf{m}_0$, und die Kovarianzmatrix des Schätzfehlers $\mathbf{P}_0 = \mathrm{E}\left([\mathbf{x}_0 - \hat{\mathbf{x}}_0][\mathbf{x}_0 - \hat{\mathbf{x}}_0]^{\mathrm{T}}\right) \geq 0$ wird als bekannt vorausgesetzt. Wie schon beim Kalmanfilter und auch beim Extended Kalmanfilter wird angenommen, dass $\mathrm{E}\left(\mathbf{x}_0 \mathbf{w}_k^{\mathrm{T}}\right) = \mathbf{0}$ und $\mathrm{E}\left(\mathbf{x}_0 \mathbf{v}_k^{\mathrm{T}}\right) = \mathbf{0}$ gilt.

Für dieses System soll nun auf Basis der Unscented-Transformation ein Kalmanfilter, das sogenannte Unscented Kalmanfilter (in der Literatur auch als Sigmapunkt Kalmanfilter bezeichnet) dargestellt werden. Dazu beachte man, dass die Störung $\mathbf{w}_k$ und das Messrauschen $\mathbf{v}_k$ eine nichtlineare Transformation erfahren, weswegen auch deren statistische Eigenschaften anhand der Unscented-Transformation erfasst werden müssen. Dazu definiert man den erweiterten Zustand $\mathbf{x}_k^a \in \mathbb{R}^{n^a}$

$$\mathbf{x}_k^a = \begin{bmatrix} \mathbf{x}_k \\ \mathbf{w}_k \\ \mathbf{v}_k \end{bmatrix} \tag{2.168}$$

und berechnet für diesen erweiterten Zustand die Unscented-Transformation.

Damit kann folgende Iteration des Unscented Kalmanfilters formuliert werden, siehe [2.4], [2.5], [2.3] für eine detaillierte Herleitung:

1. **Initialisierung:**

$$\hat{\mathbf{x}}_0 = \mathrm{E}(\mathbf{x}_0) = \mathbf{m}_0 \tag{2.169a}$$

$$\mathbf{P}_0 = \mathrm{E}\Big([\mathbf{x}_0 - \hat{\mathbf{x}}_0][\mathbf{x}_0 - \hat{\mathbf{x}}_0]^{\mathrm{T}}\Big) \tag{2.169b}$$

$$\hat{\mathbf{x}}_0^{a+} = \mathrm{E}(\mathbf{x}_0^a) = \begin{bmatrix} \hat{\mathbf{x}}_0 \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} \tag{2.169c}$$

$$\mathbf{P}_0^{a+} = \mathrm{E}\Big(\Big[\mathbf{x}_0^a - \hat{\mathbf{x}}_0^{a+}\Big]\Big[\mathbf{x}_0^a - \hat{\mathbf{x}}_0^{a+}\Big]^{\mathrm{T}}\Big) = \begin{bmatrix} \mathbf{P}_0 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{R} \end{bmatrix} \tag{2.169d}$$

2. **Berechnung der Sigmapunkte:**

$$\boldsymbol{\xi}_{0,k-1}^{a+} = \hat{\mathbf{x}}_{k-1}^{a+} \tag{2.170a}$$

$$\boldsymbol{\xi}_{i,k-1}^{a+} = \begin{cases} \hat{\mathbf{x}}_{k-1}^{a+} + \sqrt{\frac{n^a}{1-\lambda}}\Big(\sqrt{\mathbf{P}_{k-1}^{a+}}\Big)_i^{\mathrm{T}} & \text{für } i = 1, \ldots, n^a \\ \hat{\mathbf{x}}_{k-1}^{a+} - \sqrt{\frac{n^a}{1-\lambda}}\Big(\sqrt{\mathbf{P}_{k-1}^{a+}}\Big)_{i-n^a}^{\mathrm{T}} & \text{für } i = n^a+1, \ldots, 2n^a \end{cases} \tag{2.170b}$$

mit

$$\boldsymbol{\xi}_{i,k-1}^{a+} = \begin{bmatrix} \boldsymbol{\xi}_{i,k-1}^{x+} \\ \boldsymbol{\xi}_{i,k-1}^{w+} \\ \boldsymbol{\xi}_{i,k-1}^{v+} \end{bmatrix}. \tag{2.171}$$

3. **Zustands- und Kovarianzmatrixextrapolation:**

Prädizierte Sigmapunkte:

$$\boldsymbol{\xi}_{i,k}^{x-} = \mathbf{F}_{k-1}\Big(\boldsymbol{\xi}_{i,k-1}^{x+}, \mathbf{u}_{k-1}, \boldsymbol{\xi}_{i,k-1}^{w+}\Big) \tag{2.172}$$

Prädizierter Mittelwert und prädizierte Fehlerkovarianzmatrix:

$$\hat{\mathbf{x}}_k^- = \sum_{i=0}^{2n^a} W_i \boldsymbol{\xi}_{i,k}^{x-} \tag{2.173a}$$

$$\mathbf{P}_k^- = \sum_{i=0}^{2n^a} W_i \Big(\boldsymbol{\xi}_{i,k}^{x-} - \hat{\mathbf{x}}_k^-\Big)\Big(\boldsymbol{\xi}_{i,k}^{x-} - \hat{\mathbf{x}}_k^-\Big)^{\mathrm{T}} \tag{2.173b}$$

Prädizierter Ausgang (Messung):

$$\boldsymbol{\eta}_{i,k}^{-} = \mathbf{h}_k\left(\boldsymbol{\xi}_{i,k}^{x-}, \mathbf{u}_k, \boldsymbol{\xi}_{i,k-1}^{v+}\right) \tag{2.174a}$$

$$\hat{\mathbf{y}}_k^{-} = \sum_{i=0}^{2n^a} W_i \boldsymbol{\eta}_{i,k}^{-} \tag{2.174b}$$

4. **Messupdate:**

   Kovarianzmatrix $\mathbf{P}_{\mathbf{y}_k\mathbf{y}_k}$ zwischen der prädizierten Messungen und Kovarianzmatrix $\mathbf{P}_{\mathbf{x}_k\mathbf{y}_k}$ zwischen prädizierter Messung und Zustand:

$$\mathbf{P}_{\mathbf{y}_k\mathbf{y}_k} = \sum_{i=0}^{2n^a} W_i \left(\boldsymbol{\eta}_{i,k}^{-} - \hat{\mathbf{y}}_k^{-}\right)\left(\boldsymbol{\eta}_{i,k}^{-} - \hat{\mathbf{y}}_k^{-}\right)^{\mathrm{T}} \tag{2.175a}$$

$$\mathbf{P}_{\mathbf{x}_k\mathbf{y}_k} = \sum_{i=0}^{2n^a} W_i \left(\boldsymbol{\xi}_{i,k}^{x-} - \hat{\mathbf{x}}_k^{-}\right)\left(\boldsymbol{\eta}_{i,k}^{-} - \hat{\mathbf{y}}_k^{-}\right)^{\mathrm{T}} \tag{2.175b}$$

   Messupdate des geschätzten Zustands und der Fehlerkovarianz:

$$\mathbf{K}_k = \mathbf{P}_{\mathbf{x}_k\mathbf{y}_k}(\mathbf{P}_{\mathbf{y}_k\mathbf{y}_k})^{-1} \tag{2.176a}$$

$$\hat{\mathbf{x}}_k^{+} = \hat{\mathbf{x}}_k^{-} + \mathbf{K}_k\left(\mathbf{y}_k - \hat{\mathbf{y}}_k^{-}\right) \tag{2.176b}$$

$$\mathbf{P}_k^{+} = \mathbf{P}_k^{-} - \mathbf{K}_k\mathbf{P}_{\mathbf{y}_k\mathbf{y}_k}\mathbf{K}_k^{\mathrm{T}} \tag{2.176c}$$

5. **Rücksetzen / Initialisieren für nächste Iteration:**

$$\hat{\mathbf{x}}_k^{a+} = \begin{bmatrix} \hat{\mathbf{x}}_k^{+} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} \tag{2.177a}$$

$$\mathbf{P}_k^{a+} = \begin{bmatrix} \mathbf{P}_k^{+} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{R} \end{bmatrix} \tag{2.177b}$$

Für die Berechnung des Unscented-Kalmanfilter sind nach (2.170) $1 + 2(n + \dim(\mathbf{w}) + \dim(\mathbf{v}))$ Sigmapunkte notwendig. Dies kann bei einer hohen Systemordnung zu einem sehr großen Rechenaufwand führen. In vielen regelungstechnischen Problemen kann jedoch angenommen werden, dass sowohl die Prozessstörung $\mathbf{w}_k$ als auch das Messrauschen $\mathbf{v}_k$ additiv auf das System einwirken. Damit kann das System (2.166) zu

$$\mathbf{x}_{k+1} = \mathbf{F}_k(\mathbf{x}_k, \mathbf{u}_k) + \mathbf{w}_k \tag{2.178a}$$

$$\mathbf{y}_k = \mathbf{h}_k(\mathbf{x}_k, \mathbf{u}_k) + \mathbf{v}_k, \tag{2.178b}$$

vereinfacht werden. Für dieses vereinfachte System kann nun, ausgehend von den Iterationsgleichungen für den allgemeinen Fall (2.169)-(2.177) die folgende vereinfachte Formulierung des Uscented-Kalmanfilters gefunden werden:

1. **Initialisierung:**

$$\hat{\mathbf{x}}_0^+ = \mathrm{E}(\mathbf{x}_0) = \mathbf{m}_0 \tag{2.179a}$$

$$\mathbf{P}_0^+ = \mathrm{E}\Big([\mathbf{x}_0 - \hat{\mathbf{x}}_0][\mathbf{x}_0 - \hat{\mathbf{x}}_0]^{\mathrm{T}}\Big) \tag{2.179b}$$

2. **Berechnung der Sigmapunkte:**

$$\boldsymbol{\xi}_{0,k-1}^+ = \hat{\mathbf{x}}_{k-1}^+ \tag{2.180a}$$

$$\boldsymbol{\xi}_{i,k-1}^+ = \begin{cases} \hat{\mathbf{x}}_{k-1}^+ + \sqrt{\frac{n}{1-\lambda}}\Big(\sqrt{\mathbf{P}_{k-1}^+}\Big)_i^{\mathrm{T}} & \text{für } i = 1,\ldots,n \\ \hat{\mathbf{x}}_{k-1}^+ - \sqrt{\frac{n}{1-\lambda}}\Big(\sqrt{\mathbf{P}_{k-1}^+}\Big)_{i-n}^{\mathrm{T}} & \text{für } i = n+1,\ldots,2n \end{cases} \tag{2.180b}$$

3. **Zustands- und Kovarianzmatrixextrapolation:**

Prädizierte Sigmapunkte:

$$\boldsymbol{\xi}_{i,k}^- = \mathbf{F}_{k-1}\Big(\boldsymbol{\xi}_{i,k-1}^+, \mathbf{u}_{k-1}\Big) \tag{2.181}$$

Prädizierter Mittelwert und prädizierte Fehlerkovarianzmatrix:

$$\hat{\mathbf{x}}_k^- = \sum_{i=0}^{2n} W_i \boldsymbol{\xi}_{i,k}^- \tag{2.182a}$$

$$\mathbf{P}_k^- = \sum_{i=0}^{2n} W_i \Big(\boldsymbol{\xi}_{i,k}^- - \hat{\mathbf{x}}_k^-\Big)\Big(\boldsymbol{\xi}_{i,k}^- - \hat{\mathbf{x}}_k^-\Big)^{\mathrm{T}} + \mathbf{Q} \tag{2.182b}$$

Korrektur der Sigmapunkte:

$$\boldsymbol{\xi}_{0,k}^- = \hat{\mathbf{x}}_k^- \tag{2.183a}$$

$$\boldsymbol{\xi}_{i,k}^- = \begin{cases} \hat{\mathbf{x}}_k^- + \sqrt{\frac{n}{1-\lambda}}\left(\sqrt{\mathbf{P}_k^-}\right)_i & \text{für } i = 1,\ldots,n \\ \hat{\mathbf{x}}_k^- - \sqrt{\frac{n}{1-\lambda}}\left(\sqrt{\mathbf{P}_k^-}\right)_{i-n} & \text{für } i = n+1,\ldots,2n \end{cases} \tag{2.183b}$$

Prädizierter Ausgang (Messung):

$$\boldsymbol{\eta}_{i,k}^- = \mathbf{h}_k\Big(\boldsymbol{\xi}_{i,k}^-, \mathbf{u}_k\Big) \tag{2.184a}$$

$$\hat{\mathbf{y}}_k^- = \sum_{i=0}^{2n} W_i \boldsymbol{\eta}_{i,k}^- \tag{2.184b}$$

4. **Messupdate:**

   Kovarianzmatrix $\mathbf{P}_{\mathbf{y}_k\mathbf{y}_k}$ der prädizierten Messungen und Kovarianzmatrix $\mathbf{P}_{\mathbf{x}_k\mathbf{y}_k}$ zwischen prädizierter Messung und Zustand:

$$\mathbf{P}_{\mathbf{y}_k\mathbf{y}_k} = \sum_{i=0}^{2n} W_i \left(\boldsymbol{\eta}_{i,k}^- - \hat{\mathbf{y}}_k^-\right)\left(\boldsymbol{\eta}_{i,k}^- - \hat{\mathbf{y}}_k^-\right)^{\mathrm{T}} + \mathbf{R} \tag{2.185a}$$

$$\mathbf{P}_{\mathbf{x}_k\mathbf{y}_k} = \sum_{i=0}^{2n} W_i \left(\boldsymbol{\xi}_{i,k}^- - \hat{\mathbf{x}}_k^-\right)\left(\boldsymbol{\eta}_{i,k}^- - \hat{\mathbf{y}}_k^-\right)^{\mathrm{T}} \tag{2.185b}$$

   Messupdate des geschätzten Zustands und der Fehlerkovarianz:

$$\mathbf{K}_k = \mathbf{P}_{\mathbf{x}_k\mathbf{y}_k}(\mathbf{P}_{\mathbf{y}_k\mathbf{y}_k})^{-1} \tag{2.186a}$$

$$\hat{\mathbf{x}}_k^+ = \hat{\mathbf{x}}_k^- + \mathbf{K}_k\left(\mathbf{y}_k - \hat{\mathbf{y}}_k^-\right) \tag{2.186b}$$

$$\mathbf{P}_k^+ = \mathbf{P}_k^- - \mathbf{K}_k\mathbf{P}_{\mathbf{y}_k\mathbf{y}_k}\mathbf{K}_k^{\mathrm{T}} \tag{2.186c}$$

Bemerkung 2.4. Die Korrektur der Sigmapunkte in (2.183) ist notwendig, da der Einfluss der Prozessstörung $\mathbf{w}_k$ erst anhand der Kovarianzmatrix $\mathbf{Q}$ in der prädizierten Fehlerkovarianzmatrix $\mathbf{P}_k^-$ nach (2.182b) berücksichtigt wird. Durch diese Korrektur muss pro Iteration zweimal die Cholesky-Zerlegung einer $n \times n$ Matrix berechnet werden. Diese numerisch aufwändige Operation wird in der praktischen Implementierung häufig umgangen, indem in (2.184), (2.185) die ursprünglichen Sigmapunkte $\boldsymbol{\xi}_{i,k}^-$ aus (2.181) verwendet werden. Dies resultiert in einem reduzierten numerischen Aufwand, es muss jedoch in der praktischen Anwendung überprüft werden, ob die damit resultierenden Fehler akzeptabel sind.

Eine weitere Vorgehensweise um die Sigmapunkte zu korrigieren, besteht in der Definition einer erweiterten Menge von Sigmapunkten in der Form

$$\boldsymbol{\xi}_{i,k}^{a-} = \boldsymbol{\xi}_{i,k}^- \quad \text{für } i = 0, \ldots, 2n \tag{2.187}$$

und

$$\boldsymbol{\xi}_{i+2n,k}^{a-} = \begin{cases} \boldsymbol{\xi}_{0,k}^- + \sqrt{\frac{2n}{1-\lambda}}(\sqrt{\mathbf{Q}})_i^{\mathrm{T}} & \text{für } i = 1, \ldots, n \\ \boldsymbol{\xi}_{0,k}^- - \sqrt{\frac{2n}{1-\lambda}}(\sqrt{\mathbf{Q}})_{i-n}^{\mathrm{T}} & \text{für } i = n+1, \ldots, 2n \end{cases}. \tag{2.188}$$

Diese erweiterten Sigmapunkte $\boldsymbol{\xi}_{i,k}^{a-}$ und die an die neue Dimension $4n+1$ von $\boldsymbol{\xi}_{i,k}^{a-}$ angepassten Gewichtungen W_i^a werden anschließend in der Berechnung von (2.184), (2.185) verwendet. Da $\mathbf{Q}$ eine konstante Matrix ist, entfällt die Berechnung der Cholesky-Zerlegung in (2.188), was den numerischen Aufwand verringert. Andererseits müssen in (2.184) und (2.185) nun $4n+1$ Sigmapunkte berücksichtigt werden, was im Vergleich zu (2.183) wiederum zu einer Erhöhung des numerischen Aufwands führt.

2.6 Literatur

[2.1] H. van Trees, *Detection, Estimation and Modulation Theory: Part 1*. New York, USA: John Wiley & Sons, 2001.

[2.2] A. S. Deshpande, „Bridging the Gap in Applied Kalman Filtering - Estimating Outputs when Measurements are Correlated with the Process Noise," *IEEE Control Systems Magazine*, S. 87–93, 2017.

[2.3] D. Simon, *Optimal State Estimation*. New Jersey, USA: John Wiley & Sons, 2006.

[2.4] S. Julier und J. Uhlmann, „Unscented Filtering and Nonlinear Estimation," *Proceedings of the IEEE*, Jg. 92, Nr. 3, S. 401–422, 2004.

[2.5] E. Wan und R. van der Merwe, „Kalman Filtering and Neural Networks," in S. Haykin, Hrsg. New York, USA: John Wiley & Sons, 2001, Kap. The Unscented Kalman Filter, S. 221–280.

[2.6] G. Franklin, J. Powell und M. Workman, *Digital Control of Dynamic Systems*, 3. Aufl. Menlo Park, USA: Addison–Weseley, 1998.

[2.7] L. Ljung, *System Identification*. New Jersey, USA: Prentice Hall, 1999.

[2.8] D. Luenberger, *Optimization by Vector Space Methods*. New York, USA: John Wiley & Sons, 1969.

[2.9] O. Nelles, *Nonlinear System Identification*. Berlin, Deutschland: Springer, 2001.

[2.10] R. Isermann, *Identifikation dynamischer Systeme 1 und 2*, 2. Aufl. Berlin, Deutschland: Springer, 1992.

[2.11] K. Åström und B. Wittenmark, *Computer Controlled Systems: Theory and Design*. New York, USA: Prentice Hall, 1997.

[2.12] A. Bryson und Y. Ho, *Applied Optimal Control*. Washington, USA: He, 1975.

[2.13] P. Dorato, C. Abdallah und V. Cerone, *Linear Quadratic Control: An Introduction*. Florida, USA: Krieger Publishing Company, 2000.

[2.14] A. Gelb, *Applied Optimal Estimation*. Cambridge, USA: MIT Pre, 74.

3 Optimaler Zustandsregler

Das Ziel dieses Kapitels ist die Entwicklung eines optimalen Zustandsreglers für lineare, zeitinvariante Systeme und die Kombination dieses Zustandsreglers mit dem optimalen Zustandsbeobachter des letzten Kapitels. Der Ausgangspunkt der Betrachtungen ist das lineare, zeitinvariante, zeitdiskrete System der Form

$$\mathbf{x}_{k+1} = \mathbf{\Phi}\mathbf{x}_k + \mathbf{\Gamma}\mathbf{u}_k \qquad \mathbf{x}(0) = \mathbf{x}_0 \tag{3.1a}$$

$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{D}\mathbf{u}_k \tag{3.1b}$$

mit dem n-dimensionalen Zustand $\mathbf{x} \in \mathbb{R}^n$, dem p-dimensionalen deterministischen Eingang $\mathbf{u} \in \mathbb{R}^p$, dem q-dimensionalen Ausgang $\mathbf{y} \in \mathbb{R}^q$ sowie den Matrizen $\mathbf{\Phi} \in \mathbb{R}^{n\times n}$, $\mathbf{\Gamma} \in \mathbb{R}^{n\times p}$, $\mathbf{C} \in \mathbb{R}^{q\times n}$ und $\mathbf{D} \in \mathbb{R}^{q\times p}$. Gesucht ist nun eine Steuerfolge $\mathbf{u}_0, \mathbf{u}_1, \ldots, \mathbf{u}_{N-1}$ die das Gütefunktional

$$\begin{aligned} J(\mathbf{x}_0) &= \sum_{k=0}^{N-1} \left(\mathbf{x}_k^{\mathrm{T}}\mathbf{Q}\mathbf{x}_k + \mathbf{u}_k^{\mathrm{T}}\mathbf{R}\mathbf{u}_k + 2\mathbf{u}_k^{\mathrm{T}}\mathbf{N}\mathbf{x}_k \right) + \mathbf{x}_N^{\mathrm{T}}\mathbf{S}\mathbf{x}_N \\ &= \sum_{k=0}^{N-1} \begin{bmatrix} \mathbf{x}_k^{\mathrm{T}} & \mathbf{u}_k^{\mathrm{T}} \end{bmatrix} \underbrace{\begin{bmatrix} \mathbf{Q} & \mathbf{N}^{\mathrm{T}} \\ \mathbf{N} & \mathbf{R} \end{bmatrix}}_{\mathbf{J}} \begin{bmatrix} \mathbf{x}_k \\ \mathbf{u}_k \end{bmatrix} + \mathbf{x}_N^{\mathrm{T}}\mathbf{S}\mathbf{x}_N \end{aligned} \tag{3.2}$$

für geeignete *Gewichtungsmatrizen* $\mathbf{Q} \in \mathbb{R}^{n\times n}$, $\mathbf{R} \in \mathbb{R}^{p\times p}$, $\mathbf{N} \in \mathbb{R}^{p\times n}$ und $\mathbf{S} \in \mathbb{R}^{n\times n}$ minimiert. Wegen des quadratischen Gütekriteriums (3.2) ist dieser Reglerentwurf in der Literatur auch unter dem Namen *LQR (Linear Quadratic Regulator)-Problem* zu finden. Zur Lösung dieser Aufgabe wird die Methode der *dynamischen Programmierung nach Bellman* herangezogen.

3.1 Dynamische Programmierung nach Bellman

Die Grundlage der dynamischen Programmierung bildet das *Optimalitätsprinzip*:

Satz 3.1 (Optimalitätsprinzip)**.** *Eine optimale Lösung hat die Eigenschaft, dass bei jedem Punkt dieser Lösung beginnend die verbleibende Lösung optimal im Sinne der zu lösenden Aufgabe ist mit dem gewählten Punkt als Anfangsbedingung.*

Die Abbildung 3.1 veranschaulicht Satz 3.1. Diese Idee wird nun im Sinne der *dynamischen Programmierung nach Bellman* so verwendet, dass die Optimierungsaufgabe (3.2) vom Endzeitpunkt N beginnend rückwärts gelöst wird. Dabei kann der Wert der optimalen Steuerung für den Zeitpunkt N, also $\mathbf{u}_{N-1}$, unabhängig vom erreichten Zustand $\mathbf{x}_{N-1}$ gelöst werden. Im nächsten Schritt wird ausgehend von der optimalen Lösung $\mathbf{u}_{N-1}$

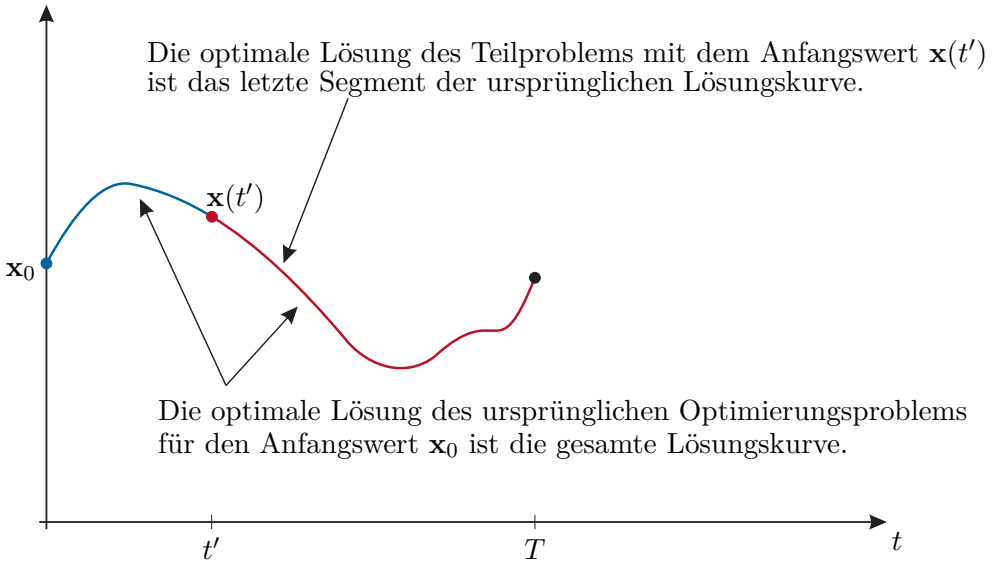


Abbildung 3.1: Zum Optimalitätsprinzip.

das optimale $\mathbf{u}_{N-2}$ berechnet. Wiederholt man diese Vorgehensweise bis $k = 0$, so ist die optimale Steuerstrategie gefunden.

Da bei der dynamischen Programmierung die Linearität der Strecke nicht notwendig ist, wird vorerst die Optimierungsaufgabe

$$\min_{(\mathbf{u}_0,\ldots,\mathbf{u}_{N-1})} J(\mathbf{x}_0) \quad \text{mit} \quad J(\mathbf{x}_0) = \sum_{k=0}^{N-1} j_k(\mathbf{x}_k, \mathbf{u}_k) + s(\mathbf{x}_N) \tag{3.3}$$

unter der Nebenbedingung

$$\mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_k, \mathbf{u}_k) \tag{3.4}$$

untersucht. Wie bereits erwähnt, wird die Optimierungsaufgabe (3.3) vom Endzeitpunkt $k = N$ beginnend rückwärts gelöst. Da $J(\mathbf{x}_N)$ unabhängig vom Stelleingang $\mathbf{u}$ ist, gilt trivialer Weise

$$J^*(\mathbf{x}_N) = s(\mathbf{x}_N), \tag{3.5}$$

wobei $J^*(\mathbf{x}_N)$ den optimalen Wert von $J(\mathbf{x}_N)$ beschreibt. Nach dem Optimalitätsprinzip gilt für die optimale Stellfolge $\mathbf{u}_0^*, \mathbf{u}_1^*, \ldots, \mathbf{u}_{N-1}^*$ mit

$$J^*(\mathbf{x}_0) = \sum_{k=0}^{N-1} j_k(\mathbf{x}_k, \mathbf{u}_k^*) + s(\mathbf{x}_N) \tag{3.6}$$

weiterhin die Beziehung

$$J^*(\mathbf{x}_0) = \sum_{k=0}^{l} j_k(\mathbf{x}_k, \mathbf{u}_k^*) + \underbrace{\sum_{k=l+1}^{N-1} j_k(\mathbf{x}_k, \mathbf{u}_k^*) + s(\mathbf{x}_N)}_{J^*(\mathbf{x}_{l+1})} \tag{3.7}$$

mit

$$J^*(\mathbf{x}_{l+1}) = \min_{(\mathbf{u}_{l+1},\ldots,\mathbf{u}_{N-1})} J(\mathbf{x}_{l+1}) \tag{3.8}$$

sowie

$$J(\mathbf{x}_{l+1}) = \sum_{k=l+1}^{N-1} j_k(\mathbf{x}_k, \mathbf{u}_k) + s(\mathbf{x}_N) \tag{3.9}$$

und unter der Nebenbedingung (3.4). Man beachte, dass (3.8) mit (3.9) für den Anfangswert $\mathbf{x}_{l+1}$ gelöst wird. Will man nun auf Basis von (3.8) einen Schritt zurückgehen und den optimalen Wert des Gütekriteriums $J^*(\mathbf{x}_l)$ bestimmen, dann folgt aus dem Optimalitätsprinzip die Ersatzaufgabe

$$J^*(\mathbf{x}_l) = \min_{\mathbf{u}_l} \left(j_l(\mathbf{x}_l, \mathbf{u}_l) + J^*\Big(\underbrace{\mathbf{x}_{l+1}}_{\mathbf{f}(\mathbf{x}_l,\mathbf{u}_l)} \Big) \right) . \tag{3.10}$$

Das Minimum bezüglich $\mathbf{u}_l$ in (3.10) kann dabei meistens aus der Beziehung

$$\frac{\partial}{\partial \mathbf{u}_l}\{j_l(\mathbf{x}_l, \mathbf{u}_l) + J^*(\mathbf{f}(\mathbf{x}_l, \mathbf{u}_l))\} = \frac{\partial}{\partial \mathbf{u}_l} j_l(\mathbf{x}_l, \mathbf{u}_l) + \frac{\partial}{\partial \mathbf{z}} J^*(\mathbf{z}) \frac{\partial}{\partial \mathbf{u}_l} \mathbf{f}(\mathbf{x}_l, \mathbf{u}_l) = \mathbf{0} \tag{3.11}$$

ermittelt werden.

Beispiel 3.1. Als nichtregelungstechnische Anwendung betrachte man ein einfaches Zuweisungsproblem. Gegeben sei eine Investitionssumme A, die auf N Projekte aufgeteilt werden soll. Es sei weiters angenommen, dass bei der Zuweisung von einer Summe u_k zum Projekt k das Projekt einen Gewinn $g_k(u_k)$ abwirft. Die zu lösende Optimierungsaufgabe lautet daher

$$\max_{(u_0,\ldots,u_{N-1})} J(x_0) \quad \text{mit} \quad J(x_0) = \sum_{k=0}^{N-1} g_k(u_k) \quad \text{unter der NB} \quad \sum_{k=0}^{N-1} u_k = A \; . \tag{3.12}$$

Die Aufgabe kann nun in eine äquivalente Regelungsaufgabe der Form

$$\max_{(u_0,\ldots,u_{N-1})} J(x_0) \quad \text{mit} \quad J(x_0) = \sum_{k=0}^{N-1} g_k(u_k) \tag{3.13}$$

unter der Nebenbedingung

$$x_{k+1} = x_k - u_k \quad \text{mit} \quad x_0 = A \quad \text{und} \quad x_N = 0 \tag{3.14}$$

umformuliert werden. Wählt man beispielsweise für $g_k(u_k) = \sqrt{u_k}$, dann erhält man mit Hilfe der dynamische Programmierung

$$\begin{aligned}
J^*(x_N) &= 0 \\
J^*(x_{N-1}) &= \max_{u_{N-1}} \left\{ \sqrt{u_{N-1}} \right\} \overset{\text{unter der NB } x_{N-1} - u_{N-1} = 0}{=} \sqrt{x_{N-1}} && u^*_{N-1} = x_{N-1} \\
J^*(x_{N-2}) &= \max_{u_{N-2}} \left\{ \sqrt{u_{N-2}} + \sqrt{x_{N-2} - u_{N-2}} \right\} = \sqrt{2x_{N-2}} && u^*_{N-2} = x_{N-2}/2 \\
J^*(x_{N-3}) &= \max_{u_{N-3}} \left\{ \sqrt{u_{N-3}} + \sqrt{2(x_{N-3} - u_{N-3})} \right\} = \sqrt{3x_{N-3}} && u^*_{N-3} = x_{N-3}/3 \\
&\vdots && \vdots \\
J^*(x_0) &= \sqrt{N x_0} && u^*_0 = x_0/N \ .
\end{aligned} \tag{3.15}$$

Aufgabe 3.1. Interpretieren Sie das Ergebnis (3.15).

3.2 Das LQR-Problem

Wendet man das Prinzip der dynamischen Programmierung auf die Aufgabe (3.2) mit der Nebenbedingung (3.1) an, so erhält man für $k = N$

$$J^*(\mathbf{x}_N) = \mathbf{x}_N^{\mathrm{T}} \mathbf{S} \mathbf{x}_N \tag{3.16}$$

und für $k = N - 1$

$$J^*(\mathbf{x}_{N-1}) = \min_{\mathbf{u}_{N-1}} \left\{ \left(\mathbf{x}_{N-1}^{\mathrm{T}} \mathbf{Q} \mathbf{x}_{N-1} + \mathbf{u}_{N-1}^{\mathrm{T}} \mathbf{R} \mathbf{u}_{N-1} + 2\mathbf{u}_{N-1}^{\mathrm{T}} \mathbf{N} \mathbf{x}_{N-1} \right) + J^* \Big(\underbrace{\mathbf{x}_N}_{\mathbf{\Phi x}_{N-1} + \mathbf{\Gamma u}_{N-1}} \Big) \right\} \tag{3.17}$$

bzw.

$$\begin{aligned}
J^*(\mathbf{x}_{N-1}) = \min_{\mathbf{u}_{N-1}} \Big\{ &\left(\mathbf{x}_{N-1}^{\mathrm{T}} \mathbf{Q} \mathbf{x}_{N-1} + \mathbf{u}_{N-1}^{\mathrm{T}} \mathbf{R} \mathbf{u}_{N-1} + 2\mathbf{u}_{N-1}^{\mathrm{T}} \mathbf{N} \mathbf{x}_{N-1} \right) + \\
&(\mathbf{\Phi x}_{N-1} + \mathbf{\Gamma u}_{N-1})^{\mathrm{T}} \mathbf{S} (\mathbf{\Phi x}_{N-1} + \mathbf{\Gamma u}_{N-1}) \Big\} \ .
\end{aligned} \tag{3.18}$$

Minimiert man (3.18) bezüglich $\mathbf{u}_{N-1}$ ergibt sich die optimale Lösung $\mathbf{u}^*_{N-1}$ von $\mathbf{u}_{N-1}$ zu

$$\mathbf{u}^*_{N-1} = -\left(\mathbf{R} + \mathbf{\Gamma}^{\mathrm{T}} \mathbf{S} \mathbf{\Gamma} \right)^{-1} \left(\mathbf{N} + \mathbf{\Gamma}^{\mathrm{T}} \mathbf{S} \mathbf{\Phi} \right) \mathbf{x}_{N-1} \ . \tag{3.19}$$

Durch Einsetzen von (3.19) in (3.18) folgt

$$\begin{aligned} &J^*(\mathbf{x}_{N-1}) \\ &= \mathbf{x}_{N-1}^{\mathrm{T}}\Big(\mathbf{Q}+\mathbf{\Phi}^{\mathrm{T}}\mathbf{S}\mathbf{\Phi}\Big)\mathbf{x}_{N-1} + (\mathbf{u}_{N-1}^*)^{\mathrm{T}}\Big(\mathbf{R}+\mathbf{\Gamma}^{\mathrm{T}}\mathbf{S}\mathbf{\Gamma}\Big)\mathbf{u}_{N-1}^* + 2(\mathbf{u}_{N-1}^*)^{\mathrm{T}}\Big(\mathbf{N}+\mathbf{\Gamma}^{\mathrm{T}}\mathbf{S}\mathbf{\Phi}\Big)\mathbf{x}_{N-1} \\ &= \mathbf{x}_{N-1}^{\mathrm{T}}\Big\{\Big(\mathbf{Q}+\mathbf{\Phi}^{\mathrm{T}}\mathbf{S}\mathbf{\Phi}\Big) - \Big(\mathbf{N}+\mathbf{\Gamma}^{\mathrm{T}}\mathbf{S}\mathbf{\Phi}\Big)^{\mathrm{T}}\Big(\mathbf{R}+\mathbf{\Gamma}^{\mathrm{T}}\mathbf{S}\mathbf{\Gamma}\Big)^{-1}\Big(\mathbf{N}+\mathbf{\Gamma}^{\mathrm{T}}\mathbf{S}\mathbf{\Phi}\Big)\Big\}\mathbf{x}_{N-1} \end{aligned} \tag{3.20}$$

Damit lässt sich unmittelbar nachfolgender Satz angeben:

Satz 3.2 (Linear Quadratic Regulator)**.** *Die eindeutige Lösung der Optimierungsaufgabe (3.2) für das lineare, zeitinvariante, zeitdiskrete System (3.1) mit der symmetrischen positiv semi-definiten Matrix* $\mathbf{S} = \mathbf{P}_N$*, der symmetrischen positiv semi-definiten Matrix*

$$\mathbf{J} = \begin{bmatrix} \mathbf{Q} & \mathbf{N}^{\mathrm{T}} \\ \mathbf{N} & \mathbf{R} \end{bmatrix} \tag{3.21}$$

und der positiv definiten Matrix $\Big(\mathbf{R}+\mathbf{\Gamma}^{\mathrm{T}}\mathbf{S}\mathbf{\Gamma}\Big)$ *ist durch das Regelgesetz*

$$\mathbf{u}_k^* = \mathbf{K}_k\mathbf{x}_k \tag{3.22}$$

mit

$$\mathbf{K}_k = -\Big(\mathbf{R}+\mathbf{\Gamma}^{\mathrm{T}}\mathbf{P}_{k+1}\mathbf{\Gamma}\Big)^{-1}\Big(\mathbf{N}+\mathbf{\Gamma}^{\mathrm{T}}\mathbf{P}_{k+1}\mathbf{\Phi}\Big) \tag{3.23}$$

und

$$\mathbf{P}_k = \Big(\mathbf{Q}+\mathbf{\Phi}^{\mathrm{T}}\mathbf{P}_{k+1}\mathbf{\Phi}\Big) - \Big(\mathbf{N}+\mathbf{\Gamma}^{\mathrm{T}}\mathbf{P}_{k+1}\mathbf{\Phi}\Big)^{\mathrm{T}}\Big(\mathbf{R}+\mathbf{\Gamma}^{\mathrm{T}}\mathbf{P}_{k+1}\mathbf{\Gamma}\Big)^{-1}\Big(\mathbf{N}+\mathbf{\Gamma}^{\mathrm{T}}\mathbf{P}_{k+1}\mathbf{\Phi}\Big) \tag{3.24}$$

gegeben. Der minimale Wert des Gütefunktionals (3.2) errechnet sich zu

$$\min_{(\mathbf{u}_0,\ldots,\mathbf{u}_{N-1})} J(\mathbf{x}_0) = J^*(\mathbf{x}_0) = \mathbf{x}_0^{\mathrm{T}}\mathbf{P}_0\mathbf{x}_0 \tag{3.25}$$

und es gilt $\mathbf{P}_k \geq 0$ *für alle* $k = 0, 1, \ldots, N$*.*

Beweis von Satz 3.2. Das Regelgesetz (3.22), (3.23) sowie die Iterationsvorschrift (3.24) und auch die Beziehung (3.25) erhält man unmittelbar durch wiederholte Anwendung der Iterationsvorschrift der dynamischen Programmierung aus den Gleichungen (3.19) und (3.20). Es bleibt damit zu zeigen, dass $\mathbf{P}_k$ für alle $k = 0, 1, \ldots, N$ positiv semi-definit ist. Dazu setze man in (3.20) für $\mathbf{u}_{N-1}^* = \mathbf{K}_{N-1}\mathbf{x}_{N-1}$ ein, womit folgt

$$\begin{aligned} J^*(\mathbf{x}_{N-1}) = \mathbf{x}_{N-1}^{\mathrm{T}}\Big(&\Big(\mathbf{Q}+\mathbf{\Phi}^{\mathrm{T}}\mathbf{P}_N\mathbf{\Phi}\Big) + \mathbf{K}_{N-1}^{\mathrm{T}}\Big(\mathbf{R}+\mathbf{\Gamma}^{\mathrm{T}}\mathbf{P}_N\mathbf{\Gamma}\Big)\mathbf{K}_{N-1} \\ &+ 2\mathbf{K}_{N-1}^{\mathrm{T}}\Big(\mathbf{N}+\mathbf{\Gamma}^{\mathrm{T}}\mathbf{P}_N\mathbf{\Phi}\Big)\Big)\mathbf{x}_{N-1} = \mathbf{x}_{N-1}^{\mathrm{T}}\mathbf{P}_{N-1}\mathbf{x}_{N-1} \end{aligned} \tag{3.26}$$

und damit für $\mathbf{P}_k$ von (3.24)

$$\mathbf{P}_k = (\mathbf{\Phi} + \mathbf{\Gamma}\mathbf{K}_k)^{\mathrm{T}}\mathbf{P}_{k+1}(\mathbf{\Phi} + \mathbf{\Gamma}\mathbf{K}_k) + \begin{bmatrix}\mathbf{E} & \mathbf{K}_k^{\mathrm{T}}\end{bmatrix} \underbrace{\begin{bmatrix}\mathbf{Q} & \mathbf{N}^{\mathrm{T}} \\ \mathbf{N} & \mathbf{R}\end{bmatrix}}_{\mathbf{J}} \begin{bmatrix}\mathbf{E} \\ \mathbf{K}_k\end{bmatrix}. \tag{3.27}$$

Da nun die Matrizen $\mathbf{P}_N = \mathbf{S}$ und $\mathbf{J}$ positiv semi-definit sind, ist auch direkt die positive Semi-definitheit von $\mathbf{P}_k$ für alle $k = 0, 1, \ldots, N$ gezeigt. □

Wie bereits beim Kalman-Filter als Beobachter (siehe (2.91)) handelt es sich auch bei Gleichung (3.24) um eine *diskrete Riccati-Gleichung*, weshalb der *zeitvariante Zustandsregler* (3.22), (3.23) auch als *Riccati-Regler* bezeichnet wird. Man beachte jedoch, dass die diskrete Riccati-Gleichung (3.24) im Gegensatz zum Kalman-Filter *rückwärts läuft*! Für eine Echtzeitimplementierung des Reglers (3.22), (3.23) muss daher die Endzeit N bekannt sein und die Matrizen $\mathbf{P}_k$ bzw. $\mathbf{K}_k$ müssen vorwegberechnet werden.

Wenn die Endzeit $N \to \infty$ geht, kann man wie bereits beim Kalman-Filter (vergleiche dazu (2.93), (2.94)) eine stationäre Lösung $\mathbf{P}_s$ und $\mathbf{K}_s$ aus (3.22)–(3.24) berechnen. Man könnte nun die stationäre Lösung $\mathbf{P}_s$ der diskreten Riccati-Gleichung (3.24) so bestimmen, dass man vom Anfangswert $\mathbf{P}_\infty = \alpha\mathbf{E}$ für $\alpha \gg 1$ beginnend so lange iteriert, bis sich $\mathbf{P}_k$ nur noch hinreichend wenig im Sinne einer Norm ändert. Eine weitere Lösungsmöglichkeit besteht darin, die zugehörige *diskrete algebraische Riccati-Gleichung* für $\mathbf{P}_{k+1} = \mathbf{P}_k = \mathbf{P}_s$ in (3.24)

$$\mathbf{P}_s = \left(\mathbf{Q} + \mathbf{\Phi}^{\mathrm{T}}\mathbf{P}_s\mathbf{\Phi}\right) - \left(\mathbf{N} + \mathbf{\Gamma}^{\mathrm{T}}\mathbf{P}_s\mathbf{\Phi}\right)^{\mathrm{T}}\left(\mathbf{R} + \mathbf{\Gamma}^{\mathrm{T}}\mathbf{P}_s\mathbf{\Gamma}\right)^{-1}\left(\mathbf{N} + \mathbf{\Gamma}^{\mathrm{T}}\mathbf{P}_s\mathbf{\Phi}\right) \tag{3.28}$$

zu lösen. Damit hat aber der *stationäre Riccati-Regler*

$$\mathbf{u}_k^* = \mathbf{K}_s\mathbf{x}_k \tag{3.29a}$$

$$\mathbf{K}_s = -\left(\mathbf{R} + \mathbf{\Gamma}^{\mathrm{T}}\mathbf{P}_s\mathbf{\Gamma}\right)^{-1}\left(\mathbf{N} + \mathbf{\Gamma}^{\mathrm{T}}\mathbf{P}_s\mathbf{\Phi}\right) \tag{3.29b}$$

die Struktur eines klassischen Zustandsreglers gemäß Abschnitt 8 des Skripts Automatisierung.

Die diskrete algebraische Riccati-Gleichung (3.28) hat eine eindeutige symmetrische positiv semi-definite Lösung $\mathbf{P}_s$ mit der Eigenschaft, dass sämtliche Eigenwerte von $(\mathbf{\Phi} + \mathbf{\Gamma}\mathbf{K}_s)$ im offenen Inneren des Einheitskreises liegen, wenn nachfolgende Bedingungen erfüllt sind:

(1) Das Paar $(\mathbf{\Phi}, \mathbf{\Gamma})$ ist *stabilisierbar*, d. h. sämtliche Eigenwerte außerhalb des Einheitskreises sind erreichbar, und

(2) das Paar $(\mathbf{C}_\mathbf{J}, \mathbf{\Phi})$ mit

$$0 \leq \mathbf{J} = \begin{bmatrix}\mathbf{Q} & \mathbf{N}^{\mathrm{T}} \\ \mathbf{N} & \mathbf{R}\end{bmatrix} = \begin{bmatrix}\mathbf{C}_\mathbf{J}^{\mathrm{T}} \\ \mathbf{D}_\mathbf{J}^{\mathrm{T}}\end{bmatrix}\begin{bmatrix}\mathbf{C}_\mathbf{J} & \mathbf{D}_\mathbf{J}\end{bmatrix} \tag{3.30}$$

ist *detektierbar*, d. h. sämtliche Eigenwerte außerhalb des Einheitskreises sind über den Ausgang $\mathbf{C}_\mathbf{J}$ beobachtbar.

Will man nun erreichen, dass sämtliche Pole des geschlossenen Kreises mit der Dynamikmatrix $(\mathbf{\Phi} + \mathbf{\Gamma}\mathbf{K}_s)$ nicht nur im Inneren des Einheitskreises liegen, sondern auf Grund von Robustheitsüberlegungen im Inneren eines Kreises mit dem Radius $r < 1$, dann muss der Reglerentwurf für das Ersatzsystem

$$\mathbf{x}_{k+1} = \tilde{\mathbf{\Phi}}\mathbf{x}_k + \tilde{\mathbf{\Gamma}}\mathbf{u}_k \tag{3.31}$$

mit

$$\tilde{\mathbf{\Phi}} = \frac{1}{r}\mathbf{\Phi} \quad \text{und} \quad \tilde{\mathbf{\Gamma}} = \frac{1}{r}\mathbf{\Gamma} \tag{3.32}$$

durchgeführt werden. Da dann die Eigenwerte der Matrix $\left(\tilde{\mathbf{\Phi}} + \tilde{\mathbf{\Gamma}}\mathbf{K}_s\right)$ im Inneren des Einheitskreises liegen, gilt wegen (3.31), dass die Eigenwerte von $(\mathbf{\Phi} + \mathbf{\Gamma}\mathbf{K}_s)$ im Inneren eines Kreises mit dem Radius r zu liegen kommen.

Aufgabe 3.2. Zeigen Sie, dass die Lösung des Optimierungsproblems

$$J(\mathbf{x}_0) = \sum_{k=0}^{\infty} \frac{1}{r^{2k}} \left(\mathbf{x}_k^{\mathrm{T}}\mathbf{Q}\mathbf{x}_k + \mathbf{u}_k^{\mathrm{T}}\mathbf{R}\mathbf{u}_k + 2\mathbf{u}_k^{\mathrm{T}}\mathbf{N}\mathbf{x}_k\right) \tag{3.33}$$

mit $0 < r < 1$ für das System (3.1) durch

$$\mathbf{u}_k^* = \mathbf{K}_s\mathbf{x}_k \tag{3.34}$$

mit

$$\mathbf{K}_s = -\left(r^2\mathbf{R} + \mathbf{\Gamma}^{\mathrm{T}}\mathbf{P}_s\mathbf{\Gamma}\right)^{-1}\left(r^2\mathbf{N} + \mathbf{\Gamma}^{\mathrm{T}}\mathbf{P}_s\mathbf{\Phi}\right) \tag{3.35}$$

und

$$r^2\mathbf{P}_s = \left(r^2\mathbf{Q} + \mathbf{\Phi}^{\mathrm{T}}\mathbf{P}_s\mathbf{\Phi}\right) - \left(r^2\mathbf{N} + \mathbf{\Gamma}^{\mathrm{T}}\mathbf{P}_s\mathbf{\Phi}\right)^{\mathrm{T}}\left(r^2\mathbf{R} + \mathbf{\Gamma}^{\mathrm{T}}\mathbf{P}_s\mathbf{\Gamma}\right)^{-1}\left(r^2\mathbf{N} + \mathbf{\Gamma}^{\mathrm{T}}\mathbf{P}_s\mathbf{\Phi}\right) \tag{3.36}$$

gegeben ist und die Eigenwerte der Dynamikmatrix des geschlossenen Kreises $(\mathbf{\Phi} + \mathbf{\Gamma}\mathbf{K}_s)$ im Inneren eines Kreises mit dem Radius $0 < r < 1$ liegen.

Wenn in dem Gütefunktional (3.2) nur der p-dimensionale Eingang $\mathbf{u}$ und der q-dimensionale Ausgang $\mathbf{y}$ gewichtet werden sollen, also

$$J(\mathbf{x}_0) = \sum_{k=0}^{N-1} \left(\mathbf{y}_k^{\mathrm{T}}\mathbf{Q}_{\mathbf{y}}\mathbf{y}_k + \mathbf{u}_k^{\mathrm{T}}\mathbf{R}\mathbf{u}_k + 2\mathbf{u}_k^{\mathrm{T}}\mathbf{N}_{\mathbf{y}}\mathbf{y}_k\right) + \mathbf{x}_N^{\mathrm{T}}\mathbf{S}\mathbf{x}_N \tag{3.37a}$$

$$= \sum_{k=0}^{N-1} \begin{bmatrix}\mathbf{y}_k^{\mathrm{T}} & \mathbf{u}_k^{\mathrm{T}}\end{bmatrix} \underbrace{\begin{bmatrix}\mathbf{Q}_{\mathbf{y}} & \mathbf{N}_{\mathbf{y}}^{\mathrm{T}} \\ \mathbf{N}_{\mathbf{y}} & \mathbf{R}\end{bmatrix}}_{\mathbf{J}} \begin{bmatrix}\mathbf{y}_k \\ \mathbf{u}_k\end{bmatrix} + \mathbf{x}_N^{\mathrm{T}}\mathbf{S}\mathbf{x}_N, \tag{3.37b}$$

dann kann (3.37) über die Beziehung $\mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{D}\mathbf{u}_k$ sowie

$$\begin{aligned}\sum_{k=0}^{N-1}\Big[\mathbf{x}_k^{\mathrm{T}}\underbrace{\mathbf{C}^{\mathrm{T}}\mathbf{Q}_{\mathbf{y}}\mathbf{C}}_{\tilde{\mathbf{Q}}}\mathbf{x}_k + \mathbf{u}_k^{\mathrm{T}}\underbrace{\left(\mathbf{R}+\mathbf{D}^{\mathrm{T}}\mathbf{Q}_{\mathbf{y}}\mathbf{D}+\mathbf{N}_{\mathbf{y}}\mathbf{D}+\mathbf{D}^{\mathrm{T}}\mathbf{N}_{\mathbf{y}}^{\mathrm{T}}\right)}_{\tilde{\mathbf{R}}}\mathbf{u}_k \\ + 2\mathbf{u}_k^{\mathrm{T}}\underbrace{\left(\mathbf{N}_{\mathbf{y}}+\mathbf{D}^{\mathrm{T}}\mathbf{Q}_{\mathbf{y}}\right)\mathbf{C}}_{\tilde{\mathbf{N}}}\mathbf{x}_k\Big] \\ = \sum_{k=0}^{N-1}\begin{bmatrix}\mathbf{x}_k^{\mathrm{T}} & \mathbf{u}_k^{\mathrm{T}}\end{bmatrix}\underbrace{\begin{bmatrix}\tilde{\mathbf{Q}} & \tilde{\mathbf{N}}^{\mathrm{T}}\\ \tilde{\mathbf{N}} & \tilde{\mathbf{R}}\end{bmatrix}}_{\tilde{\mathbf{J}}}\begin{bmatrix}\mathbf{x}_k\\ \mathbf{u}_k\end{bmatrix}\end{aligned} \tag{3.38}$$

auf die Form von (3.2)

$$J(\mathbf{x}_0) = \sum_{k=0}^{N-1}\begin{bmatrix}\mathbf{x}_k^{\mathrm{T}} & \mathbf{u}_k^{\mathrm{T}}\end{bmatrix}\underbrace{\begin{bmatrix}\tilde{\mathbf{Q}} & \tilde{\mathbf{N}}^{\mathrm{T}}\\ \tilde{\mathbf{N}} & \tilde{\mathbf{R}}\end{bmatrix}}_{\tilde{\mathbf{J}}}\begin{bmatrix}\mathbf{x}_k\\ \mathbf{u}_k\end{bmatrix} + \mathbf{x}_N^{\mathrm{T}}\mathbf{S}\mathbf{x}_N \tag{3.39}$$

überführt und mittels Satz 3.2 gelöst werden.

Hinweis: Mit Hilfe der Gewichtungsmatrix $\mathbf{J}$ des Gütefunktionals (3.2) bzw. (3.37) kann das Verhalten des geschlossenen Regelkreises gezielt beeinflusst werden. Als generelle Faustregel gilt: je größer die Einträge der Matrix $\mathbf{R}$ (Gewichtung der Stellgrößen) sind, desto kleiner werden die erforderlichen Stellgrößen. Weiters kann durch eine sehr hohe Gewichtung eines bestimmten Zustandes in $\mathbf{Q}$ bzw. $\tilde{\mathbf{Q}}$ erreicht werden, dass im geschlossenen Kreis dieser Zustand sehr schnell nach Null abklingt. Da sich diese Bewertung im Zeitdiskreten als oft sehr schwierig erweist, ist es sinnvoll, das Gütefunktional (3.2) im Zeitkontinuierlichen anzugeben und es anschließend ins Zeitdiskrete zu übersetzen. Die Idee besteht dabei darin, die Zustände und die Stellgrößen im Sinne einer „Leistung" in der Form

$$\int_0^T x^2(t)\,\mathrm{d}t \tag{3.40}$$

zu bewerten. Als Ausgangspunkt betrachte man das lineare, zeitinvariante, zeitkontinuierliche System

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} \tag{3.41a}$$

$$\mathbf{y} = \mathbf{C}\mathbf{x} + \mathbf{D}\mathbf{u} \tag{3.41b}$$

mit dem Zustand $\mathbf{x} \in \mathbb{R}^n$, dem Eingang $\mathbf{u} \in \mathbb{R}^p$ und dem Ausgang $\mathbf{y} \in \mathbb{R}^q$. Das zugehörige Abtastsystem (siehe Abschnitt 6 des Skripts Automatisierung) errechnet sich für die Abtastzeit T_a zu

$$\mathbf{x}_{k+1} = \mathbf{\Phi}\mathbf{x}_k + \mathbf{\Gamma}\mathbf{u}_k \tag{3.42a}$$

$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{D}\mathbf{u}_k \tag{3.42b}$$

mit

$$\mathbf{\Phi} = \exp(\mathbf{A}T_a) \quad \text{und} \quad \mathbf{\Gamma} = \int_0^{T_a} \exp(\mathbf{A}\tau)\,\mathrm{d}\tau\mathbf{B}\ . \tag{3.43}$$

Man wählt nun für das zeitkontinuierliche System (3.41) ein Gütefunktional der Form

$$\begin{aligned} J(\mathbf{x}_0) &= \int_0^T \Big(\mathbf{x}^{\mathrm{T}}(t)\mathbf{Q}_c\mathbf{x}(t) + \mathbf{u}^{\mathrm{T}}(t)\mathbf{R}_c\mathbf{u}(t) + 2\mathbf{u}^{\mathrm{T}}(t)\mathbf{N}_c\mathbf{x}(t)\Big)\,\mathrm{d}t + \mathbf{x}^{\mathrm{T}}(T)\mathbf{S}_c\mathbf{x}(T) \\ &= \int_0^T \begin{bmatrix}\mathbf{x}^{\mathrm{T}}(t) & \mathbf{u}^{\mathrm{T}}(t)\end{bmatrix} \underbrace{\begin{bmatrix}\mathbf{Q}_c & \mathbf{N}_c^{\mathrm{T}} \\ \mathbf{N}_c & \mathbf{R}_c\end{bmatrix}}_{\mathbf{J}_c} \begin{bmatrix}\mathbf{x}(t) \\ \mathbf{u}(t)\end{bmatrix} \mathrm{d}t + \mathbf{x}^{\mathrm{T}}(T)\mathbf{S}_c\mathbf{x}(T) \end{aligned} \tag{3.44}$$

mit den symmetrischen, positiv semi-definiten Gewichtungsmatrizen $\mathbf{J}_c$ und $\mathbf{S}_c$ sowie der Endzeit $T = NT_a$. Um (3.44) in die Form von (3.2) überzuführen, schreibt man vorerst (3.44) wie folgt

$$\begin{aligned} J(\mathbf{x}_0) = &\sum_{k=0}^{N-1} \int_{kT_a}^{(k+1)T_a} \Big(\mathbf{x}^{\mathrm{T}}(t)\mathbf{Q}_c\mathbf{x}(t) + \mathbf{u}^{\mathrm{T}}(t)\mathbf{R}_c\mathbf{u}(t) + 2\mathbf{u}^{\mathrm{T}}(t)\mathbf{N}_c\mathbf{x}(t)\Big)\,\mathrm{d}t \\ &+ \mathbf{x}^{\mathrm{T}}(NT_a)\mathbf{S}_c\mathbf{x}(NT_a) \end{aligned} \tag{3.45}$$

um. Berücksichtigt man die Tatsache, dass die Stellgröße $\mathbf{u}(t) = \mathbf{u}_k$ im Abtastintervall $kT_a \le t < (k+1)T_a$ konstant ist und für den Zustand $\mathbf{x}(t)$ gilt

$$\mathbf{x}(t) = \underbrace{\exp(\mathbf{A}(t-kT_a))}_{\mathbf{\Phi}(t-kT_a)}\mathbf{x}_k + \underbrace{\int_{kT_a}^{t} \exp(\mathbf{A}(t-\tau))\,\mathrm{d}\tau\mathbf{B}}_{\mathbf{\Gamma}(t-kT_a)}\mathbf{u}_k, \quad kT_a \le t < (k+1)T_a, \tag{3.46}$$

dann erhält man für (3.45)

$$J(\mathbf{x}_0) = \sum_{k=0}^{N-1} \Big(\mathbf{x}_k^T\mathbf{Q}\mathbf{x}_k + \mathbf{u}_k^{\mathrm{T}}\mathbf{R}\mathbf{u}_k + 2\mathbf{u}_k^{\mathrm{T}}\mathbf{N}\mathbf{x}_k\Big) + \mathbf{x}_N^{\mathrm{T}}\mathbf{S}\mathbf{x}_N \tag{3.47}$$

mit

$$\mathbf{Q} = \int_{kT_a}^{(k+1)T_a} \mathbf{\Phi}^{\mathrm{T}}(t-kT_a)\mathbf{Q}_c\mathbf{\Phi}(t-kT_a)\,\mathrm{d}t = \int_0^{T_a} \mathbf{\Phi}^{\mathrm{T}}(t)\mathbf{Q}_c\mathbf{\Phi}(t)\,\mathrm{d}t \tag{3.48a}$$

$$\begin{aligned}\mathbf{R} &= \int_{kT_a}^{(k+1)T_a} \Big(\mathbf{\Gamma}^{\mathrm{T}}(t-kT_a)\mathbf{Q}_c\mathbf{\Gamma}(t-kT_a) + 2\mathbf{N}_c\mathbf{\Gamma}(t-kT_a) + \mathbf{R}_c\Big)\,\mathrm{d}t \\ &= \int_0^{T_a} \Big(\mathbf{\Gamma}^{\mathrm{T}}(t)\mathbf{Q}_c\mathbf{\Gamma}(t) + 2\mathbf{N}_c\mathbf{\Gamma}(t) + \mathbf{R}_c\Big)\,\mathrm{d}t\end{aligned} \tag{3.48b}$$

$$\begin{aligned}\mathbf{N} &= \int_{kT_a}^{(k+1)T_a} \Big(\mathbf{\Gamma}^{\mathrm{T}}(t-kT_a)\mathbf{Q}_c + \mathbf{N}_c\Big)\mathbf{\Phi}(t-kT_a)\,\mathrm{d}t \\ &= \int_0^{T_a} \Big(\mathbf{\Gamma}^{\mathrm{T}}(t)\mathbf{Q}_c + \mathbf{N}_c\Big)\mathbf{\Phi}(t)\,\mathrm{d}t\end{aligned} \tag{3.48c}$$

$$\mathbf{S} = \mathbf{S}_c \tag{3.48d}$$

sowie

$$\mathbf{\Phi}(t) = \exp(\mathbf{A}t) \quad \text{und} \quad \mathbf{\Gamma}(t) = \int_0^t \exp(\mathbf{A}\tau)\,\mathrm{d}\tau\mathbf{B}\ . \tag{3.49}$$

Man beachte, dass hier im Allgemeinen die Kopplungsmatrix $\mathbf{N}$ von Null verschieden ist, selbst wenn $\mathbf{N}_c = \mathbf{0}$ gilt.

Aufgabe 3.3. Vergleichen Sie die MATLAB-Befehle `lqrd`, `dlqr` und `dlqry`. Was leisten diese Befehle? Sehen Sie sich den jeweiligen `help`-Text an und stellen Sie anschließend die Verbindung zur bisher gezeigten Theorie her.

3.3 Das LQR-Problem mit stochastischer Störung

Das Modell (3.1) gemäß (2.72) wird nun um die r-dimensionale stochastische Störung $\mathbf{w}$ erweitert

$$\mathbf{x}_{k+1} = \mathbf{\Phi}\mathbf{x}_k + \mathbf{\Gamma}\mathbf{u}_k + \mathbf{G}\mathbf{w}_k \qquad \mathbf{x}(0) = \mathbf{x}_0 \tag{3.50}$$

mit dem n-dimensionalen Zustand $\mathbf{x} \in \mathbb{R}^n$, dem p-dimensionalen deterministischen Eingang $\mathbf{u} \in \mathbb{R}^p$ und den Matrizen $\mathbf{\Phi} \in \mathbb{R}^{n\times n}$, $\mathbf{\Gamma} \in \mathbb{R}^{n\times p}$ und $\mathbf{G} \in \mathbb{R}^{n\times r}$. Es gelten nun folgende Annahmen:

(1) Von der Störung $\mathbf{w}$ wird vorausgesetzt, dass gilt

$$\mathrm{E}(\mathbf{w}_k) = \mathbf{0} \qquad \mathrm{E}\Big(\mathbf{w}_k\mathbf{w}_j^{\mathrm{T}}\Big) = \hat{\mathbf{Q}}\delta_{kj} \tag{3.51}$$

mit der Kovarianzmatrix $\hat{\mathbf{Q}} \geq 0$ und dem Kroneckersymbol $\delta_{kj} = 1$ für $k = j$ und $\delta_{kj} = 0$ sonst.

(2) Der Erwartungswert und die Kovarianzmatrix des Anfangswertes $\mathbf{x}_0$ sind durch

$$\mathrm{E}(\mathbf{x}_0) = \mathbf{m}_0 \qquad \mathrm{cov}(\mathbf{x}_0) = \mathrm{E}\Big([\mathbf{x}_0 - \mathbf{m}_0][\mathbf{x}_0 - \mathbf{m}_0]^{\mathrm{T}}\Big) = \hat{\mathbf{P}}_0 \geq 0 \tag{3.52}$$

gegeben.

(3) Die Störung $\mathbf{w}_k$, $k \geq 0$, ist mit dem Anfangswert $\mathbf{x}_0$ nicht korreliert, d. h. es gilt

$$\mathrm{E}\Big(\mathbf{w}_k \mathbf{x}_0^{\mathrm{T}}\Big) = \mathbf{0} \ . \tag{3.53}$$

Damit folgt aber analog zu (2.76) und (2.77) wegen

$$\mathbf{x}_j = \mathbf{\Phi}^j \mathbf{x}_0 + \sum_{l=0}^{j-1} \mathbf{\Phi}^l (\mathbf{\Gamma} \mathbf{u}_{j-1-l} + \mathbf{G} \mathbf{w}_{j-1-l}) \tag{3.54}$$

und (3.52) sowie (3.53) auch die Beziehung

$$\mathrm{E}\Big(\mathbf{w}_k \mathbf{x}_j^{\mathrm{T}}\Big) = \mathbf{0} \quad \text{für} \quad k \geq j \ . \tag{3.55}$$

Da nun $\mathbf{x}_k$, $k \geq 0$, im Gegensatz zu (3.1) ein stochastisches Signal ist, muss das Gütefunktional (3.2) durch

$$\begin{aligned} J(\mathbf{x}_0) &= \mathrm{E}\left(\sum_{k=0}^{N-1} \Big(\mathbf{x}_k^{\mathrm{T}} \mathbf{Q} \mathbf{x}_k + \mathbf{u}_k^{\mathrm{T}} \mathbf{R} \mathbf{u}_k + 2 \mathbf{u}_k^{\mathrm{T}} \mathbf{N} \mathbf{x}_k \Big) \right) + \mathrm{E}\Big(\mathbf{x}_N^{\mathrm{T}} \mathbf{S} \mathbf{x}_N \Big) \\ &= \mathrm{E}\left(\sum_{k=0}^{N-1} \begin{bmatrix} \mathbf{x}_k^{\mathrm{T}} & \mathbf{u}_k^{\mathrm{T}} \end{bmatrix} \underbrace{\begin{bmatrix} \mathbf{Q} & \mathbf{N}^{\mathrm{T}} \\ \mathbf{N} & \mathbf{R} \end{bmatrix}}_{\mathbf{J}} \begin{bmatrix} \mathbf{x}_k \\ \mathbf{u}_k \end{bmatrix} \right) + \mathrm{E}\Big(\mathbf{x}_N^{\mathrm{T}} \mathbf{S} \mathbf{x}_N \Big) \end{aligned} \tag{3.56}$$

ersetzt werden. Zur Berechnung des optimalen Regelgesetzes, also der Minimierung von (3.56) bezüglich $\mathbf{u}_0, \mathbf{u}_1, \ldots, \mathbf{u}_{N-1}$ unter der Nebenbedingung (3.50), bedient man sich wiederum der Methode der dynamischen Programmierung. Für $k = N$ gilt

$$J^*(\mathbf{x}_N) = \mathrm{E}\Big(\mathbf{x}_N^{\mathrm{T}} \mathbf{S} \mathbf{x}_N \Big) \tag{3.57}$$

und für $k = N - 1$ folgt

$$\begin{aligned} &J^*(\mathbf{x}_{N-1}) = \\ &\min_{\mathbf{u}_{N-1}} \mathrm{E}\left(\mathbf{x}_{N-1}^{\mathrm{T}} \mathbf{Q} \mathbf{x}_{N-1} + \mathbf{u}_{N-1}^{\mathrm{T}} \mathbf{R} \mathbf{u}_{N-1} + 2 \mathbf{u}_{N-1}^{\mathrm{T}} \mathbf{N} \mathbf{x}_{N-1} + J^*\left(\underbrace{\mathbf{x}_N}_{\mathbf{\Phi}\mathbf{x}_{N-1} + \mathbf{\Gamma}\mathbf{u}_{N-1} + \mathbf{G}\mathbf{w}_{N-1}} \right) \right) \end{aligned} \tag{3.58}$$

mit

$$J^*(\mathbf{x}_N) = \mathrm{E}\Big((\mathbf{\Phi}\mathbf{x}_{N-1} + \mathbf{\Gamma}\mathbf{u}_{N-1} + \mathbf{G}\mathbf{w}_{N-1})^{\mathrm{T}} \mathbf{S} (\mathbf{\Phi}\mathbf{x}_{N-1} + \mathbf{\Gamma}\mathbf{u}_{N-1} + \mathbf{G}\mathbf{w}_{N-1}) \Big) \ . \tag{3.59}$$

Minimiert man (3.58) bezüglich $\mathbf{u}_{N-1}$, ergibt sich die optimale Lösung $\mathbf{u}_{N-1}^*$ von $\mathbf{u}_{N-1}$ zu

$$\mathbf{u}_{N-1}^* = \mathbf{K}_{N-1} \mathbf{x}_{N-1} \tag{3.60}$$

mit

$$\mathbf{K}_{N-1} = -\Big(\mathbf{R} + \mathbf{\Gamma}^{\mathrm{T}}\mathbf{P}_N\mathbf{\Gamma}\Big)^{-1}\Big(\mathbf{N} + \mathbf{\Gamma}^{\mathrm{T}}\mathbf{P}_N\mathbf{\Phi}\Big), \tag{3.61}$$

wobei $\mathbf{S} = \mathbf{P}_N$ gesetzt wurde und die Bedingung $\Big(\mathbf{R} + \mathbf{\Gamma}^{\mathrm{T}}\mathbf{P}_N\mathbf{\Gamma}\Big) > 0$ erfüllt sein muss. Durch Einsetzen von (3.60) in (3.58) erhält man

$$\begin{aligned} J^*(\mathbf{x}_{N-1}) = \\ \mathrm{E}\Bigg(\mathbf{x}_{N-1}^{\mathrm{T}} \underbrace{\Big(\Big(\mathbf{Q} + \mathbf{\Phi}^{\mathrm{T}}\mathbf{P}_N\mathbf{\Phi}\Big) + \mathbf{K}_{N-1}^{\mathrm{T}}\Big(\mathbf{R} + \mathbf{\Gamma}^{\mathrm{T}}\mathbf{P}_N\mathbf{\Gamma}\Big)\mathbf{K}_{N-1} + 2\mathbf{K}_{N-1}^{\mathrm{T}}\Big(\mathbf{N} + \mathbf{\Gamma}^{\mathrm{T}}\mathbf{P}_N\mathbf{\Phi}\Big)\Big)}_{\mathbf{P}_{N-1}} \mathbf{x}_{N-1} \Bigg) \\ + \underbrace{2\,\mathrm{E}\Big(\mathbf{w}_{N-1}^{\mathrm{T}}\mathbf{G}^{\mathrm{T}}\mathbf{P}_N(\mathbf{\Phi}\mathbf{x}_{N-1} + \mathbf{\Gamma}\mathbf{u}_{N-1})\Big)}_{2\,\mathrm{spur}\left(\mathbf{P}_N\mathbf{G}\,\mathrm{E}\left(\mathbf{w}_{N-1}\mathbf{x}_{N-1}^{\mathrm{T}}\right)\mathbf{\Phi}^{\mathrm{T}}\right)+2\,\mathrm{E}\left(\mathbf{w}_{N-1}^{\mathrm{T}}\right)\mathbf{G}^{\mathrm{T}}\mathbf{P}_N\mathbf{\Gamma}\mathbf{u}_{N-1}} + \underbrace{\mathrm{E}\Big(\mathbf{w}_{N-1}^{\mathrm{T}}\mathbf{G}^{\mathrm{T}}\mathbf{P}_N\mathbf{G}\mathbf{w}_{N-1}\Big)}_{\mathrm{spur}\left(\mathbf{P}_N\mathbf{G}\,\mathrm{E}\left(\mathbf{w}_{N-1}\mathbf{w}_{N-1}^{\mathrm{T}}\right)\mathbf{G}^{\mathrm{T}}\right)} . \end{aligned} \tag{3.62}$$

Damit lässt sich unmittelbar folgendes Ergebnis angeben: Das optimale Regelgesetz $\mathbf{u}_k^*$ für das System mit stochastischer Störung (3.50) entspricht dem des ungestörten Systems und lautet (siehe (3.22))

$$\mathbf{u}_k^* = \mathbf{K}_k\mathbf{x}_k \tag{3.63}$$

mit

$$\mathbf{K}_k = -\Big(\mathbf{R} + \mathbf{\Gamma}^{\mathrm{T}}\mathbf{P}_{k+1}\mathbf{\Gamma}\Big)^{-1}\Big(\mathbf{N} + \mathbf{\Gamma}^{\mathrm{T}}\mathbf{P}_{k+1}\mathbf{\Phi}\Big) \tag{3.64a}$$

$$\mathbf{P}_k = \Big(\mathbf{Q} + \mathbf{\Phi}^{\mathrm{T}}\mathbf{P}_{k+1}\mathbf{\Phi}\Big) - \Big(\mathbf{N} + \mathbf{\Gamma}^{\mathrm{T}}\mathbf{P}_{k+1}\mathbf{\Phi}\Big)^{\mathrm{T}}\Big(\mathbf{R} + \mathbf{\Gamma}^{\mathrm{T}}\mathbf{P}_{k+1}\mathbf{\Gamma}\Big)^{-1}\Big(\mathbf{N} + \mathbf{\Gamma}^{\mathrm{T}}\mathbf{P}_{k+1}\mathbf{\Phi}\Big) . \tag{3.64b}$$

Der minimale Wert des Gütefunktionals (3.56) errechnet sich nach (3.62) zu

$$\min_{(\mathbf{u}_0,\ldots,\mathbf{u}_{N-1})} J(\mathbf{x}_0) = J^*(\mathbf{x}_0) = \mathrm{E}\Big(\mathbf{x}_0^{\mathrm{T}}\mathbf{P}_0\mathbf{x}_0\Big) + \sum_{k=0}^{N-1} \mathrm{spur}\Big(\mathbf{P}_{k+1}\mathbf{G}\hat{\mathbf{Q}}\mathbf{G}^{\mathrm{T}}\Big) . \tag{3.65}$$

Aufgabe 3.4. Beweisen Sie die Gültigkeit von Beziehung (3.65). Zeigen Sie weiters, dass der Ausdruck $\mathrm{E}\Big(\mathbf{x}_0^{\mathrm{T}}\mathbf{P}_0\mathbf{x}_0\Big)$ wie folgt

$$\mathrm{E}\Big(\mathbf{x}_0^{\mathrm{T}}\mathbf{P}_0\mathbf{x}_0\Big) = \mathbf{m}_0^{\mathrm{T}}\mathbf{P}_0\mathbf{m}_0 + \mathrm{spur}\Big(\mathbf{P}_0\hat{\mathbf{P}}_0\Big) \tag{3.66}$$

mit

$$\hat{\mathbf{P}}_0 = \mathrm{cov}(\mathbf{x}_0) \quad \text{und} \quad \mathbf{m}_0 = \mathrm{E}(\mathbf{x}_0) \tag{3.67}$$

vereinfacht werden kann.

Aus (3.65), (3.66) und (3.67) folgt also, dass der minimale Wert des Gütefunktionals bei dem System mit stochastischer Störung größer als der des ungestörten Systems (siehe (3.25)) ist. Hinzu kommen nämlich zwei zusätzliche Terme: ein Term mit $\hat{\mathbf{Q}}$ zufolge der

stochastischen Störung $\mathbf{w}$ und ein Term mit $\hat{\mathbf{P}}_0$ zufolge des stochastischen Charakters des Anfangswertes $\mathbf{x}_0$. In Satz 3.2 wurde gezeigt, dass $\mathbf{P}_k \geq 0$ ist für alle $k \geq 0$, sofern $\mathbf{S}$ und $\mathbf{J}$ von (3.56) positiv semi-definit sind. Damit kann aber, wie man aus (3.65) erkennt, das Minimierungsproblem mit dem Gütefunktional (3.56) für $N \to \infty$ nicht gelöst werden. Um trotzdem auch für den Fall mit stochastischer Störung den stationären Riccati-Regler (3.29) berechnen zu können, setzt man das Gütefunktional in der Form

$$J(\mathbf{x}_0) = \lim_{N\to\infty} \frac{1}{N} \mathrm{E}\left(\sum_{k=0}^{N-1} \left(\mathbf{x}_k^{\mathrm{T}} \mathbf{Q} \mathbf{x}_k + \mathbf{u}_k^{\mathrm{T}} \mathbf{R} \mathbf{u}_k + 2\mathbf{u}_k^{\mathrm{T}} \mathbf{N} \mathbf{x}_k \right) \right) \tag{3.68}$$

an.

Aufgabe 3.5. Zeigen Sie, dass der stationäre Riccati-Regler (3.29) das Gütefunktional (3.68) mit der Nebenbedingung (3.50) minimiert und für den minimalen Wert des Gütefunktionals gilt

$$J^*(\mathbf{x}_0) = \mathrm{spur}\left(\mathbf{P}_s \mathbf{G} \hat{\mathbf{Q}} \mathbf{G}^{\mathrm{T}} \right)$$

mit $\mathbf{P}_s$ als Lösung der diskreten algebraischen Riccati-Gleichung (3.28).

3.4 Das LQG-Regelungsproblem

Im Folgenden sei angenommen, dass lediglich der Ausgang $\mathbf{y}$ gemessen werden kann, der Zustand $\mathbf{x}$ hingegen nicht zur Verfügung steht. Dazu betrachte man das auch dem Kalman-Filter zugrundeliegende (siehe (2.72)) zeitdiskrete, lineare, zeitinvariante System der Form

$$\mathbf{x}_{k+1} = \mathbf{\Phi} \mathbf{x}_k + \mathbf{\Gamma} \mathbf{u}_k + \mathbf{G} \mathbf{w}_k \qquad \mathbf{x}(0) = \mathbf{x}_0 \tag{3.69a}$$

$$\mathbf{y}_k = \mathbf{C} \mathbf{x}_k + \mathbf{D} \mathbf{u}_k + \mathbf{v}_k \tag{3.69b}$$

mit dem n-dimensionalen Zustand $\mathbf{x} \in \mathbb{R}^n$, dem p-dimensionalen deterministischen Eingang $\mathbf{u} \in \mathbb{R}^p$, dem q-dimensionalen Ausgang $\mathbf{y} \in \mathbb{R}^q$, der r-dimensionalen Störung $\mathbf{w} \in \mathbb{R}^r$, dem Messrauschen $\mathbf{v}$ sowie den Matrizen $\mathbf{\Phi} \in \mathbb{R}^{n\times n}$, $\mathbf{\Gamma} \in \mathbb{R}^{n\times p}$, $\mathbf{G} \in \mathbb{R}^{n\times r}$, $\mathbf{C} \in \mathbb{R}^{q\times n}$ und $\mathbf{D} \in \mathbb{R}^{q\times p}$. Es gelten nun wieder folgende Annahmen:

(1) Von der Störung $\mathbf{w}$ und vom Messrauschen $\mathbf{v}$ wird vorausgesetzt, dass gilt

$$\mathrm{E}(\mathbf{v}_k) = \mathbf{0} \qquad \mathrm{E}\left(\mathbf{w}_k \mathbf{w}_j^{\mathrm{T}} \right) = \hat{\mathbf{Q}} \delta_{kj} \tag{3.70a}$$

$$\mathrm{E}(\mathbf{w}_k) = \mathbf{0} \qquad \mathrm{E}\left(\mathbf{v}_k \mathbf{v}_j^{\mathrm{T}} \right) = \hat{\mathbf{R}} \delta_{kj} \tag{3.70b}$$

$$\mathrm{E}\left(\mathbf{w}_k \mathbf{v}_j^{\mathrm{T}} \right) = \mathbf{0} \tag{3.70c}$$

mit $\hat{\mathbf{Q}} \geq 0$ sowie $\hat{\mathbf{R}} > 0$ und dem Kroneckersymbol $\delta_{kj} = 1$ für $k = j$ und $\delta_{kj} = 0$ sonst.

(2) Der Erwartungswert des Anfangswertes und die Kovarianzmatrix des Anfangsfehlers sind durch

$$\mathrm{E}(\mathbf{x}_0) = \mathbf{m}_0 \qquad \mathrm{E}\left([\mathbf{x}_0 - \hat{\mathbf{x}}_0][\mathbf{x}_0 - \hat{\mathbf{x}}_0]^{\mathrm{T}} \right) = \hat{\mathbf{P}}_0 \geq 0 \tag{3.71}$$

mit dem Schätzwert $\hat{\mathbf{x}}_0$ des Anfangswertes $\mathbf{x}_0$ gegeben.

(3) Die Störung $\mathbf{w}_k$, $k \geq 0$, und das Messrauschen $\mathbf{v}_l$, $l \geq 0$, sind mit dem Anfangswert $\mathbf{x}_0$ nicht korreliert, d. h. es gilt

$$\mathrm{E}\Big(\mathbf{w}_k \mathbf{x}_0^{\mathrm{T}}\Big) = \mathbf{0} \tag{3.72a}$$

$$\mathrm{E}\Big(\mathbf{v}_l \mathbf{x}_0^{\mathrm{T}}\Big) = \mathbf{0}\ . \tag{3.72b}$$

Damit folgt aber wegen (3.54) und (3.70) auch die Beziehung

$$\mathrm{E}\Big(\mathbf{w}_k \mathbf{x}_j^{\mathrm{T}}\Big) = \mathbf{0} \quad \text{für} \quad k \geq j \tag{3.73a}$$

$$\mathrm{E}\Big(\mathbf{v}_l \mathbf{x}_j^{\mathrm{T}}\Big) = \mathbf{0} \quad \text{für alle } l, j\ . \tag{3.73b}$$

Die Regelungsaufgabe besteht nun darin, das Gütefunktional

$$\begin{aligned} J(\mathbf{x}_0) &= \mathrm{E}\left(\sum_{k=0}^{N-1}\Big(\mathbf{x}_k^{\mathrm{T}}\mathbf{Q}\mathbf{x}_k + \mathbf{u}_k^{\mathrm{T}}\mathbf{R}\mathbf{u}_k + 2\mathbf{u}_k^{\mathrm{T}}\mathbf{N}\mathbf{x}_k\Big)\right) + \mathrm{E}\Big(\mathbf{x}_N^{\mathrm{T}}\mathbf{S}\mathbf{x}_N\Big) \\ &= \mathrm{E}\left(\sum_{k=0}^{N-1}\begin{bmatrix}\mathbf{x}_k^{\mathrm{T}} & \mathbf{u}_k^{\mathrm{T}}\end{bmatrix}\underbrace{\begin{bmatrix}\mathbf{Q} & \mathbf{N}^{\mathrm{T}} \\ \mathbf{N} & \mathbf{R}\end{bmatrix}}_{\mathbf{J}}\begin{bmatrix}\mathbf{x}_k \\ \mathbf{u}_k\end{bmatrix}\right) + \mathrm{E}\Big(\mathbf{x}_N^{\mathrm{T}}\mathbf{S}\mathbf{x}_N\Big) \end{aligned} \tag{3.74}$$

für die positiv semi-definiten Gewichtungsmatrizen $\mathbf{J}$ und $\mathbf{S}$ bezüglich $\mathbf{u}_0, \mathbf{u}_1, \ldots, \mathbf{u}_{N-1}$ unter der Nebenbedingung (3.69) so zu minimieren, dass das Regelgesetz nur von den messbaren Ausgangsgrößen $\mathbf{y}$ abhängt.

Bezeichnet man mit $\hat{\mathbf{x}}_k = \hat{\mathbf{x}}(k|k-1)$ den Schätzwert von $\mathbf{x}_k$ unter Zuhilfenahme von $0, 1, \ldots, k-1$ Messungen (vgl. Definition 2.3), dann folgt

$$\begin{aligned} &\mathrm{E}\Big(\mathbf{x}_k^{\mathrm{T}}\mathbf{Q}\mathbf{x}_k\Big) = \mathrm{E}\Big((\hat{\mathbf{x}}_k - (\hat{\mathbf{x}}_k - \mathbf{x}_k))^{\mathrm{T}}\mathbf{Q}(\hat{\mathbf{x}}_k - (\hat{\mathbf{x}}_k - \mathbf{x}_k))\Big) \\ &= \mathrm{E}\Big(\hat{\mathbf{x}}_k^{\mathrm{T}}\mathbf{Q}\hat{\mathbf{x}}_k\Big) - 2\,\mathrm{E}\Big(\hat{\mathbf{x}}_k^{\mathrm{T}}\mathbf{Q}(\hat{\mathbf{x}}_k - \mathbf{x}_k)\Big) + \mathrm{E}\Big((\hat{\mathbf{x}}_k - \mathbf{x}_k)^{\mathrm{T}}\mathbf{Q}(\hat{\mathbf{x}}_k - \mathbf{x}_k)\Big) \\ &= \mathrm{E}\Big(\hat{\mathbf{x}}_k^{\mathrm{T}}\mathbf{Q}\hat{\mathbf{x}}_k\Big) - 2\,\mathrm{spur}\Big(\mathbf{Q}\,\mathrm{E}\Big(\hat{\mathbf{x}}_k(\hat{\mathbf{x}}_k - \mathbf{x}_k)^{\mathrm{T}}\Big)\Big) + \mathrm{spur}\Big(\mathbf{Q}\,\mathrm{E}\Big((\hat{\mathbf{x}}_k - \mathbf{x}_k)(\hat{\mathbf{x}}_k - \mathbf{x}_k)^{\mathrm{T}}\Big)\Big)\ . \end{aligned} \tag{3.75}$$

Im Weiteren werden nur Schätzer in Betracht gezogen, für die die Beziehung

$$\mathrm{E}\Big(\hat{\mathbf{x}}_k(\hat{\mathbf{x}}_k - \mathbf{x}_k)^{\mathrm{T}}\Big) = \mathbf{0} \tag{3.76}$$

gilt, also der zweite Term in Ausdruck (3.75) Null wird. Nach Satz 2.4 bzw. Aufgabe 2.5 erfüllt dies der Minimum-Varianz-Schätzer und damit auch der Gauß-Markov-Schätzer und natürlich das Kalman-Filter, vgl. dazu Aufgabe 2.5. Mittels der dynamischen Programmierung erhält man nun in einem ersten Schritt für $\mathbf{S} = \mathbf{P}_N$ den minimalen Wert des Gütefunktionals $J^*(\mathbf{x}_N)$ von $J(\mathbf{x}_N)$ zu

$$J^*(\mathbf{x}_N) = \mathrm{E}\Big(\mathbf{x}_N^{\mathrm{T}}\mathbf{P}_N\mathbf{x}_N\Big) = \mathrm{spur}\Big(\mathbf{P}_N\,\mathrm{E}\Big(\mathbf{x}_N\mathbf{x}_N^{\mathrm{T}}\Big)\Big) \tag{3.77}$$

bzw. mit $\mathbf{x}_N = \mathbf{x}_N - \hat{\mathbf{x}}_N + \hat{\mathbf{x}}_N$ auch

$$J^*\Big(\hat{\mathbf{x}}_N, \hat{\mathbf{P}}_N\Big) = \mathrm{E}\Big(\mathbf{x}_N^{\mathrm{T}}\mathbf{P}_N\mathbf{x}_N\Big) = \mathrm{E}\Big(\hat{\mathbf{x}}_N^{\mathrm{T}}\mathbf{P}_N\hat{\mathbf{x}}_N\Big) + \mathrm{spur}\Big(\mathbf{P}_N\hat{\mathbf{P}}_N\Big) \tag{3.78}$$

mit der Kovarianzmatrix des Schätzfehlers

$$\hat{\mathbf{P}}_N = \mathrm{cov}(\hat{\mathbf{x}}_N - \mathbf{x}_N) = \mathrm{E}\Big((\hat{\mathbf{x}}_N - \mathbf{x}_N)(\hat{\mathbf{x}}_N - \mathbf{x}_N)^{\mathrm{T}}\Big)\ . \tag{3.79}$$

Im nächsten Schritt der dynamischen Programmierung muss nun nachfolgende Minimierungsaufgabe

$$J^*(\mathbf{x}_{N-1}) = \min_{\mathbf{u}_{N-1}} \mathrm{E}\left(\mathbf{x}_{N-1}^{\mathrm{T}}\mathbf{Q}\mathbf{x}_{N-1} + \mathbf{u}_{N-1}^{\mathrm{T}}\mathbf{R}\mathbf{u}_{N-1} + 2\mathbf{u}_{N-1}^{\mathrm{T}}\mathbf{N}\mathbf{x}_{N-1} + J^*\Big(\underbrace{\mathbf{x}_N}_{\mathbf{\Phi}\mathbf{x}_{N-1}+\mathbf{\Gamma}\mathbf{u}_{N-1}+\mathbf{G}\mathbf{w}_{N-1}}\Big)\right) \tag{3.80}$$

mit

$$J^*(\mathbf{x}_N) = \mathrm{E}\Big((\mathbf{\Phi}\mathbf{x}_{N-1} + \mathbf{\Gamma}\mathbf{u}_{N-1} + \mathbf{G}\mathbf{w}_{N-1})^{\mathrm{T}}\mathbf{P}_N(\mathbf{\Phi}\mathbf{x}_{N-1} + \mathbf{\Gamma}\mathbf{u}_{N-1} + \mathbf{G}\mathbf{w}_{N-1})\Big) \tag{3.81}$$

gelöst werden. Da $\mathbf{u}_{N-1}$ aber nicht von $\mathbf{x}_{N-1}$ abhängen darf, löst man die Minimierungsaufgabe bezüglich eines Schätzwertes $\hat{\mathbf{x}}_{N-1}$, indem man in (3.80), (3.81)

$$\mathbf{x}_{N-1} = \mathbf{x}_{N-1} - \hat{\mathbf{x}}_{N-1} + \hat{\mathbf{x}}_{N-1} \tag{3.82}$$

setzt. Das Regelgesetz mit der Reglerverstärkungsmatrix $\mathbf{K}_{N-1}$ ist identisch zu (3.60) mit (3.61) und lautet

$$\mathbf{u}_{N-1}^* = \mathbf{K}_{N-1}\hat{\mathbf{x}}_{N-1} \tag{3.83}$$

und

$$\mathbf{K}_{N-1} = -\Big(\mathbf{R} + \mathbf{\Gamma}^{\mathrm{T}}\mathbf{P}_N\mathbf{\Gamma}\Big)^{-1}\Big(\mathbf{N} + \mathbf{\Gamma}^{\mathrm{T}}\mathbf{P}_N\mathbf{\Phi}\Big) \tag{3.84}$$

mit dem minimalen Wert des Gütefunktionals

$$\begin{aligned} J^*(\mathbf{x}_{N-1}) = \mathrm{E}\Bigg(\mathbf{x}_{N-1}^{\mathrm{T}} \underbrace{\Big(\Big(\mathbf{Q} + \mathbf{\Phi}^{\mathrm{T}}\mathbf{P}_N\mathbf{\Phi}\Big) + \mathbf{K}_{N-1}^{\mathrm{T}}\Big(\mathbf{R} + \mathbf{\Gamma}^{\mathrm{T}}\mathbf{P}_N\mathbf{\Gamma}\Big)\mathbf{K}_{N-1} + 2\mathbf{K}_{N-1}^{\mathrm{T}}\Big(\mathbf{N} + \mathbf{\Gamma}^{\mathrm{T}}\mathbf{P}_N\mathbf{\Phi}\Big)\Big)}_{\mathbf{P}_{N-1}} \mathbf{x}_{N-1}\Bigg) \\ + \underbrace{2\,\mathrm{E}\Big(\mathbf{w}_{N-1}^{\mathrm{T}}\mathbf{G}^{\mathrm{T}}\mathbf{P}_N(\mathbf{\Phi}\mathbf{x}_{N-1} + \mathbf{\Gamma}\mathbf{u}_{N-1})\Big)}_{2\,\mathrm{spur}\left(\mathbf{P}_N\mathbf{G}\,\mathrm{E}\left(\mathbf{w}_{N-1}\mathbf{x}_{N-1}^{\mathrm{T}}\right)\mathbf{\Phi}^{\mathrm{T}}\right)+2\,\mathrm{spur}\left(\mathbf{P}_N\mathbf{G}\,\mathrm{E}\left(\mathbf{w}_{N-1}\hat{\mathbf{x}}_{N-1}^{\mathrm{T}}\right)\mathbf{K}_{N-1}^{\mathrm{T}}\mathbf{\Gamma}^{\mathrm{T}}\right)} + \underbrace{\mathrm{E}\Big(\mathbf{w}_{N-1}^{\mathrm{T}}\mathbf{G}^{\mathrm{T}}\mathbf{P}_N\mathbf{G}\mathbf{w}_{N-1}\Big)}_{\mathrm{spur}\left(\mathbf{P}_N\mathbf{G}\,\mathrm{E}\left(\mathbf{w}_{N-1}\mathbf{w}_{N-1}^{\mathrm{T}}\right)\mathbf{G}^{\mathrm{T}}\right)} \\ + \underbrace{\mathrm{E}\Big((\mathbf{x}_{N-1} - \hat{\mathbf{x}}_{N-1})^{\mathrm{T}}\mathbf{K}_{N-1}^{\mathrm{T}}\Big(\mathbf{R} + \mathbf{\Gamma}^{\mathrm{T}}\mathbf{P}_N\mathbf{\Gamma}\Big)\mathbf{K}_{N-1}(\mathbf{x}_{N-1} - \hat{\mathbf{x}}_{N-1})\Big)}_{\mathrm{spur}\left(\mathrm{E}\left((\mathbf{x}_{N-1}-\hat{\mathbf{x}}_{N-1})(\mathbf{x}_{N-1}-\hat{\mathbf{x}}_{N-1})^{\mathrm{T}}\right)\mathbf{K}_{N-1}^{\mathrm{T}}(\mathbf{R}+\mathbf{\Gamma}^{\mathrm{T}}\mathbf{P}_N\mathbf{\Gamma})\mathbf{K}_{N-1}\right)} \end{aligned} \tag{3.85}$$

Aus der Rekursion erhält man schlussendlich mit $\mathrm{E}\left(\mathbf{w}_k\mathbf{w}_k^{\mathrm{T}}\right) = \hat{\mathbf{Q}}$ und (3.79) den minimalen Wert des Gütefunktionals für alle Stellfolgenwerte zu (man vergleiche das Ergebnis mit (3.65))

$$\begin{aligned}\min_{(\mathbf{u}_0,\ldots,\mathbf{u}_{N-1})} J(\mathbf{x}_0) = J^*(\mathbf{x}_0) = \mathrm{E}\left(\mathbf{x}_0^{\mathrm{T}}\mathbf{P}_0\mathbf{x}_0\right) + \sum_{k=0}^{N-1} \mathrm{spur}\left(\mathbf{P}_{k+1}\mathbf{G}\hat{\mathbf{Q}}\mathbf{G}^{\mathrm{T}}\right) \\ + \sum_{k=0}^{N-1} \mathrm{spur}\left(\hat{\mathbf{P}}_k\mathbf{K}_k^{\mathrm{T}}\left(\mathbf{R} + \boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{P}_{k+1}\boldsymbol{\Gamma}\right)\mathbf{K}_k\right) .\end{aligned} \tag{3.86}$$

Man erkennt, dass (3.86) unabhängig von der Art des Schätzers ist, sofern dieser die Bedingung (3.76) sowie $\mathrm{E}\left(\mathbf{w}_j\hat{\mathbf{x}}_0^{\mathrm{T}}\right)$ erfüllt. Diese Eigenschaft gemeinsam mit der Tatsache, dass die Pole beim Zustandsbeobachter- und Zustandsreglerentwurf getrennt voneinander vorgegeben werden können, bezeichnet man auch als das *Separationstheorem für den optimalen Zustandsbeobachter- und Zustandsreglerentwurf.*

Da der Ausdruck $\mathbf{K}_k^{\mathrm{T}}\left(\mathbf{R} + \boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{P}_{k+1}\boldsymbol{\Gamma}\right)\mathbf{K}_k$ für alle $k \geq 0$ positiv definit ist, kann man den dritten Term in (3.86) (bzw. den letzten Term in (3.85)) auch in der Form

$$\sum_{k=0}^{N-1} \mathrm{spur}\left(\hat{\mathbf{P}}_k\mathbf{K}_k^{\mathrm{T}}\left(\mathbf{R} + \boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{P}_{k+1}\boldsymbol{\Gamma}\right)\mathbf{K}_k\right) = \sum_{k=0}^{N-1} \mathrm{E}\left((\mathbf{x}_k - \hat{\mathbf{x}}_k)^{\mathrm{T}}\bar{\mathbf{K}}_k^{\mathrm{T}}\bar{\mathbf{K}}_k(\mathbf{x}_k - \hat{\mathbf{x}}_k)\right) \tag{3.87}$$

mit der Cholesky-Zerlegung

$$\bar{\mathbf{K}}_k^{\mathrm{T}}\bar{\mathbf{K}}_k = \mathbf{K}_k^{\mathrm{T}}\left(\mathbf{R} + \boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{P}_{k+1}\boldsymbol{\Gamma}\right)\mathbf{K}_k \tag{3.88}$$

anschreiben. Nach Satz 2.5 ist die lineare Minimum-Varianz-Schätzung einer linearen Funktion eines Parametervektors äquivalent der linearen Funktion der Minimum-Varianz-Schätzung des Parametervektors selbst, weswegen der Wert des Gütefunktionals (3.86) durch die Wahl eines Minimum-Varianz-Schätzers kleinstmöglich gemacht werden kann. Da das Kalman-Filter auf der rekursiven Minimum-Varianz-Schätzung beruht, wird das optimale LQG-Regelungsproblem durch die Kombination eines LQR-Zustandsreglers und eines Kalman-Filter Beobachters gelöst.

Dazu sei der *dynamische Regler mit optimaler Ausgangsrückführung* für das System (3.69) nach (3.83)–(3.85) und Satz 2.7 im Folgenden nochmals in der Form

$$\hat{\mathbf{x}}_{k+1} = \boldsymbol{\Phi}\hat{\mathbf{x}}_k + \boldsymbol{\Gamma}\mathbf{u}_k + \hat{\mathbf{K}}_k(\mathbf{y}_k - \mathbf{C}\hat{\mathbf{x}}_k - \mathbf{D}\mathbf{u}_k) \qquad \hat{\mathbf{x}}(0) = \hat{\mathbf{x}}_0 \tag{3.89a}$$

$$\mathbf{u}_k = \mathbf{K}_k\hat{\mathbf{x}}_k \tag{3.89b}$$

mit

$$\mathbf{P}_k = \left(\mathbf{Q} + \boldsymbol{\Phi}^{\mathrm{T}}\mathbf{P}_{k+1}\boldsymbol{\Phi}\right) - \left(\mathbf{N} + \boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{P}_{k+1}\boldsymbol{\Phi}\right)^{\mathrm{T}}\left(\mathbf{R} + \boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{P}_{k+1}\boldsymbol{\Gamma}\right)^{-1}\left(\mathbf{N} + \boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{P}_{k+1}\boldsymbol{\Phi}\right) \tag{3.90a}$$

$$\mathbf{K}_k = -\left(\mathbf{R} + \boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{P}_{k+1}\boldsymbol{\Gamma}\right)^{-1}\left(\mathbf{N} + \boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{P}_{k+1}\boldsymbol{\Phi}\right) \tag{3.90b}$$

$$\hat{\mathbf{P}}_{k+1} = \boldsymbol{\Phi}\hat{\mathbf{P}}_k\boldsymbol{\Phi}^{\mathrm{T}} + \mathbf{G}\hat{\mathbf{Q}}\mathbf{G}^{\mathrm{T}} - \boldsymbol{\Phi}\hat{\mathbf{P}}_k\mathbf{C}^{\mathrm{T}}\left(\mathbf{C}\hat{\mathbf{P}}_k\mathbf{C}^{\mathrm{T}} + \hat{\mathbf{R}}\right)^{-1}\mathbf{C}\hat{\mathbf{P}}_k\boldsymbol{\Phi}^{\mathrm{T}} \tag{3.90c}$$

$$\hat{\mathbf{K}}_k = \boldsymbol{\Phi}\hat{\mathbf{P}}_k\mathbf{C}^{\mathrm{T}}\left(\mathbf{C}\hat{\mathbf{P}}_k\mathbf{C}^{\mathrm{T}} + \hat{\mathbf{R}}\right)^{-1}, \tag{3.90d}$$

der Randbedingung $\mathbf{P}_N = \mathbf{S} \geq 0$ und der Anfangsbedingung $\hat{\mathbf{P}}_0 \geq 0$ zusammengefasst.

Man überzeugt sich leicht, dass für $\mathbf{N} = \mathbf{0}$ und $\mathbf{G} = \mathbf{E}$ der *LQR-Zustandsregler* und das *Kalman-Filter* gemäß der Definition im Skript Automatisierung *dual* sind. Es gilt nämlich dann für (3.90)

$$\mathbf{P}_k = \mathbf{Q} + \mathbf{\Phi}^{\mathrm{T}}\mathbf{P}_{k+1}\mathbf{\Phi} - \mathbf{\Phi}^{\mathrm{T}}\mathbf{P}_{k+1}\mathbf{\Gamma}\Big(\mathbf{R} + \mathbf{\Gamma}^{\mathrm{T}}\mathbf{P}_{k+1}\mathbf{\Gamma}\Big)^{-1}\mathbf{\Gamma}^{\mathrm{T}}\mathbf{P}_{k+1}\mathbf{\Phi} \tag{3.91a}$$

$$\mathbf{K}_k = -\Big(\mathbf{R} + \mathbf{\Gamma}^{\mathrm{T}}\mathbf{P}_{k+1}\mathbf{\Gamma}\Big)^{-1}\mathbf{\Gamma}^{\mathrm{T}}\mathbf{P}_{k+1}\mathbf{\Phi} \tag{3.91b}$$

$$\hat{\mathbf{P}}_{k+1} = \hat{\mathbf{Q}} + \mathbf{\Phi}\hat{\mathbf{P}}_k\mathbf{\Phi}^{\mathrm{T}} - \mathbf{\Phi}\hat{\mathbf{P}}_k\mathbf{C}^{\mathrm{T}}\Big(\hat{\mathbf{R}} + \mathbf{C}\hat{\mathbf{P}}_k\mathbf{C}^{\mathrm{T}}\Big)^{-1}\mathbf{C}\hat{\mathbf{P}}_k\mathbf{\Phi}^{\mathrm{T}} \tag{3.91c}$$

$$\hat{\mathbf{K}}_k = \mathbf{\Phi}\hat{\mathbf{P}}_k\mathbf{C}^{\mathrm{T}}\Big(\hat{\mathbf{R}} + \mathbf{C}\hat{\mathbf{P}}_k\mathbf{C}^{\mathrm{T}}\Big)^{-1}, \tag{3.91d}$$

d. h. die diskrete Riccati-Gleichung des Zustandsreglers (3.91a) geht für $\mathbf{\Phi}^{\mathrm{T}} = \mathbf{\Phi}$, $\mathbf{\Gamma} = \mathbf{C}^{\mathrm{T}}$, $\mathbf{P} = \hat{\mathbf{P}}$, $\mathbf{R} = \hat{\mathbf{R}}$ und $\mathbf{Q} = \hat{\mathbf{Q}}$ in die diskrete Riccati-Gleichung des Kalman-Filters (3.91c) über. Der einzige Unterschied besteht darin, dass die Riccati-Gleichung des Zustandsreglers rückwärts und die des Kalman-Filters vorwärts läuft. Weiters gilt für $\mathbf{\Phi}^{\mathrm{T}} = \mathbf{\Phi}$, $\mathbf{\Gamma} = \mathbf{C}^{\mathrm{T}}$, $\mathbf{P} = \hat{\mathbf{P}}$ und $\mathbf{R} = \hat{\mathbf{R}}$ die Beziehung $\mathbf{K}_k = \hat{\mathbf{K}}_k^{\mathrm{T}}$.

Wie bereits im Abschnitt 3.3 beim LQR-Regler mit stochastischer Störung gezeigt wurde, ist das Gütefunktional (3.74) für $N \to \infty$ nicht sinnvoll, weshalb dem stationären LQG-Problem (stationärer LQR-Regler und stationäres Kalman-Filter) das Gütekriterium

$$\lim_{N\to\infty} \frac{1}{N}\,\mathrm{E}\left(\sum_{k=0}^{N-1}\Big(\mathbf{x}_k^{\mathrm{T}}\mathbf{Q}\mathbf{x}_k + \mathbf{u}_k^{\mathrm{T}}\mathbf{R}\mathbf{u}_k + 2\mathbf{u}_k^{\mathrm{T}}\mathbf{N}\mathbf{x}_k\Big)\right) \tag{3.92}$$

zugrunde gelegt wird.

Man beachte, dass sich das stationäre Kalman-Filter und der stationäre Riccati-Regler gemeinsam

$$\hat{\mathbf{x}}_{k+1} = \mathbf{\Phi}\hat{\mathbf{x}}_k + \mathbf{\Gamma}\mathbf{u}_k + \hat{\mathbf{K}}(\mathbf{y}_k - \mathbf{C}\hat{\mathbf{x}}_k - \mathbf{D}\mathbf{u}_k) \qquad \hat{\mathbf{x}}(0) = \mathbf{0} \tag{3.93a}$$

$$\mathbf{u}_k = \mathbf{K}\hat{\mathbf{x}}_k \tag{3.93b}$$

auch als Reglerübertragungsmatrix $\mathbf{R}_{LQG}(z)$ mit dem q-dimensionalen Eingang $\mathbf{y}$ und dem p-dimensionalen Ausgang $\mathbf{u}$ in der Form

$$\mathbf{R}_{LQG}(z) = \frac{\mathbf{u}_z(z)}{\mathbf{y}_z(z)} = \mathbf{K}\Big(z\mathbf{E} - \Big(\mathbf{\Phi} + \mathbf{\Gamma}\mathbf{K} - \hat{\mathbf{K}}(\mathbf{C} + \mathbf{D}\mathbf{K})\Big)\Big)^{-1}\hat{\mathbf{K}} \tag{3.94}$$

schreiben lässt.

Aufgabe 3.6. Gegeben ist das vereinfachte Modell eines Musikkassetten-Antriebs nach Abbildung 3.2. Die beiden Gleichstrommotoren können unabhängig voneinander angesteuert werden, wodurch sowohl die Bandposition x_3 (Position des Lesekopfes) als auch die Zugkraft f_e getrennt geregelt werden können.

Das Massenträgheitsmoment der Motoren inklusive der Scheiben ist durch $J = 6.375 \cdot 10^{-3}\,\mathrm{kgm}^2$ gegeben, der Radius der Scheiben beträgt $r = 0.1\,\mathrm{m}$, die Ankerkonstante der Gleichstrommotoren hat den Wert $k_m = 0.544\,\mathrm{Nm/A}$ und das masselose Band kann näherungsweise durch eine lineare Feder mit der Federkonstanten $c = 2.113 \cdot 10^3\,\mathrm{N/m}$ und einen linearen Dämpfer mit der Dämpfungskonstanten

$d = 3.75\,\mathrm{Ns/m}$ modelliert werden.

Berechnen Sie das mathematische Modell mit den Eingangsgrößen $\mathbf{u}^{\mathrm{T}} = \begin{bmatrix} i_1 & i_2 \end{bmatrix}$ und den Ausgangsgrößen $\mathbf{y}^{\mathrm{T}} = \begin{bmatrix} x_3 & f_e \end{bmatrix}$. Wählen Sie eine geeignete Abtastzeit T_a und bestimmen Sie das zugehörige Abtastsystem. Entwerfen Sie ein Kalman-Filter sowie einen LQR-Zustandsregler mit Hilfe von MATLAB und simulieren Sie das Ergebnis in MATLAB/SIMULINK.

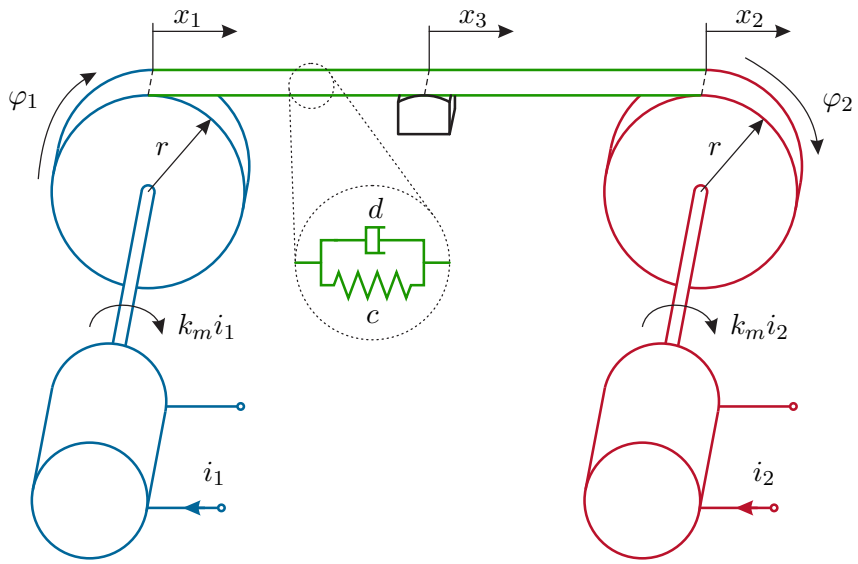


Abbildung 3.2: Zur Aufgabe 3.6 des Musikkassetten-Antriebs.

Hinweis: (zu Aufgabe 3.6) Das zeitkontinuierliche mathematische Modell folgt aus den Beziehungen

$$
\begin{aligned}
J\frac{\mathrm{d}^2}{\mathrm{d}t^2}\varphi_1 &= k_m i_1 + f_e r \\
J\frac{\mathrm{d}^2}{\mathrm{d}t^2}\varphi_2 &= k_m i_2 - f_e r \\
f_e &= c(x_2 - x_1) + d(\dot{x}_2 - \dot{x}_1) \\
x_3 &= \frac{x_1 + x_2}{2}\,.
\end{aligned}
$$

Wählen Sie als Zustandsgrößen die Winkel φ_1 und φ_2 sowie die zugehörigen Winkelgeschwindigkeiten ω_1 und ω_2.

Aufgabe 3.7. Abbildung 3.3 zeigt das vereinfachte Modell einer Magnetlagerung. Die Aufgabe der Regelung besteht darin, die Masse m durch die Magnetkraft f_{mag} in einem konstanten Abstand s (Geschwindigkeit $\dot{s} = w$) zu halten. Für die weitere Berechnung seien folgende Parameter gegeben: Masse $m = 500\,\mathrm{kg}$, Windungsanzahl der Spule $N = 500$, Widerstand der Spule $R = 2\Omega$, Permeabilitätskonstante der Luft $\mu_0 = 4\pi 10^{-7}\frac{\mathrm{Vs}}{\mathrm{Am}}$ sowie die Fläche des Luftspalts zur Berechnung der Magnetkraft $A = 0.04\,\mathrm{m}^2$.
Bestimmen Sie das nichtlineare mathematische Modell in der Form

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x} = \mathbf{f}(\mathbf{x}, u)$$
$$y = s$$

mit dem Zustand $\mathbf{x}^{\mathrm{T}} = \begin{bmatrix} s & w & i \end{bmatrix}$, der Eingangsspannung u als Eingangsgröße und der Position s als Ausgangsgröße. Linearisieren Sie das mathematische Modell um die Ruhelage $(\mathbf{x}_0, u_0)$, die durch die stationäre Spannung $u_0 = 60\,\mathrm{V}$ festgelegt ist. Entwerfen Sie einen zeitkontinuierlichen LQR-Zustandsregler und ein zeitkontinuierliches Kalman-Filter mit Hilfe von MATLAB und simulieren Sie das Ergebnis in MATLAB/SIMULINK. Verwenden Sie dazu die Befehle `ss`, `lqr`, `lqe`, `reg` und `lqgreg` sowie `lqg` der ROBUST CONTROL TOOLBOX.

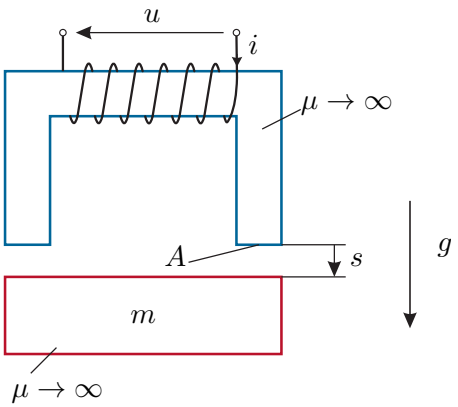


Abbildung 3.3: Zur Aufgabe 3.7 der Magnetlagerung.

Hinweis: (zu Aufgabe 3.7) Das nichtlineare mathematische Modell lautet

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} s \\ w \\ i \end{bmatrix} = \begin{bmatrix} w \\ g - \frac{ki^2}{2ms^2} \\ \frac{s}{k}\left(u - Ri + \frac{iwk}{s^2}\right) \end{bmatrix} \quad \text{mit} \quad k = \frac{1}{2}A\mu_0 N^2 \ .$$

Aufgabe 3.8. Betrachten Sie die Temperaturregelstrecke nach Abbildung 3.4 mit einem konstanten Massenstrom $\dot{m}_{zu} = \dot{m}_{ab} = \dot{m}$.

Die Temperatur im Tank T_{tank} mit der konstanten Wassermenge m_{tank} kann über das Mischventil mit der Ausgangstemperatur T_m beeinflusst werden. Für die Temperatur direkt am Tankeingang T_{zu} gilt wegen der Vernachlässigung der Wärmeverluste der Rohrleitung

$$T_{zu} = T_m(t - T_t)$$

mit der durch den Transportprozess durch die Rohrleitung bedingten Totzeit T_t.

Bestimmen Sie das mathematische Modell der Temperaturregelstrecke mit der Zustandsgröße $x = T_{tank}$, der Eingangsgröße $u = T_m$ und der Ausgangsgröße $y = T_{ab} = T_{tank}$. Berechnen Sie weiters allgemein die Transporttotzeit T_t für eine reibungsfreie Rohrleitung mit dem Innendurchmesser D und der Länge L bei gegebenen Massenstrom $\dot{m}$ und Dichte ρ der Flüssigkeit.

Die in der Wassermenge m_{tank} mit der Temperatur T_{tank} und der spezifischen Wärmekapazität von Wasser c_W gespeicherte Energie E_{tank} lautet

$$E_{tank} = c_W m_{tank} T_{tank} \ .$$

Wählen Sie geeignete Parameter und implementieren Sie das mathematische Modell in MATLAB/SIMULINK. Entwerfen Sie für eine geeignete Abtastzeit T_a ein Kalman-Filter sowie einen LQR-Zustandsregler und simulieren Sie den geschlossenen Regelkreis mit der zeitkontinuierlichen Regelstrecke in MATLAB/SIMULINK.

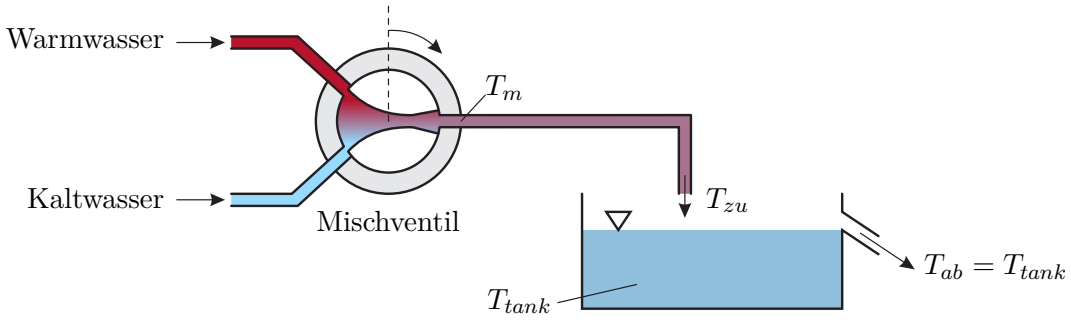


Abbildung 3.4: Zur Aufgabe 3.8 der Temperaturregelstrecke.

3.5 Erweiterte Konzepte der Zustandsregelung

3.5.1 Feedforward der geschätzten Störung

Bereits beim Kalman-Filter (siehe Abschnitt 2.3.2) wurde dargestellt, wie man deterministische Störungen durch Hinzunahme eines Störmodells systematisch berücksichtigen kann. Im Folgenden wird dieses Konzept auf den kombinierten Zustandsregler- und Zustandsbeobachterentwurf erweitert. Dazu betrachte man das lineare, zeitinvariante, zeitkontinuierliche System

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} + \mathbf{G}\mathbf{w} \qquad \mathbf{x}(0) = \mathbf{x}_0 \tag{3.95a}$$

$$\mathbf{y} = \mathbf{C}\mathbf{x} \tag{3.95b}$$

mit dem n-dimensionalen Zustand $\mathbf{x} \in \mathbb{R}^n$, dem p-dimensionalen deterministischen Eingang $\mathbf{u} \in \mathbb{R}^p$, dem q-dimensionalen Ausgang $\mathbf{y} \in \mathbb{R}^q$, der r-dimensionalen deterministischen Störung $\mathbf{w} \in \mathbb{R}^r$ sowie den Matrizen $\mathbf{A} \in \mathbb{R}^{n\times n}$, $\mathbf{B} \in \mathbb{R}^{n\times p}$, $\mathbf{G} \in \mathbb{R}^{n\times r}$ und $\mathbf{C} \in \mathbb{R}^{q\times n}$. Es sei angenommen, dass die Störung $\mathbf{w}$ durch das Störmodell

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{z} = \mathbf{A}_\mathbf{z}\mathbf{z} \qquad \mathbf{z}(0) = \mathbf{z}_0 \tag{3.96a}$$

$$\mathbf{w} = \mathbf{C}_\mathbf{z}\mathbf{z} \tag{3.96b}$$

beschrieben werden kann. In einem ersten Schritt wird nun für das erweiterte System

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix}\mathbf{x}\\ \mathbf{z}\end{bmatrix} = \begin{bmatrix}\mathbf{A} & \mathbf{G}\mathbf{C}_\mathbf{z}\\ \mathbf{0} & \mathbf{A}_\mathbf{z}\end{bmatrix}\begin{bmatrix}\mathbf{x}\\ \mathbf{z}\end{bmatrix} + \begin{bmatrix}\mathbf{B}\\ \mathbf{0}\end{bmatrix}\mathbf{u} \tag{3.97a}$$

$$\mathbf{y} = \begin{bmatrix}\mathbf{C} & \mathbf{0}\end{bmatrix}\begin{bmatrix}\mathbf{x}\\ \mathbf{z}\end{bmatrix} \tag{3.97b}$$

das zugehörige Abtastsystem für die Abtastzeit T_a in der Form

$$\begin{bmatrix}\mathbf{x}_{k+1}\\ \mathbf{z}_{k+1}\end{bmatrix} = \begin{bmatrix}\mathbf{\Phi} & \mathbf{\Phi}_{\mathbf{xz}}\\ \mathbf{0} & \mathbf{\Phi}_\mathbf{z}\end{bmatrix}\begin{bmatrix}\mathbf{x}_k\\ \mathbf{z}_k\end{bmatrix} + \begin{bmatrix}\mathbf{\Gamma}\\ \mathbf{0}\end{bmatrix}\mathbf{u}_k \tag{3.98a}$$

$$\mathbf{y}_k = \begin{bmatrix}\mathbf{C} & \mathbf{0}\end{bmatrix}\begin{bmatrix}\mathbf{x}_k\\ \mathbf{z}_k\end{bmatrix} \tag{3.98b}$$

berechnet.

> *Aufgabe* 3.9. Zeigen Sie, dass das Abtastsystem von (3.97) die Struktur von (3.98) haben muss.

Wie man sieht, ist der Zustand des Störmodells $\mathbf{z}$ über den Eingang $\mathbf{u}$ nicht erreichbar. Wenn das System (3.98) vollständig beobachtbar ist, kann der Zustand $\mathbf{x}$ sowie der Zustand des Störmodells $\mathbf{z}$ über einen Beobachter

$$\begin{bmatrix}\hat{\mathbf{x}}_{k+1}\\ \hat{\mathbf{z}}_{k+1}\end{bmatrix} = \begin{bmatrix}\mathbf{\Phi} & \mathbf{\Phi}_{\mathbf{xz}}\\ \mathbf{0} & \mathbf{\Phi}_\mathbf{z}\end{bmatrix}\begin{bmatrix}\hat{\mathbf{x}}_k\\ \hat{\mathbf{z}}_k\end{bmatrix} + \begin{bmatrix}\mathbf{\Gamma}\\ \mathbf{0}\end{bmatrix}\mathbf{u}_k + \begin{bmatrix}\hat{\mathbf{K}}\\ \hat{\mathbf{K}}_\mathbf{z}\end{bmatrix}(\mathbf{y}_k - \mathbf{C}\hat{\mathbf{x}}_k) \tag{3.99}$$

beobachtet werden. Im nächsten Schritt wird ein Zustandsregler

$$\mathbf{u}_k = \mathbf{K}\mathbf{x}_k \tag{3.100}$$

für das System

$$\mathbf{x}_{k+1} = \mathbf{\Phi}\mathbf{x}_k + \mathbf{\Gamma}\mathbf{u}_k \tag{3.101}$$

entworfen und in der Form

$$\mathbf{u}_k = \mathbf{K}\hat{\mathbf{x}}_k + \mathbf{K}_\mathbf{z}\hat{\mathbf{z}}_k \tag{3.102}$$

erweitert und implementiert. Damit lautet der geschlossene Regelkreis (3.98), (3.99) und (3.102)

$$\mathbf{x}_{k+1} = (\mathbf{\Phi} + \mathbf{\Gamma}\mathbf{K})\mathbf{x}_k + (\mathbf{\Phi}_{\mathbf{xz}} + \mathbf{\Gamma}\mathbf{K}_\mathbf{z})\mathbf{z}_k - \mathbf{\Gamma}\mathbf{K}\tilde{\mathbf{x}}_k - \mathbf{\Gamma}\mathbf{K}_\mathbf{z}\tilde{\mathbf{z}}_k \tag{3.103a}$$

$$\mathbf{z}_{k+1} = \mathbf{\Phi}_\mathbf{z}\mathbf{z}_k \tag{3.103b}$$

mit der Dynamik der Beobachtungsfehler $\tilde{\mathbf{x}}_k = \mathbf{x}_k - \hat{\mathbf{x}}_k$ und $\tilde{\mathbf{z}}_k = \mathbf{z}_k - \hat{\mathbf{z}}_k$

$$\begin{bmatrix}\tilde{\mathbf{x}}_{k+1}\\ \tilde{\mathbf{z}}_{k+1}\end{bmatrix} = \begin{bmatrix}\mathbf{\Phi} - \hat{\mathbf{K}}\mathbf{C} & \mathbf{\Phi}_{\mathbf{xz}}\\ -\hat{\mathbf{K}}_\mathbf{z}\mathbf{C} & \mathbf{\Phi}_\mathbf{z}\end{bmatrix}\begin{bmatrix}\tilde{\mathbf{x}}_k\\ \tilde{\mathbf{z}}_k\end{bmatrix}. \tag{3.104}$$

Man beachte, dass mit der Reglermatrix $\mathbf{K}$ die Dynamik festgelegt wird, mit der ein Anfangsfehler $\mathbf{x}(0)$ bei verschwindender Störung, also $\mathbf{w} = \mathbf{0}$, zu Null geregelt wird. Im Weiteren legen die Beobachtermatrizen $\hat{\mathbf{K}}$ und $\hat{\mathbf{K}}_\mathbf{z}$ die Fehlerdynamik fest und mit Hilfe von $\mathbf{K}_\mathbf{z}$ kann der Einfluss der Störung $\mathbf{w}$ gezielt unterdrückt werden. Wie man aus (3.103) erkennt, ist die optimale Wahl für $\mathbf{K}_\mathbf{z}$, falls möglich, durch

$$\mathbf{\Phi}_{\mathbf{xz}} + \mathbf{\Gamma}\mathbf{K}_\mathbf{z} = \mathbf{0} \tag{3.105}$$

gegeben. Man nennt diese Strategie zur Störunterdrückung auch *Feedforward der geschätzten Störung*. Abbildung 3.5 zeigt das zugehörige Blockschaltbild.

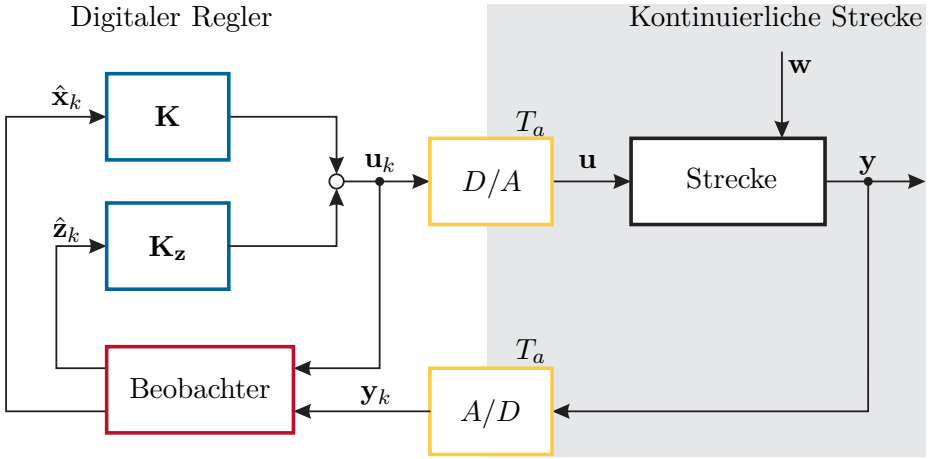


Abbildung 3.5: Blockschaltbild zum Feedforward der geschätzten Störung.

3.5.2 Zustandsregler- und Zustandsbeobachterentwurf mit Integralanteil

Es sei nun angenommen, dass die Störung $\mathbf{w}$ in (3.95) konstant aber unbekannt ist. Dann lautet das zugehörige Störmodell (3.96)

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{z} = \mathbf{0} \qquad \mathbf{z}(0) = \mathbf{w} \tag{3.106a}$$

$$\mathbf{w} = \mathbf{z} \tag{3.106b}$$

und das Abtastsystem (3.98) ergibt sich zu

$$\begin{bmatrix}\mathbf{x}_{k+1}\\ \mathbf{z}_{k+1}\end{bmatrix} = \begin{bmatrix}\boldsymbol{\Phi} & \boldsymbol{\Phi}_{\mathbf{xz}}\\ \mathbf{0} & \mathbf{E}\end{bmatrix}\begin{bmatrix}\mathbf{x}_k\\ \mathbf{z}_k\end{bmatrix} + \begin{bmatrix}\boldsymbol{\Gamma}\\ \mathbf{0}\end{bmatrix}\mathbf{u}_k \tag{3.107a}$$

$$\mathbf{y}_k = \begin{bmatrix}\mathbf{C} & \mathbf{0}\end{bmatrix}\begin{bmatrix}\mathbf{x}_k\\ \mathbf{z}_k\end{bmatrix} \tag{3.107b}$$

mit

$$\boldsymbol{\Phi} = \exp(\mathbf{A}T_a) \tag{3.108a}$$

$$\boldsymbol{\Phi}_{\mathbf{xz}} = \int_0^{T_a} \exp(\mathbf{A}\tau)\,\mathrm{d}\tau\mathbf{G} \tag{3.108b}$$

$$\boldsymbol{\Gamma} = \int_0^{T_a} \exp(\mathbf{A}\tau)\,\mathrm{d}\tau\mathbf{B}\ . \tag{3.108c}$$

Aufgabe 3.10. Zeigen Sie die Gültigkeit von (3.108).

Nimmt man nun an, dass $\boldsymbol{\Phi}_{\mathbf{xz}} = \boldsymbol{\Gamma}$ bzw. $\mathbf{G} = \mathbf{B}$ gilt, d. h. die Störung $\mathbf{z}_k$ wirkt beim Abtastsystem vor dem Eingang $\mathbf{u}_k$ auf das System, dann lässt sich (3.105) exakt lösen. Es gilt nämlich $\mathbf{K}_\mathbf{z} = -\mathbf{E}$. Damit lautet der Regler von (3.102)

$$\mathbf{u}_k = \mathbf{K}\hat{\mathbf{x}}_k - \hat{\mathbf{z}}_k \tag{3.109}$$

und der Beobachter von (3.99) hat die Form

$$\begin{bmatrix}\hat{\mathbf{x}}_{k+1}\\ \hat{\mathbf{z}}_{k+1}\end{bmatrix} = \begin{bmatrix}\boldsymbol{\Phi} & \boldsymbol{\Gamma}\\ \mathbf{0} & \mathbf{E}\end{bmatrix}\begin{bmatrix}\hat{\mathbf{x}}_k\\ \hat{\mathbf{z}}_k\end{bmatrix} + \begin{bmatrix}\boldsymbol{\Gamma}\\ \mathbf{0}\end{bmatrix}\mathbf{u}_k + \begin{bmatrix}\hat{\mathbf{K}}\\ \hat{\mathbf{K}}_\mathbf{z}\end{bmatrix}(\mathbf{y}_k - \mathbf{C}\hat{\mathbf{x}}_k)\ . \tag{3.110}$$

Man erkennt durch Umschreiben von (3.109) und (3.110)

$$\hat{\mathbf{x}}_{k+1} = (\boldsymbol{\Phi} + \boldsymbol{\Gamma}\mathbf{K})\hat{\mathbf{x}}_k + \hat{\mathbf{K}}(\mathbf{y}_k - \mathbf{C}\hat{\mathbf{x}}_k) \tag{3.111a}$$

$$\hat{\mathbf{z}}_{k+1} = \hat{\mathbf{z}}_k + \hat{\mathbf{K}}_\mathbf{z}(\mathbf{y}_k - \mathbf{C}\hat{\mathbf{x}}_k) \tag{3.111b}$$

$$\mathbf{u}_k = \mathbf{K}\hat{\mathbf{x}}_k - \hat{\mathbf{z}}_k, \tag{3.111c}$$

dass sich der Regler aus der Rückführung des geschätzten Zustandes $\hat{\mathbf{x}}_k$ mit der Reglermatrix $\mathbf{K}$ und dem durch die Matrix $\hat{\mathbf{K}}_\mathbf{z}$ gewichteten integrierten Ausgangsfehler zusammensetzt. Abbildung 3.6 zeigt die zugehörige Reglerstruktur in Form eines Blockschaltbildes.

Zusammenfassend sieht der Entwurf mit Integralanteil folgendermaßen aus:

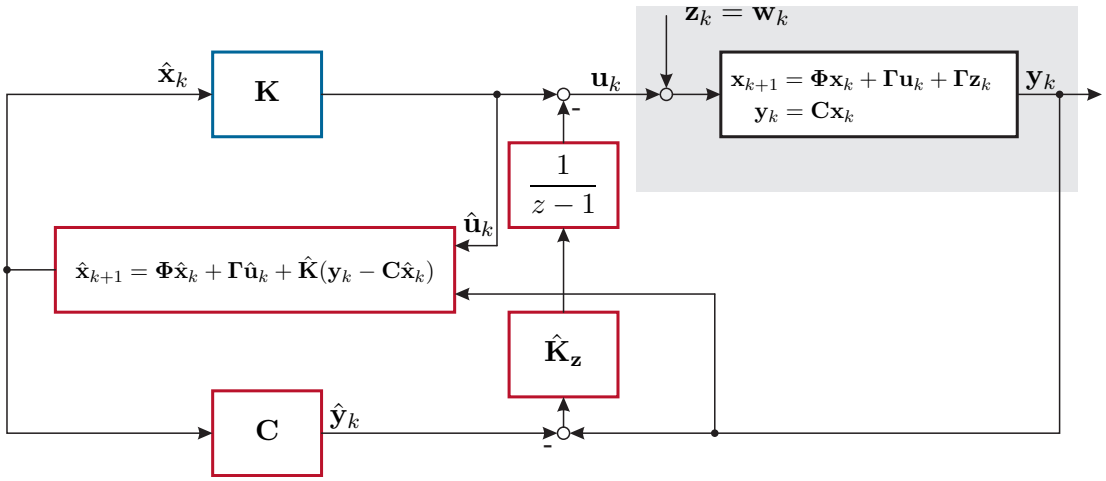


Abbildung 3.6: Zum Zustandsregler-/Zustandsbeobachterentwurf mit Integralanteil.

(1) Man entwirft für das System

$$\mathbf{x}_{k+1} = \mathbf{\Phi}\mathbf{x}_k + \mathbf{\Gamma}\mathbf{u}_k \tag{3.112a}$$

$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k \tag{3.112b}$$

einen Zustandsregler der Form

$$\mathbf{u}_k = \mathbf{K}\mathbf{x}_k \ . \tag{3.113}$$

(2) Im zweiten Schritt erweitert man das System (3.112) um eine konstante, aber unbekannte Störung $\mathbf{w}$, die am Eingang der Strecke wirkt, also

$$\begin{bmatrix}\mathbf{x}_{k+1}\\ \mathbf{z}_{k+1}\end{bmatrix} = \begin{bmatrix}\mathbf{\Phi} & \mathbf{\Gamma}\\ \mathbf{0} & \mathbf{E}\end{bmatrix}\begin{bmatrix}\mathbf{x}_k\\ \mathbf{z}_k\end{bmatrix} + \begin{bmatrix}\mathbf{\Gamma}\\ \mathbf{0}\end{bmatrix}\mathbf{u}_k \qquad \mathbf{z}(0) = \mathbf{w} \tag{3.114a}$$

$$\mathbf{y}_k = \begin{bmatrix}\mathbf{C} & \mathbf{0}\end{bmatrix}\begin{bmatrix}\mathbf{x}_k\\ \mathbf{z}_k\end{bmatrix} \tag{3.114b}$$

und entwirft dafür die Beobachterverstärkungsmatrizen $\hat{\mathbf{K}}$ und $\hat{\mathbf{K}}_\mathbf{z}$ eines vollständigen Beobachters (siehe (3.110)).

(3) Der Regler folgt dann gemäß (3.109) zu

$$\mathbf{u}_k = \mathbf{K}\hat{\mathbf{x}}_k - \hat{\mathbf{z}}_k \ . \tag{3.115}$$

3.5.3 Zustandsregler- und Zustandsbeobachterentwurf mit Führungsgrößen

In diesem Abschnitt wird gezeigt, wie man beim Zustandsregler- und Zustandsbeobachterentwurf systematisch Führungsgrößen mitberücksichtigen kann. In der englischsprachigen Literatur wird diese Aufgabenstellung oft auch als das *Servoproblem* bezeichnet. Es wird angenommen, dass die Größen $\bar{\mathbf{y}}_k \in \mathbb{R}^p$

$$\bar{\mathbf{y}}_k = \mathbf{C}_\mathbf{r}\mathbf{x}_k \tag{3.116}$$

des Systems

$$\mathbf{x}_{k+1} = \mathbf{\Phi}\mathbf{x}_k + \mathbf{\Gamma}\mathbf{u}_k \tag{3.117a}$$
$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k, \tag{3.117b}$$

mit dem n-dimensionalen Zustand $\mathbf{x} \in \mathbb{R}^n$, dem p-dimensionalen Eingang $\mathbf{u} \in \mathbb{R}^p$, dem q-dimensionalen Ausgang $\mathbf{y} \in \mathbb{R}^q$ sowie den Matrizen $\mathbf{\Phi} \in \mathbb{R}^{n\times n}$, $\mathbf{\Gamma} \in \mathbb{R}^{n\times p}$ und $\mathbf{C} \in \mathbb{R}^{q\times n}$, auf einen vorgegebenen stationären Referenzwert $\mathbf{r}_s \in \mathbb{R}^p$

$$\bar{\mathbf{y}}_s = \mathbf{C}_\mathbf{r}\mathbf{x}_s = \mathbf{r}_s \tag{3.118}$$

geregelt werden sollen. In einem ersten Schritt wird ein Zustandsregler

$$\mathbf{u}_k = \mathbf{K}\mathbf{x}_k \tag{3.119}$$

entworfen, beispielsweise als stationärer Riccati-Regler, und in einem zweiten Schritt in der Form

$$\mathbf{u}_k = -\mathbf{K}(\mathbf{L}_\mathbf{r}\mathbf{r}_k - \mathbf{x}_k) + \mathbf{L}_\mathbf{u}\mathbf{r}_k \tag{3.120}$$

mit $\mathbf{L}_\mathbf{r} \in \mathbb{R}^{n\times p}$ und $\mathbf{L}_\mathbf{u} \in \mathbb{R}^{p\times p}$ erweitert. Die Matrizen $\mathbf{L}_\mathbf{r}$ und $\mathbf{L}_\mathbf{u}$ sollen dabei nachfolgende Bedingungen für den stationären Zustand

$$\mathbf{L}_\mathbf{r}\mathbf{r}_s = \mathbf{x}_s \tag{3.121a}$$
$$\mathbf{L}_\mathbf{u}\mathbf{r}_s = \mathbf{u}_s \tag{3.121b}$$

erfüllen. Aus (3.117) folgt im stationären Zustand

$$(\mathbf{E} - \mathbf{\Phi})\mathbf{x}_s - \mathbf{\Gamma}\mathbf{u}_s = \mathbf{0} \tag{3.122}$$

und durch Einsetzen von (3.121) in (3.118) und (3.122) erhält man

$$((\mathbf{E} - \mathbf{\Phi})\mathbf{L}_\mathbf{r} - \mathbf{\Gamma}\mathbf{L}_\mathbf{u})\mathbf{r}_s = \mathbf{0} \tag{3.123a}$$
$$\bar{\mathbf{y}}_s = \mathbf{C}_\mathbf{r}\mathbf{L}_\mathbf{r}\mathbf{r}_s = \mathbf{r}_s \tag{3.123b}$$

bzw. für $\mathbf{r}_s \neq \mathbf{0}$

$$\underbrace{\begin{bmatrix} \mathbf{E} - \mathbf{\Phi} & -\mathbf{\Gamma} \\ \mathbf{C}_\mathbf{r} & \mathbf{0} \end{bmatrix}}_{\mathbf{X}} \begin{bmatrix} \mathbf{L}_\mathbf{r} \\ \mathbf{L}_\mathbf{u} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{E} \end{bmatrix}. \tag{3.124}$$

Wenn die Matrix $\mathbf{X}$ invertierbar ist, dann können aus (3.124) die Matrizen $\mathbf{L}_\mathbf{r}$ und $\mathbf{L}_\mathbf{u}$ berechnet werden. In Abbildung 3.7 ist die zugehörige Reglerstruktur dargestellt.

Man erkennt, dass (3.120) auch in der Form

$$\mathbf{u}_k = \mathbf{K}\mathbf{x}_k + \mathbf{L}\mathbf{r}_k \quad \text{mit} \quad \mathbf{L} = \mathbf{L}_\mathbf{u} - \mathbf{K}\mathbf{L}_\mathbf{r} \tag{3.125}$$

vereinfacht werden kann.

Nimmt man nun an, dass der Zustand $\mathbf{x}_k$ nicht messbar ist, dann wird zusätzlich ein Zustandsbeobachter der Form

$$\hat{\mathbf{x}}_{k+1} = \mathbf{\Phi}\hat{\mathbf{x}}_k + \mathbf{\Gamma}\mathbf{u}_k + \hat{\mathbf{K}}(\mathbf{y}_k - \mathbf{C}\hat{\mathbf{x}}_k) \tag{3.126}$$

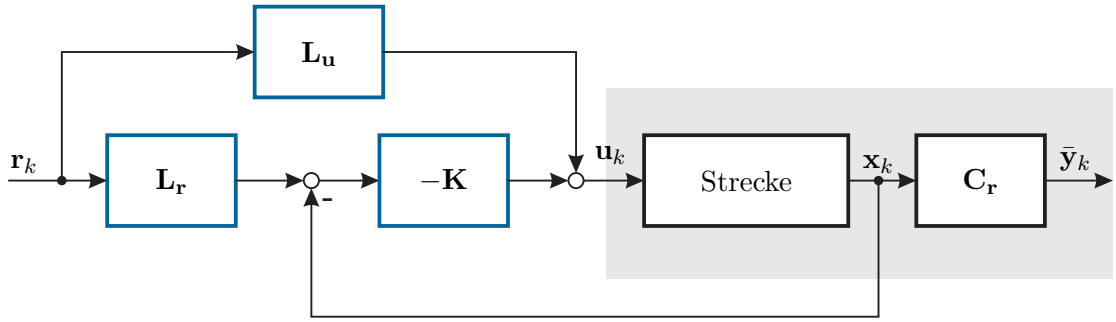


Abbildung 3.7: Reglerstruktur bei Zustandsregelung mit Führungsgrößen.

benötigt. Der geschlossene Kreis (3.117), (3.125) und (3.126) lautet dann mit dem Beobachtungsfehler $\mathbf{e}_k = \mathbf{x}_k - \hat{\mathbf{x}}_k$

$$\mathbf{x}_{k+1} = (\mathbf{\Phi} + \mathbf{\Gamma K})\mathbf{x}_k - \mathbf{\Gamma K}\mathbf{e}_k + \mathbf{\Gamma L}\mathbf{r}_k \tag{3.127a}$$

$$\mathbf{e}_{k+1} = \left(\mathbf{\Phi} - \hat{\mathbf{K}}\mathbf{C}\right)\mathbf{e}_k \tag{3.127b}$$

$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k \ . \tag{3.127c}$$

Wie zu erwarten, ist der Beobachtungsfehler $\mathbf{e}_k$ über den Referenzeingang $\mathbf{r}_k$ nicht erreichbar.

Aufgabe 3.11. Auf analoge Art und Weise wie in Abschnitt 3.5.2 gezeigt, kann man den Zustandsregler- und Zustandsbeobachterentwurf mit Führungsgrößen um einen Integralanteil erweitern. Zeigen Sie, dass sich in diesem Fall das Regelgesetz wie folgt

$$\begin{aligned}
\hat{\mathbf{x}}_{k+1} &= (\mathbf{\Phi} + \mathbf{\Gamma K})\hat{\mathbf{x}}_k + \hat{\mathbf{K}}(\mathbf{y}_k - \mathbf{C}\hat{\mathbf{x}}_k) + \mathbf{\Gamma L}\mathbf{r}_k \\
\hat{\mathbf{z}}_{k+1} &= \hat{\mathbf{z}}_k + \hat{\mathbf{K}}_{\mathbf{z}}(\mathbf{y}_k - \mathbf{C}\hat{\mathbf{x}}_k) \\
\mathbf{u}_k &= \mathbf{K}\hat{\mathbf{x}}_k - \hat{\mathbf{z}}_k + \mathbf{L}\mathbf{r}_k
\end{aligned}$$

errechnet.

3.5.4 Das Feedforward-Konzept

Wenn sich das geregelte System wie ein Referenzmodell der Form

$$\bar{\mathbf{x}}_{k+1} = \mathbf{\Phi}_m\bar{\mathbf{x}}_k + \mathbf{\Gamma}_m\mathbf{r}_k \tag{3.128a}$$

$$\bar{\mathbf{y}}_k = \mathbf{C}_m\bar{\mathbf{x}}_k \tag{3.128b}$$

verhalten soll, dann kann man folgende Vorgangsweise wählen: Man simuliert das Referenzmodell (3.128) im Computer mit und setzt das Regelgesetz wie folgt

$$\mathbf{u}_k = -\mathbf{K}(\bar{\mathbf{x}}_k - \mathbf{x}_k) + \bar{\mathbf{u}}_k \tag{3.129}$$

an.

Das so genannte *(Stellgrößen-)Feedforward-Signal* $\bar{\mathbf{u}}_k$ wird dabei so bestimmt, dass bei idealer Übereinstimmung von Referenzmodell und geregeltem System, also $\bar{\mathbf{x}}_k = \mathbf{x}_k$, auch die Ausgangsgrößen des Referenzmodells und des geregelten Systems übereinstimmen, also $\bar{\mathbf{y}}_k = \mathbf{y}_k$.

Die Berechnung des Feedforward-Signals im Mehrgrößenfall gestaltet sich im Allgemeinen relativ schwierig. Im Eingrößenfall kann mit Hilfe der z-Übertragungsfunktionen von (3.117) und (3.128)

$$G(z) = \frac{y_z(z)}{u_z(z)} = \mathbf{c}^{\mathrm{T}}(z\mathbf{E} - \mathbf{\Phi})^{-1}\mathbf{\Gamma} \tag{3.130a}$$

$$G_m(z) = \frac{\bar{y}_z(z)}{r_z(z)} = \mathbf{c}_m^{\mathrm{T}}(z\mathbf{E} - \mathbf{\Phi}_m)^{-1}\mathbf{\Gamma}_m \tag{3.130b}$$

aus der Bedingung

$$G(z)u_z(z) = y_z(z) = \bar{y}_z(z) = G_m(z)r_z(z) \tag{3.131}$$

die z-Transformierte $\bar{u}_z(z)$ des Feedforward-Signals $(\bar{u}_k)$ wie folgt

$$\bar{u}_z(z) = \frac{G_m(z)}{G(z)} r_z(z) \tag{3.132}$$

errechnet werden. Man erkennt aus (3.132), dass dies nur dann möglich ist, wenn die Graddifferenz von $G_m(z)$ größer gleich der Graddifferenz von $G(z)$ ist, $G_m(z)$ BIBO-stabil ist und sämtliche Nullstellen im geschlossenen Äußeren des Einheitskreises von $G(z)$ auch Nullstellen von $G_m(z)$ sind.

Im Fall, dass die Zählerpolynome von $G(z)$ und $G_m(z)$ übereinstimmen und die Ordnungen der Nennerpolynome gleich sind, also

$$G(z) = \frac{z_G(z)}{n_G(z)} = \frac{b_0 + b_1 z + \cdots + b_{n-1}z^{n-1} + b_n z^n}{a_0 + a_1 z + \cdots + a_{n-1}z^{n-1} + a_n z^n} \tag{3.133a}$$

$$G_m(z) = \frac{z_{G_m}(z)}{n_{G_m}(z)} = V\frac{b_0 + b_1 z + \cdots + b_{n-1}z^{n-1} + b_n z^n}{\bar{a}_0 + \bar{a}_1 z + \cdots + \bar{a}_{n-1}z^{n-1} + z^n}, \tag{3.133b}$$

folgt $\bar{u}_z(z)$ einfach zu

$$\bar{u}_z(z) = V\frac{a_0 + a_1 z + \cdots + a_{n-1}z^{n-1} + a_n z^n}{\bar{a}_0 + \bar{a}_1 z + \cdots + \bar{a}_{n-1}z^{n-1} + z^n} r_z(z)\ . \tag{3.134}$$

Der Faktor V wird beispielsweise so gewählt, dass gilt $\lim_{z\to 1} G_m(z) = 1$. Liegt nun das Referenzmodell (3.128) in der ersten Standardform vor, dann lautet die Zustandsrealisierung

von (3.134) in der ersten Standardform im Eingrößenfall

$$\underbrace{\begin{bmatrix} \bar{x}_{1,k+1} \\ \bar{x}_{2,k+1} \\ \vdots \\ \bar{x}_{n-1,k+1} \\ \bar{x}_{n,k+1} \end{bmatrix}}_{\bar{\mathbf{x}}_{k+1}} = \underbrace{\begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 1 \\ -\bar{a}_0 & -\bar{a}_1 & \dots & -\bar{a}_{n-2} & -\bar{a}_{n-1} \end{bmatrix}}_{\mathbf{\Phi}_m} \underbrace{\begin{bmatrix} \bar{x}_{1,k} \\ \bar{x}_{2,k} \\ \vdots \\ \bar{x}_{n-1,k} \\ \bar{x}_{n,k} \end{bmatrix}}_{\bar{\mathbf{x}}_k} + \underbrace{\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}}_{\mathbf{\Gamma}_m} r_k \tag{3.135a}$$

$$\bar{u}_k = \underbrace{V \begin{bmatrix} a_0 - \bar{a}_0 a_n & a_1 - \bar{a}_1 a_n & \dots & a_{n-1} - \bar{a}_{n-1} a_n \end{bmatrix}}_{\bar{\mathbf{c}}_m^T} \underbrace{\begin{bmatrix} \bar{x}_{1,k} \\ \bar{x}_{2,k} \\ \vdots \\ \bar{x}_{n-1,k} \\ \bar{x}_{n,k} \end{bmatrix}}_{\bar{\mathbf{x}}_k} + V a_n r_k \ . \tag{3.135b}$$

Nun kann man sämtliche bisher gewonnenen Ergebnisse in einer gemeinsamen Struktur zusammenfassen. Ein Zustandsregler- und Zustandsbeobachter mit Integralanteil und Feedforward über ein Referenzmodell im Falle eines Eingrößensystems setzt sich aus dem *Referenzmodell* (siehe (3.128))

$$\bar{\mathbf{x}}_{k+1} = \mathbf{\Phi}_m \bar{\mathbf{x}}_k + \mathbf{\Gamma}_m r_k \tag{3.136a}$$

$$\bar{y}_k = \mathbf{c}_m^{\mathrm{T}} \bar{\mathbf{x}}_k \tag{3.136b}$$

aus dem *Zustands- und Störgrößenbeobachter* (siehe (3.110))

$$\begin{bmatrix} \hat{\mathbf{x}}_{k+1} \\ \hat{z}_{k+1} \end{bmatrix} = \begin{bmatrix} \mathbf{\Phi} & \mathbf{\Gamma} \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{x}}_k \\ \hat{z}_k \end{bmatrix} + \begin{bmatrix} \mathbf{\Gamma} \\ 0 \end{bmatrix} u_k + \begin{bmatrix} \hat{\mathbf{k}} \\ \hat{k}_z \end{bmatrix} \Big(y_k - \mathbf{c}^{\mathrm{T}} \hat{\mathbf{x}}_k \Big) \tag{3.137}$$

und aus dem *Stellgrößengesetz*, bestehend aus dem *Rückkopplungsanteil* $-\mathbf{k}^{\mathrm{T}}(\bar{\mathbf{x}}_k - \hat{\mathbf{x}}_k) - \hat{z}_k$ und dem *Feedforward-Anteil* $\bar{u}_k = \bar{\mathbf{c}}_m^{\mathrm{T}} \bar{\mathbf{x}}_k + V a_n r_k$ (siehe (3.119), (3.129) und (3.135)),

$$u_k = -\mathbf{k}^{\mathrm{T}}(\bar{\mathbf{x}}_k - \hat{\mathbf{x}}_k) - \hat{z}_k + \underbrace{\bar{\mathbf{c}}_m^{\mathrm{T}} \bar{\mathbf{x}}_k + V a_n r_k}_{\bar{u}_k} \tag{3.138}$$

zusammen. Durch Vereinfachung der Gleichungen (3.136)–(3.138) erhält man schlussendlich

$$\bar{\mathbf{x}}_{k+1} = \mathbf{\Phi}_m \bar{\mathbf{x}}_k + \mathbf{\Gamma}_m r_k \tag{3.139a}$$

$$\hat{\mathbf{x}}_{k+1} = \Big(\mathbf{\Phi} + \mathbf{\Gamma} \mathbf{k}^{\mathrm{T}} - \hat{\mathbf{k}} \mathbf{c}^{\mathrm{T}} \Big) \hat{\mathbf{x}}_k + \mathbf{\Gamma} \Big(\bar{\mathbf{c}}_m^{\mathrm{T}} - \mathbf{k}^{\mathrm{T}} \Big) \bar{\mathbf{x}}_k + V a_n \mathbf{\Gamma} r_k + \hat{\mathbf{k}} y_k \tag{3.139b}$$

$$\hat{z}_{k+1} = \hat{z}_k + \hat{k}_z \Big(y_k - \mathbf{c}^{\mathrm{T}} \hat{\mathbf{x}}_k \Big) \tag{3.139c}$$

$$u_k = \Big(\bar{\mathbf{c}}_m^{\mathrm{T}} - \mathbf{k}^{\mathrm{T}} \Big) \bar{\mathbf{x}}_k + \mathbf{k}^{\mathrm{T}} \hat{\mathbf{x}}_k - \hat{z}_k + V a_n r_k \ . \tag{3.139d}$$

Die Abbildung 3.8 zeigt die Struktur des Regelkonzeptes (3.139).

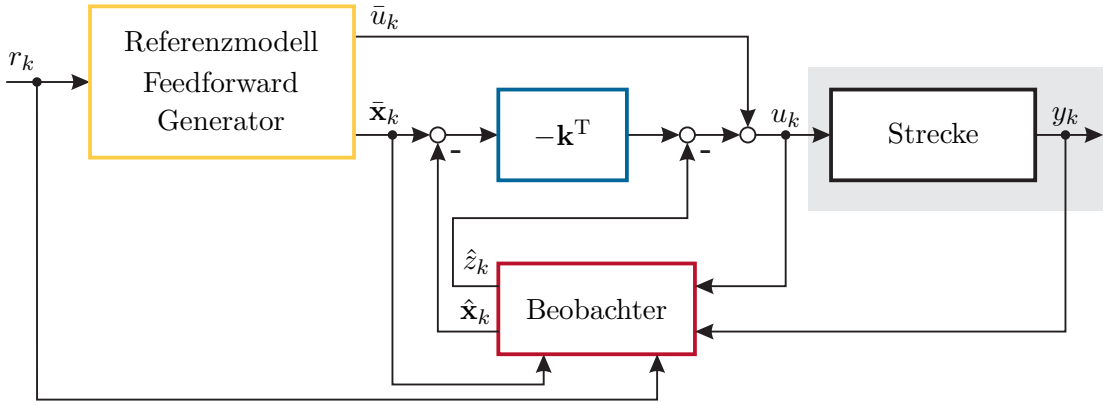


Abbildung 3.8: Zur Regelstruktur des Feedforward-Konzeptes.

Aufgabe 3.12. Betrachten Sie die Übertragungsfunktion eines Doppelintegrators

$$G(s) = \frac{1}{s^2} \ .$$

Bestimmen Sie für die Abtastzeit $T_a = 0.1\,\mathrm{s}$ das zugehörige Abtastmodell. Wenden Sie für dieses Abtastmodell, sofern möglich, sämtliche Regelkonzepte, die in diesem Kapitel vorgestellt wurden, an. Berücksichtigen Sie dabei einmal eine konstante, aber unbekannte Störung am Eingang und einmal eine sinusförmige Eingangsstörung der Form

$$w(t) = A\sin(2t + \varphi)$$

mit unbekannter Amplitude A und unbekannter Phase φ. Im Weiteren sollen sprungförmige und sinusförmige Referenzsignale vorgegeben werden, denen das System zumindest stationär ohne Regelfehler folgen kann. Implementieren Sie sämtliche Regelkonzepte in MATLAB/SIMULINK.

Aufgabe 3.13. Gegeben ist das durch einen fremderregten Gleichstrommotor angetriebene mechanische System von Abbildung 3.9. Bestimmen Sie das mathematische Modell unter der Annahme, dass die Dynamik des Gleichstrommotors vernachlässigt werden kann. Verwenden Sie als Zustandsgrößen die Drehwinkelgeschwindigkeiten ω_1, ω_2 und den Differenzwinkel $\Delta\varphi = \varphi_1 - \varphi_2$, als Eingangsgröße den Ankerstrom i_a und als Ausgangsgröße ω_2. Die Ankerkreiskonstante hat den Wert $k_a = 1\,\mathrm{Nm/A}$, die Massenträgheitsmomente der beiden Massen sind $J_1 = 1.11\,\mathrm{kgm}^2$ und $J_2 = 10\,\mathrm{kgm}^2$ und die Feder- sowie die Dämpfungskonstante der Torsionswelle sind durch $c = 1\,\mathrm{Nm/rad}$ und $d = 0.1\,\mathrm{Nms/rad}$ gegeben.

Bestimmen Sie für eine Abtastzeit $T_a = 0.5\,\mathrm{s}$ das zugehörige Abtastmodell. Wenden Sie für dieses Abtastmodell, sofern möglich, sämtliche in diesem Kapitel vorgestellten Regelkonzepte an und testen Sie Ihre Ergebnisse in MATLAB/SIMULINK.

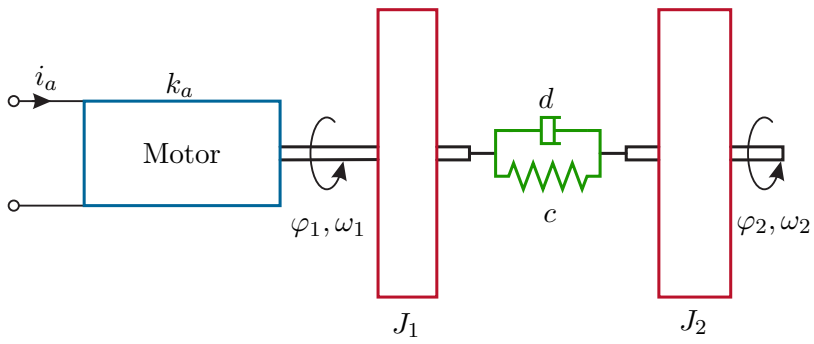


Abbildung 3.9: Mechanisches System mit Torsionswelle.

3.6 Literatur

[3.1] G. Franklin, J. Powell und M. Workman, *Digital Control of Dynamic Systems*, 3. Aufl. Menlo Park, USA: Addison–Weseley, 1998.

[3.2] K. Åström und B. Wittenmark, *Computer Controlled Systems: Theory and Design*. New York, USA: Prentice Hall, 1997.

[3.3] A. Bryson und Y. Ho, *Applied Optimal Control*. Washington, USA: He, 1975.

[3.4] P. Dorato, C. Abdallah und V. Cerone, *Linear Quadratic Control: An Introduction*. Florida, USA: Krieger Publishing Company, 2000.

A Grundlagen der Stochastik

Satz A.1 (Axiome der Wahrscheinlichkeit)**.** *Folgende Axiome der Wahrscheinlichkeit können definiert werden:*

(1) Die Wahrscheinlichkeit $\mathrm{P}(A)$ *eines Ergebnisses* A *bei einem Experiment ist eine eindeutig bestimmte reelle nichtnegative Zahl, die höchstens gleich 1 werden kann, es gilt also*

$$0 \leq \mathrm{P}(A) \leq 1 \ . \tag{A.1}$$

(2) Für ein sicheres Ereignis A *eines Experiments gilt*

$$\mathrm{P}(A) = 1 \ . \tag{A.2}$$

Für äquivalente Ereignisse B *und* C *bei einem Experiment gilt*

$$\mathrm{P}(B) = \mathrm{P}(C) \ . \tag{A.3}$$

(3) Schließen sich zwei Ereignisse B *und* C *bei einem Experiment gegenseitig aus, dann gilt*

$$\mathrm{P}(B + C) = \mathrm{P}(B) + \mathrm{P}(C) \ . \tag{A.4}$$

Definition A.1 (Zufallsvariable, stochastische Variable)**.** Eine Funktion X heißt Zufallsvariable oder stochastische Variable, wenn sie einem Zufallsexperiment zugeordnet ist und folgende Eigenschaften besitzt:

(1) Die Werte von X sind reelle Zahlen und

(2) für jede Zahl a und jedes Intervall I auf der Zahlengeraden ist die Wahrscheinlichkeit des Ereignisses "X hat den Wert a" oder "X liegt im Intervall I" im Einklang mit den Axiomen der Wahrscheinlichkeit.

Unter einem Zufallsexperiment versteht man in diesem Zusammenhang ein Experiment, bei dem das Ergebnis einer einzelnen Ausführung jeweils durch eine einzelne Zahl ausgedrückt werden kann.

Wenn X eine *diskrete Zufallsvariable* ist, dann können allen Werten von X, im Weiteren mit $x_1, x_2, \ldots$ bezeichnet, die zugehörigen Wahrscheinlichkeiten $p_1 = \mathrm{P}(X = x_1)$, $p_2 = \mathrm{P}(X = x_2), \ldots$ zugeordnet werden. Die Funktion

$$f(x) = \begin{cases} p_j & \text{für } x = x_j \\ 0 & \text{für alle übrigen } x \end{cases} \tag{A.5}$$

bezeichnet man dann als *Wahrscheinlichkeitsfunktion*. Da die Zufallsvariable X immer einen Wert x_j annimmt, muss auch gelten, dass die Summe aller Wahrscheinlichkeiten

$$\sum_j f(x_j) = 1 \tag{A.6}$$

ist. Die Wahrscheinlichkeit, dass die Zufallsvariable X im Intervall $a < X \leq b$ liegt, errechnet sich dann auf einfache Weise in der Form

$$\mathrm{P}(a < X \leq b) = \sum_{a<x_j\leq b} f(x_j) \ . \tag{A.7}$$

Trägt man die Wahrscheinlichkeit $\mathrm{P}(X \leq x)$ als Funktion von x auf, also die Wahrscheinlichkeit eines Experimentes, dessen Ergebnisse $X \leq x$ sind, dann erhält man die *Wahrscheinlichkeitsverteilungsfunktion*

$$F(x) = \mathrm{P}(X \leq x) = \sum_{x_j\leq x} f(x_j) \ . \tag{A.8}$$

Bei einer *stetigen Zufallsvariablen* X lässt sich die Wahrscheinlichkeitsverteilungsfunktion in Integralform

$$F(x) = \int_{-\infty}^{x} f(v)\,\mathrm{d}v \quad \text{mit} \quad F(\infty) = 1 \tag{A.9}$$

darstellen, wobei der Integrand $f(v)$ eine nichtnegative und bis auf endlich viele Punkte stetige Funktion ist, die man auch als *Wahrscheinlichkeitsdichtefunktion* bezeichnet.

Aufgabe A.1. Zeigen Sie, dass für die stetige Zufallsvariable X die Beziehung

$$\mathrm{P}(a < X \leq b) = \int_a^b f(v)\,\mathrm{d}v = F(b) - F(a)$$

gilt.

Abbildung A.1 zeigt den typischen Verlauf der Wahrscheinlichkeitsverteilungsfunktion einer diskreten und einer stetigen Zufallsvariablen.

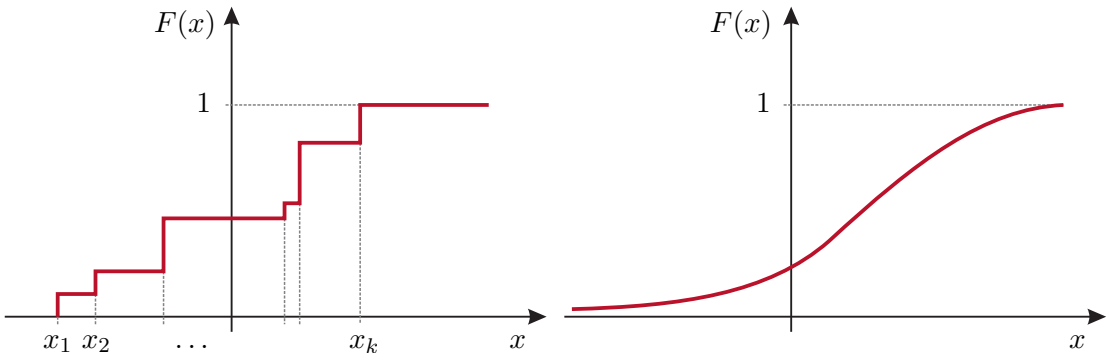


Abbildung A.1: Zur Wahrscheinlichkeitsverteilungsfunktion.

Definition A.2 (Erwartungswert)**.** Unter dem *Erwartungswert* $\mathrm{E}(X)$ einer Zufallsvariablen X (auch Mittelwert oder 1. Moment genannt) versteht man im Falle einer diskreten Verteilung

$$\mathrm{E}(X) = \sum_j x_j f(x_j) \tag{A.10}$$

und im Falle einer stetigen Verteilung

$$\mathrm{E}(X) = \int_{-\infty}^{\infty} x f(x)\,\mathrm{d}x\ . \tag{A.11}$$

Für eine Funktion $g(X)$ der Zufallsvariablen X gilt allgemein, dass sich der Erwartungswert der Funktion $g(X)$ in der Form

$$\mathrm{E}(g(X)) = \sum_j g(x_j) f(x_j) \quad \text{bzw.} \quad \mathrm{E}(g(X)) = \int_{-\infty}^{\infty} g(x) f(x)\,\mathrm{d}x \tag{A.12}$$

errechnet.

Aufgabe A.2. Zeigen Sie die Gültigkeit nachfolgender wichtiger Formel

$$\mathrm{E}(\alpha g(X) + \beta h(X)) = \alpha\,\mathrm{E}(g(X)) + \beta\,\mathrm{E}(h(X))$$

mit den Konstanten α und β. Zeigen Sie, dass damit auch

$$\mathrm{E}(\mathrm{E}(X)) = \mathrm{E}(X) \quad \text{und} \quad \mathrm{E}(X\,\mathrm{E}(X)) = \mathrm{E}(X)^2$$

gilt.

Definition A.3 (Varianz)**.** Die *Varianz* σ_X^2 einer Zufallsvariablen X misst die quadratische Abweichung vom Erwartungswert und ist in der Form

$$\sigma_X^2 = \mathrm{E}\Big([X - \mathrm{E}(X)]^2\Big) = \mathrm{E}\Big(X^2 - 2X\,\mathrm{E}(X) + \mathrm{E}(X)^2\Big) = \mathrm{E}\Big(X^2\Big) - \mathrm{E}(X)^2 \tag{A.13}$$

definiert. Die Varianz wird auch als zweites zentrales Moment und deren positive Quadratwurzel σ_X als Standardabweichung bezeichnet.

Die bisherigen Betrachtungen lassen sich nun einfach auf *mehrere Zufallsvariablen* übertragen. Bei zwei Zufallsvariablen X und Y nimmt die Wahrscheinlichkeitsverteilungsfunktion im diskreten Fall die Form

$$F(x,y) = \mathrm{P}(X \le x, Y \le y) = \sum_{x_k \le x} \sum_{y_j \le y} f(x_k, y_j) \tag{A.14}$$

mit

$$f(x,y) = \begin{cases} p_{kj} & \text{für } x = x_k,\ y = y_j \\ 0 & \text{für alle übrigen } (x,y) \end{cases} \quad \text{mit} \quad \sum_k \sum_j f(x_k, y_j) = 1 \tag{A.15}$$

an und im stetigen Fall gilt

$$F(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(v,w)\,\mathrm{d}v\,\mathrm{d}w \quad \text{mit} \quad F(\infty,\infty) = 1\ . \tag{A.16}$$

Die so genannten *Randverteilungen* sind durch nachfolgende Ausdrücke

$$F_1(x) = \mathrm{P}(X \le x, Y \text{ beliebig}) = \sum_{x_k \le x} \underbrace{\sum_j f(x_k, y_j)}_{f_1(x_k)} \tag{A.17}$$

bzw.

$$F_1(x) = \mathrm{P}(X \le x, Y \text{ beliebig}) = \int_{-\infty}^{x} \underbrace{\int_{-\infty}^{\infty} f(v,w)\,\mathrm{d}v}_{f_1(w)}\,\mathrm{d}w \tag{A.18}$$

und

$$F_2(y) = \mathrm{P}(X \text{ beliebig}, Y \le y) = \sum_{y_j \le y} \underbrace{\sum_k f(x_k, y_j)}_{f_2(y_j)} \tag{A.19}$$

bzw.

$$F_2(y) = \mathrm{P}(X \text{ beliebig}, Y \le y) = \int_{-\infty}^{y} \underbrace{\int_{-\infty}^{\infty} f(v,w)\,\mathrm{d}w}_{f_2(v)}\,\mathrm{d}v \tag{A.20}$$

gegeben.

Definition A.4 (Unabhängigkeit von Zufallsvariablen)**.** Zwei Zufallsvariablen X und Y sind genau dann unabhängig, wenn für jedes Paar von Ereignissen der Form

$$a_1 < X \leq b_1 \quad \text{und} \quad a_2 < Y \leq b_2 \tag{A.21}$$

die Beziehung

$$\mathrm{P}(a_1 < X \leq b_1, a_2 < Y \leq b_2) = \mathrm{P}(a_1 < X \leq b_1)\,\mathrm{P}(a_2 < Y \leq b_2) \tag{A.22}$$

und damit auch

$$f(x,y) = f_1(x) f_2(y) \quad \text{sowie} \quad F(x,y) = F_1(x) F_2(y) \tag{A.23}$$

mit $f_1(x)$, $f_2(y)$, $F_1(x)$ und $F_2(y)$ nach (A.17)–(A.20) gilt.

Satz A.2 (Erwartungswert und Kovarianz zweier Zufallsvariablen)**.** *Für den Erwartungswert einer Funktion $g(X,Y)$ in den zwei Zufallsvariablen X und Y gilt*

$$\mathrm{E}(g(X,Y)) = \sum_k \sum_j g(x_k, y_j) f(x_k, y_j) \tag{A.24}$$

bzw.

$$\mathrm{E}(g(X,Y)) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) f(x,y)\,\mathrm{d}x\,\mathrm{d}y \ . \tag{A.25}$$

Der Erwartungswert $\mathrm{E}(X+Y)$ einer Summe von Zufallsvariablen X und Y errechnet sich zu

$$\mathrm{E}(X+Y) = \mathrm{E}(X) + \mathrm{E}(Y) \ . \tag{A.26}$$

Sind zwei Zufallsvariablen X und Y unabhängig, dann gilt die Beziehung (hier für den diskreten Fall mit Hilfe von (A.19) gezeigt)

$$\mathrm{E}(XY) = \sum_k \sum_j x_k y_j f(x_k, y_j) = \sum_k \sum_j x_k y_j f_1(x_k) f_2(y_j) = \mathrm{E}(X)\,\mathrm{E}(Y) \ . \tag{A.27}$$

Wenn eine Zufallsvariable Z sich als Summe zweier Zufallsvariablen X und Y ergibt, dann folgt nach (A.13) für die Varianz von Z

$$\begin{aligned} \sigma_Z^2 &= \mathrm{E}\left(Z^2\right) - \mathrm{E}(Z)^2 = \mathrm{E}\left(X^2\right) + 2\,\mathrm{E}(XY) + \mathrm{E}\left(Y^2\right) \\ &\quad - \mathrm{E}(X)^2 - 2\,\mathrm{E}(X)\,\mathrm{E}(Y) - \mathrm{E}(Y)^2 = \sigma_X^2 + \sigma_Y^2 + 2\sigma_{XY} \end{aligned} \tag{A.28}$$

mit der so genannten Kovarianz

$$\sigma_{XY} = \mathrm{E}(XY) - \mathrm{E}(X)\,\mathrm{E}(Y) \ . \tag{A.29}$$

Man überzeugt sich leicht, dass (A.29) auch in der Form

$$\sigma_{XY} = \mathrm{E}([X - \mathrm{E}(X)][Y - \mathrm{E}(Y)]) \tag{A.30}$$

angeschrieben werden kann. Sind die Zufallsvariablen X und Y unabhängig, so gilt nach (A.27) $\sigma_{XY} = 0$.

Definition A.5 (Korrelationskoeffizient und Kovarianzmatrix)**.** Der Quotient

$$r = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} \tag{A.31}$$

wird als *Korrelationskoeffizient* zwischen den Zufallsvariablen X und Y bezeichnet. Ist $r = 0$, dann sind X und Y *unkorreliert*. Man erkennt auch, dass zwei unabhängige Zufallsvariablen X und Y wegen $\sigma_{XY} = 0$ auch unkorreliert sein müssen. Der Korrelationskoeffizient $-1 \leq r \leq 1$ gibt ein Maß für die *lineare Abhängigkeit* von X und Y an.

Für eine vektorwertige Zufallsvariable $\mathbf{X}^{\mathrm{T}} = \begin{bmatrix} X_1 & X_2 & \dots & X_n \end{bmatrix}$ gilt nun

$$\mathrm{E}\left(\mathbf{X}^{\mathrm{T}}\right) = \begin{bmatrix} \mathrm{E}(X_1) & \mathrm{E}(X_2) & \dots & \mathrm{E}(X_n) \end{bmatrix} \tag{A.32}$$

und unter der *Kovarianzmatrix* der vektorwertigen Zufallsvariablen $\mathbf{X}$ versteht man die Matrix

$$\mathrm{cov}(\mathbf{X}) = \mathrm{E}\Big([\mathbf{X} - \mathrm{E}(\mathbf{X})][\mathbf{X} - \mathrm{E}(\mathbf{X})]^{\mathrm{T}}\Big) = \begin{bmatrix} \sigma_{X_1}^2 & \sigma_{X_1X_2} & \cdots & \sigma_{X_1X_n} \\ \sigma_{X_2X_1} & \sigma_{X_2}^2 & \cdots & \sigma_{X_2X_n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{X_nX_1} & \sigma_{X_nX_2} & \cdots & \sigma_{X_n}^2 \end{bmatrix} . \tag{A.33}$$

Man erkennt, dass die Kovarianzmatrix symmetrisch ist.

Aufgabe A.3. Zeigen Sie, dass die Beziehung

$$\mathrm{E}\Big(\|\mathbf{X} - \mathrm{E}(\mathbf{X})\|_2^2\Big) = \mathrm{E}\Big([\mathbf{X} - \mathrm{E}(\mathbf{X})]^{\mathrm{T}}[\mathbf{X} - \mathrm{E}(\mathbf{X})]\Big) = \mathrm{spur}(\mathrm{cov}(\mathbf{X}))$$

mit $\mathrm{spur}(\mathbf{S}) = \sum_i s_{ii}$ gilt.

Aus dem bisher Gesagten folgt, dass die Kovarianzmatrix einer vektorwertigen Zufallsvariablen $\mathbf{X}$, deren Komponenten X_i, X_j für $i \neq j = 1, \dots, n$ unkorreliert sind, eine Diagonalmatrix ist.

Für stochastische Zeitsignale, die aus statistisch identischen Signalquellen stammen, existiert nicht nur eine einzige Realisierung $x_1(t)$ sondern eine ganze Familie (Ensemble) von Zufallszeitfunktionen $\{x_j(t)\}$. Dieses Ensemble von Zufallszeitfunktionen bzw. Zufallsfolgen im zeitdiskreten Fall wird als *zeitkontinuierlicher* bzw. *zeitdiskreter stochastischer Prozess* $\mathrm{x}(t)$ bezeichnet. Eine einzelne Realisierung $x_j(t)$ für festes j nennt man auch *Musterfunktion*. Für jeden festen Zeitpunkt t ist $\mathrm{x}(t)$ eine Zufallsvariable mit einer Wahrscheinlichkeitsverteilungsfunktion (siehe (A.8))

$$F(x, t) = \mathrm{P}(\mathrm{x}(t) \leq x) \ . \tag{A.34}$$

Die Wahrscheinlichkeitsdichtefunktion errechnet sich dann gemäß (A.9) zu

$$f(x,t) = \frac{\partial F(x,t)}{\partial x} \ . \tag{A.35}$$

Definition A.6 (Mittelwert, Auto- und Kreuzkorrelationsfunktion eines stochastischen Prozesses)**.** Der Mittelwert $\eta_{\mathrm{x}}(t)$ eines stochastischen Prozesses $\mathrm{x}(t)$ lautet (siehe auch (A.11))

$$\eta_{\mathrm{x}}(t) = \mathrm{E}(\mathrm{x}(t)) = \int_{-\infty}^{\infty} x f(x,t)\,\mathrm{d}x \ . \tag{A.36}$$

Unter der *Autokorrelationsfunktion* $\Phi_{\mathrm{xx}}(t_1,t_2)$ eines stochastischen Prozesses $\mathrm{x}(t)$ versteht man den Erwartungswert des Produktes $\mathrm{x}(t_1)\mathrm{x}(t_2)$

$$\Phi_{\mathrm{xx}}(t_1,t_2) = \mathrm{E}(\mathrm{x}(t_1)\mathrm{x}(t_2)) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 f(x_1,t_1) x_2 f(x_2,t_2)\,\mathrm{d}x_1\,\mathrm{d}x_2 \ . \tag{A.37}$$

Die *Kreuzkorrelationsfunktion* $\Phi_{\mathrm{xy}}(t_1,t_2)$ zweier stochastischer Prozesse $\mathrm{x}(t)$ und $\mathrm{y}(t)$ ist durch die Beziehung

$$\Phi_{\mathrm{xy}}(t_1,t_2) = \mathrm{E}(\mathrm{x}(t_1)\mathrm{y}(t_2)) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x f(x,t_1) y f(y,t_2)\,\mathrm{d}x\,\mathrm{d}y \tag{A.38}$$

gegeben.

Definition A.7 (Auto- und Kreuzkovarianzfunktion)**.** Die *Autokovarianzfunktion* $C_{\mathrm{xx}}(t_1,t_2)$ eines stochastischen Prozesses $\mathrm{x}(t)$ lautet

$$C_{\mathrm{xx}}(t_1,t_2) = \mathrm{E}([\mathrm{x}(t_1)-\eta_{\mathrm{x}}(t_1)][\mathrm{x}(t_2)-\eta_{\mathrm{x}}(t_2)]) = \Phi_{\mathrm{xx}}(t_1,t_2) - \eta_{\mathrm{x}}(t_1)\eta_{\mathrm{x}}(t_2) \ . \tag{A.39}$$

Die *Kreuzkovarianzfunktion* $C_{\mathrm{xy}}(t_1,t_2)$ zweier stochastischer Prozesse $\mathrm{x}(t)$ und $\mathrm{y}(t)$ ergibt sich analog zu

$$C_{\mathrm{xy}}(t_1,t_2) = \mathrm{E}([\mathrm{x}(t_1)-\eta_{\mathrm{x}}(t_1)][\mathrm{y}(t_2)-\eta_{\mathrm{y}}(t_2)]) = \Phi_{\mathrm{xy}}(t_1,t_2) - \eta_{\mathrm{x}}(t_1)\eta_{\mathrm{y}}(t_2) \ . \tag{A.40}$$

Aufgabe A.4. Zeigen Sie die Gültigkeit der rechten Identität von (A.39).

Definition A.8 (Stationärer stochastischer Prozess)**.** Man bezeichnet einen stochastischen Prozess $\mathrm{x}(t)$ als *stationär im strengen Sinne*, wenn die statistischen Eigenschaften invariant gegenüber Zeitverschiebungen sind, d. h. $\mathrm{x}(t)$ und $\mathrm{x}(t+c)$ haben für alle c die identischen statistischen Eigenschaften. Der stochastische Prozess $\mathrm{x}(t)$ ist *stationär im weiteren Sinne*, wenn der Mittelwert $\eta_{\mathrm{x}}(t) = \eta_{\mathrm{x}}$ konstant ist und die Autokorrelationsfunktion nur von der Zeitdifferenz $\tau = t_2 - t_1$ abhängt, also gilt $\Phi_{\mathrm{xx}}(\tau) = \mathrm{E}(\mathrm{x}(t)\mathrm{x}(t+\tau))$.

Die bisher gezeigten Funktionen (A.36)–(A.40) berücksichtigen sämtliche Realisierungen, also das gesamte Ensemble, eines stochastischen Prozesses. Nach der so genannten

Ergoden-Hypothese kann man die Funktionen (A.36)–(A.40) auch mit Hilfe einer einzigen Musterfunktion anschreiben, falls *unendlich lange Zeitabschnitte* betrachtet werden. Ergodische Prozesse sind auch stationär, die Umkehrung gilt im Allgemeinen aber nicht.

Damit ergeben sich die nachfolgende Ausdrücke, jeweils links für den zeitkontinuierlichen und rechts für den zeitdiskreten Fall:

(1) Mittelwert

$$\eta_x = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t)\,\mathrm{d}t \quad \text{bzw.} \quad \eta_x = \lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} x_k \tag{A.41}$$

(2) Autokorrelationsfunktion

$$\Phi_{xx}(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t)x(t+\tau)\,\mathrm{d}t \quad \text{bzw.} \quad \Phi_{xx}(\tau) = \lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} x_k x_{k+\tau} \tag{A.42}$$

(3) Kreuzkorrelationsfunktion

$$\Phi_{xy}(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t)y(t+\tau)\,\mathrm{d}t \quad \text{bzw.} \quad \Phi_{xy}(\tau) = \lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} x_k y_{k+\tau} \tag{A.43}$$

Aufgabe A.5. Zeigen Sie, dass $\Phi_{xx}(\tau)$ folgende Eigenschaften erfüllt:

1. $\Phi_{xx}(\tau) = \Phi_{xx}(-\tau)$
2. $\Phi_{xx}(0) = \mathrm{E}\left(x(t)^2\right)$
3. $\Phi_{xx}(\infty) = \mathrm{E}(x(t))^2$
4. $\Phi_{xx}(\tau) \leq \Phi_{xx}(0)$

Aufgabe A.6. Zeigen Sie, dass $\Phi_{xy}(\tau)$ folgende Eigenschaften erfüllt:

1. $\Phi_{xy}(\tau) = \Phi_{yx}(-\tau)$
2. $\Phi_{xy}(0) = \mathrm{E}(x(t)y(t))$
3. $\Phi_{xy}(\infty) = \mathrm{E}(x(t))\,\mathrm{E}(y(t))$
4. $\Phi_{xy}(\tau) \leq \frac{1}{2}[\Phi_{xx}(0) + \Phi_{yy}(0)]$

Aufgabe A.7. Wie lauten im ergodischen Fall die Ausdrücke für die Autovarianz- und Autokovarianzfunktion?

Definition A.9 (Weißes Rauschen)**.** Ein stochastischer Prozess $\mathrm{v}(t)$ wird als *(striktes) weißes Rauschen* bezeichnet, wenn für alle Zeitpunkte $t_1 \neq t_2$ $\mathrm{v}(t_1)$ und $\mathrm{v}(t_2)$ statistisch unabhängig sind bzw.

$$C_{\mathrm{vv}}(\tau) = C_0\delta(\tau) \quad \text{mit} \quad \delta(\tau) = \begin{cases} 1 & \text{für } \tau = 0 \\ 0 & \text{sonst} \end{cases} \quad \text{und } C_0 > 0\ . \tag{A.44}$$

gilt. Es wird weiterhin angenommen, dass der Mittelwert $\eta_{\mathrm{v}} = 0$ ist.

Hinweis: Man kann zeigen, dass im zeitkontinuierlichen Fall die mittlere Leistung des weißen Rauschens unendlich ist und somit dieser Prozess nicht ideal realisiert werden kann. Im Gegensatz dazu bleibt im Zeitdiskreten, wo die Signalfolgenwerte im endlichen Abstand zueinander stehen, die Leistung endlich, womit ideales weißes Rauschen auch realisierbar ist.

A.1 Literatur

[A.1] R. Isermann, *Identifikation dynamischer Systeme 1 und 2*, 2. Aufl. Berlin, Deutschland: Springer, 1992.

[A.2] E. Kreyszig, *Statistische Methoden und ihre Anwendungen*, 7. Aufl. Göttingen, Deutschland: Vandenhoeck & Ruprecht, 1998.

[A.3] D. Luenberger, *Optimization by Vector Space Methods*. New York, USA: John Wiley & Sons, 1969.

[A.4] A. Papoulis und S. Pillai, *Probability, Random Variables and Stochastic Processes*, 4. Aufl. New York, USA: McGraw-Hill, 2002.