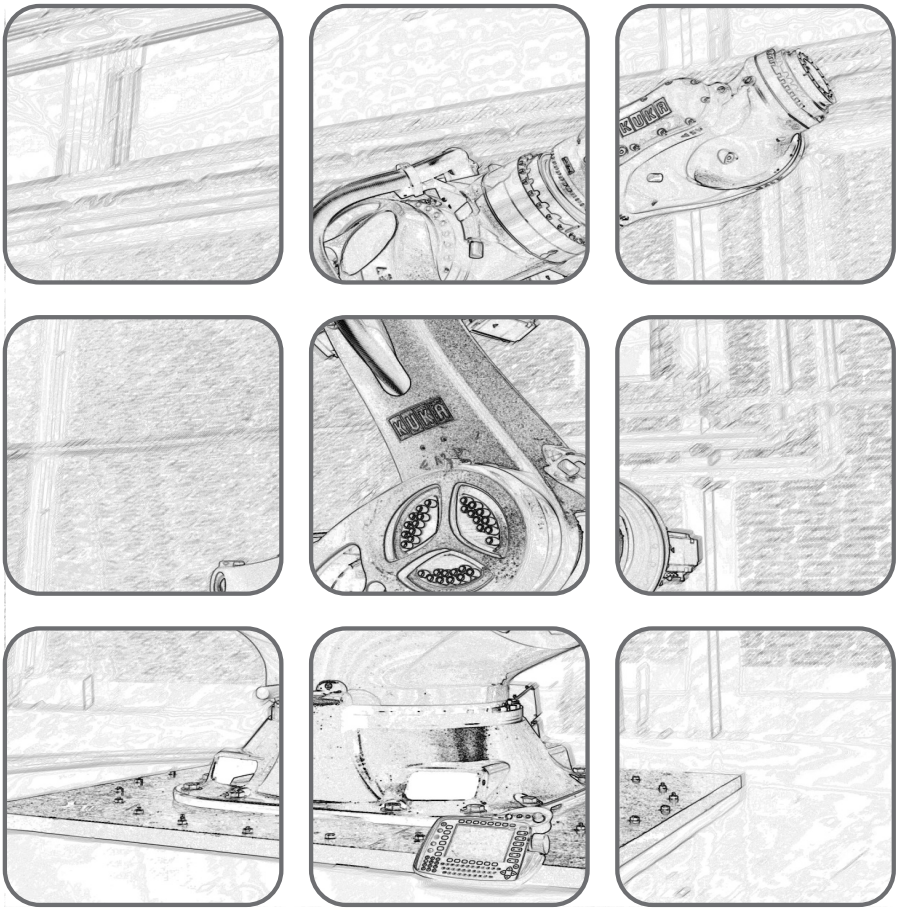
TECHNISCHE UNIVERSITÄT WIEN

AUTOMATION & CONTROL INSTITUTE
INSTITUT FÜR AUTOMATISIERUNGS- & REGELUNGSTECHNIK

Vorlesung
WS 2026/2027

REGELUNGSSYSTEME

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



Regelungssysteme

Vorlesung
WS 2026/2027

Univ.-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

© Institut für Automatisierungs- und Regelungstechnik, TU Wien

Inhaltsverzeichnis

1 Identifikationsverfahren **1**
- 1.1 Allgemeines . . . 1
- 1.2 Nicht-parametrische Verfahren . . . 4
  - 1.2.1 Impulsantwortanalyse . . . 4
  - 1.2.2 Frequenzbereichsanalyse - harmonische Anregung . . . 5
  - 1.2.3 Frequenzbereichsanalyse - ETFE . . . 8
- 1.3 Parametrische Verfahren . . . 12
  - 1.3.1 Modellstrukturen . . . 12
  - 1.3.2 Least Squares . . . 17
  - 1.3.3 Least-Squares Identifikation . . . 23
  - 1.3.4 Rekursive Least-Squares (RLS) Identifikation . . . 28
  - 1.3.5 Methode der gewichteten kleinsten Quadrate . . . 31
  - 1.3.6 Least-Mean Squares (LMS) Identifikation . . . 34
- 1.4 Literatur . . . 37

2 Optimale Schätzer **38**
- 2.1 Gauß-Markov-Schätzung . . . 42
  - 2.1.1 Quadratische Minimierung mit affinen NB . . . 44
- 2.2 Minimum-Varianz-Schätzung . . . 48
  - 2.2.1 Rekursive Minimum-Varianz-Schätzung . . . 52
- 2.3 Das Kalman-Filter . . . 55
  - 2.3.1 Das Kalman-Filter als optimaler Beobachter . . . 58
  - 2.3.2 Eigenschaften des stationären Kalman-Filters . . . 61
- 2.4 Das Extended Kalman-Filter . . . 65
- 2.5 Das Unscented Kalman-Filter . . . 72
  - 2.5.1 Erwartungswert und Kovarianz von nichtlinearen Transformationen 72
  - 2.5.2 Die Unscented-Transformation . . . 77
  - 2.5.3 Zustandsschätzung dynamischer Systeme mit Hilfe der Unscented-Transformation . . . 83
- 2.6 Literatur . . . 88

3 Optimaler Zustandsregler **89**
- 3.1 Dyn. Programmierung nach Bellman . . . 89
- 3.2 Das LQR-Problem . . . 92
- 3.3 Das LQR-Problem mit stochastischer Störung . . . 98
- 3.4 Das LQG-Regelungsproblem . . . 101
- 3.5 Erweiterte Konzepte der Zustandsregelung . . . 109
  - 3.5.1 Feedforward der geschätzten Störung . . . 109

©W. Kemmetmüller, A. Kugi, Institut für Automatisierungs- und Regelungstechnik, TU Wien

3.5.2 Zustandsregler- und Zustandsbeobachterentwurf mit Integralanteil 111
3.5.3 Zustandsregler- und Zustandsbeobachterentwurf mit Führungsgrößen 112
3.5.4 Das Feedforward-Konzept 114
3.6 Literatur 119

A Grundlagen der Stochastik **120**
A.1 Literatur 129

©W. Kemmetmüller, A. Kugi, Institut für Automatisierungs- und Regelungstechnik, TU Wien

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

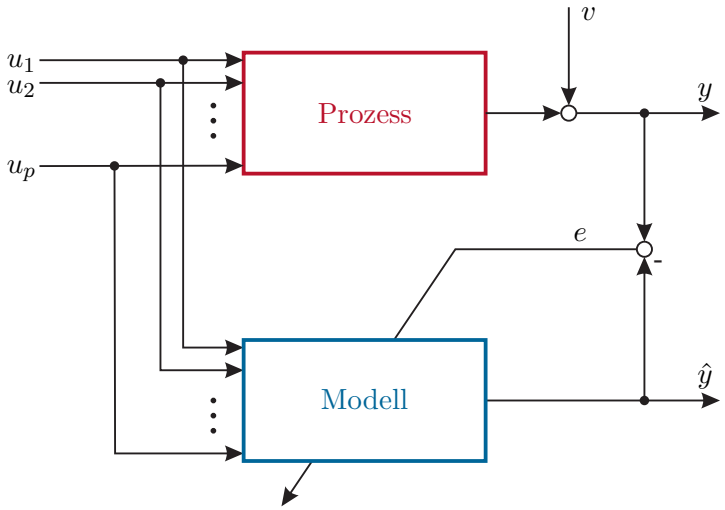


Abbildung 1.1: Zum Konzept der Prozessidentifikation.

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:

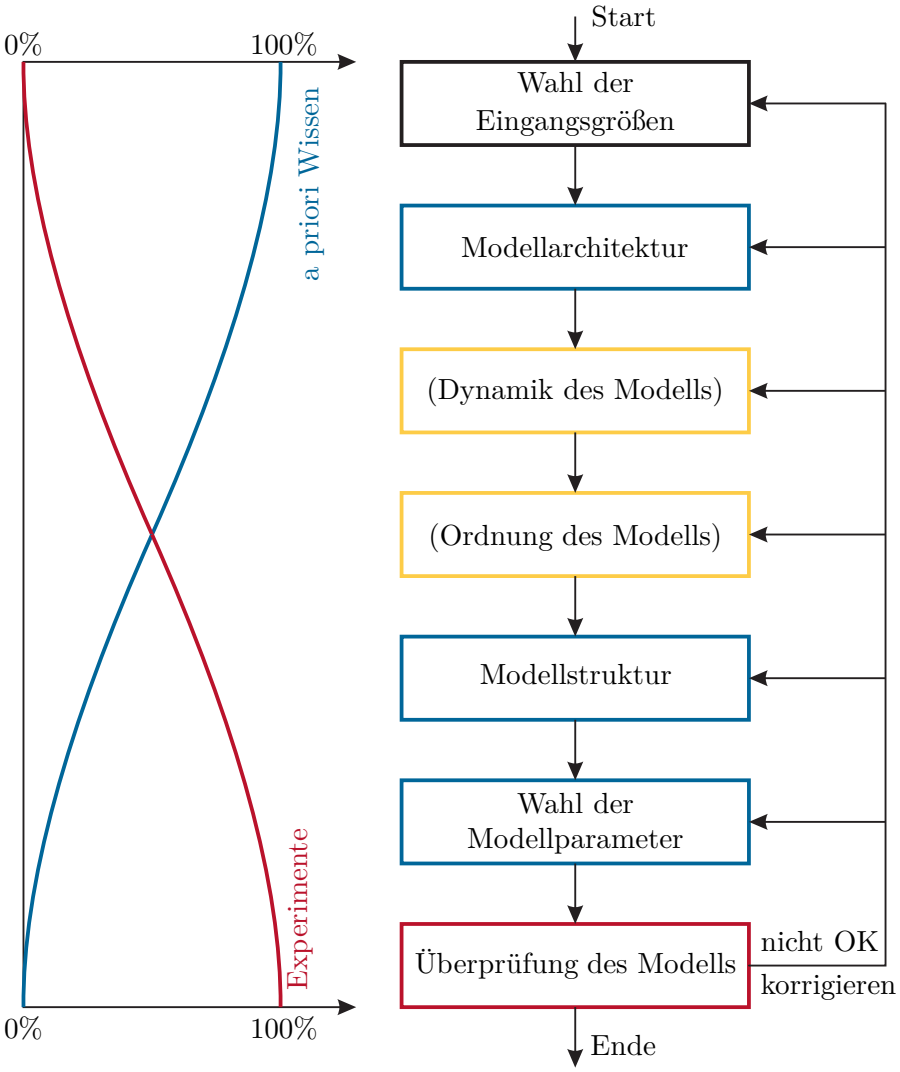


Abbildung 1.2: Ablauf der Identifikation.

©W. Kemmetmüller, A. Kugi, Institut für Automatisierungs- und Regelungstechnik, TU Wien

(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

©W. Kemmetmüller, A. Kugi, Institut für Automatisierungs- und Regelungstechnik, TU Wien

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.

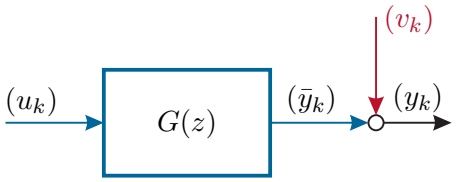


Abbildung 1.3: Betrachtetes System für die nicht-parametrische Identifikation.

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

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

©W. Kemmetmüller, A. Kugi, Institut für Automatisierungs- und Regelungstechnik, TU Wien

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

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

©W. Kemmetmüller, A. Kugi, Institut für Automatisierungs- und Regelungstechnik, TU Wien

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

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

©W. Kemmetmüller, A. Kugi, Institut für Automatisierungs- und Regelungstechnik, TU Wien

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)

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

©W. Kemmetmüller, A. Kugi, Institut für Automatisierungs- und Regelungstechnik, TU Wien

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

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

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

©W. Kemmetmüller, A. Kugi, Institut für Automatisierungs- und Regelungstechnik, TU Wien

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

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

©W. Kemmetmüller, A. Kugi, Institut für Automatisierungs- und Regelungstechnik, TU Wien

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

©W. Kemmetmüller, A. Kugi, Institut für Automatisierungs- und Regelungstechnik, TU Wien

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.

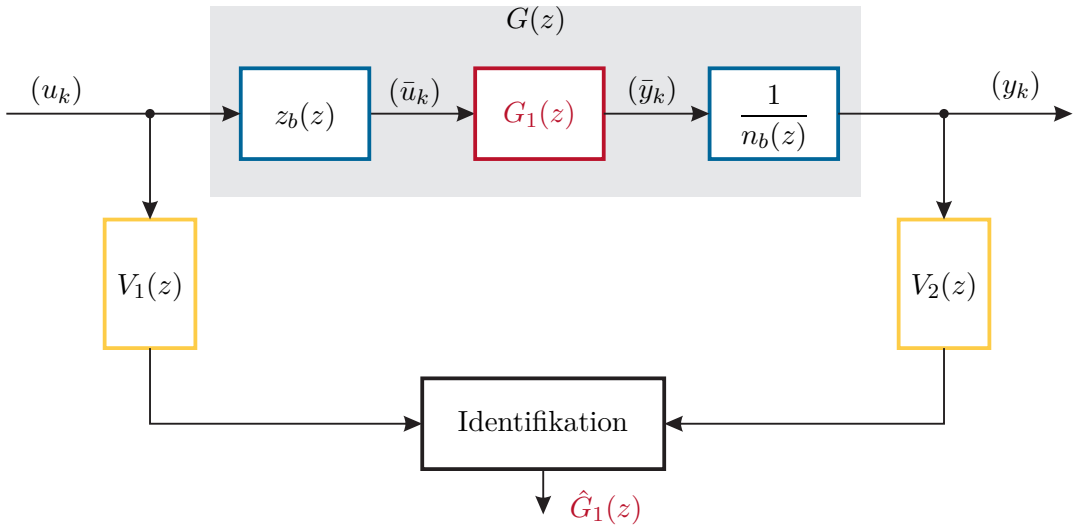


Abbildung 1.4: Identifikation unter Berücksichtigung bekannter Teile der Übertragungsfunktion (Verfahren nach Clary).

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

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

©W. Kemmetmüller, A. Kugi, Institut für Automatisierungs- und Regelungstechnik, TU Wien

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 ausreichend, 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 $w_{1,k} = r$.

©W. Kemmetmüller, A. Kugi, Institut für Automatisierungs- und Regelungstechnik, TU Wien

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.

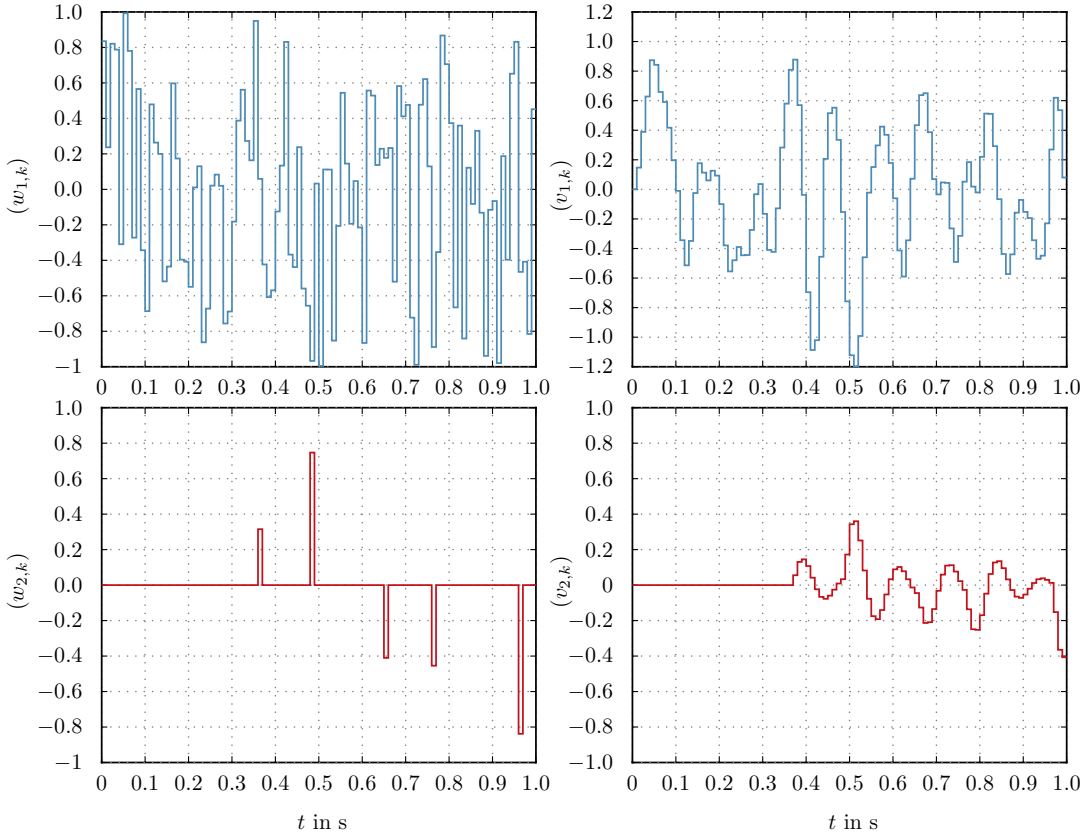


Abbildung 1.5: Antworten $(v_{1,k})$ und $(v_{2,k})$ des linearen, zeitinvarianten Abtastsystems zweiter Ordnung (1.41) auf weißes Rauschen $(w_{1,k})$ bzw. $(w_{2,k})$ mit unterschiedlicher Wahrscheinlichkeitsdichtefunktion.

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

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

©W. Kemmetmüller, A. Kugi, Institut für Automatisierungs- und Regelungstechnik, TU Wien

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.

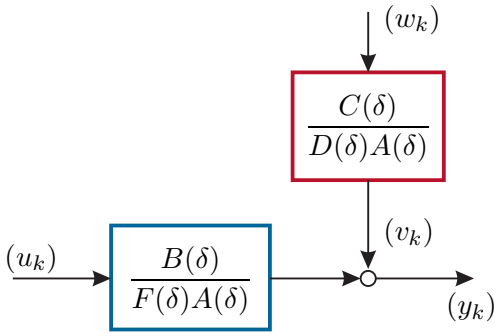


Abbildung 1.6: Zu den Modellstrukturen der parametrischen Identifikation.

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:

©W. Kemmetmüller, A. Kugi, Institut für Automatisierungs- und Regelungstechnik, TU Wien

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.

©W. Kemmetmüller, A. Kugi, Institut für Automatisierungs- und Regelungstechnik, TU Wien

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.

©W. Kemmetmüller, A. Kugi, Institut für Automatisierungs- und Regelungstechnik, TU Wien

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

©W. Kemmetmüller, A. Kugi, Institut für Automatisierungs- und Regelungstechnik, TU Wien

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

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

©W. Kemmetmüller, A. Kugi, Institut für Automatisierungs- und Regelungstechnik, TU Wien

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 $n = 2$.

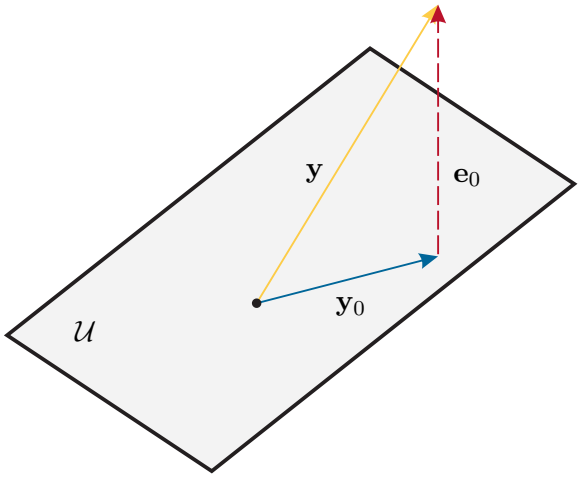


Abbildung 1.7: Zur Methode der kleinsten Fehlerquadrate.

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

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

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

©W. Kemmetmüller, A. Kugi, Institut für Automatisierungs- und Regelungstechnik, TU Wien

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

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

©W. Kemmetmüller, A. Kugi, Institut für Automatisierungs- und Regelungstechnik, TU Wien

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

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

©W. Kemmetmüller, A. Kugi, Institut für Automatisierungs- und Regelungstechnik, TU Wien

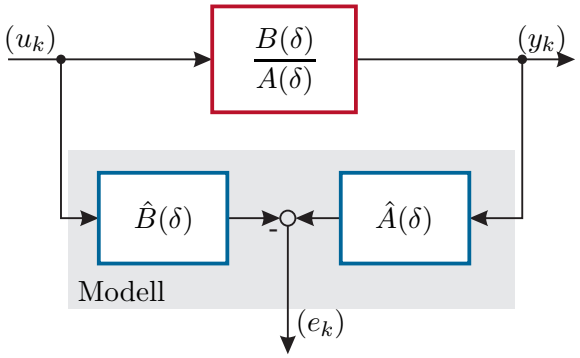


Abbildung 1.8: Zum verallgemeinerten Gleichungsfehler.

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 $\mathrm{rang}(\mathbf{S}_N) \neq \mathrm{rang}([\mathbf{S}_N, \mathbf{y}_N])$ 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.

©W. Kemmetmüller, A. Kugi, Institut für Automatisierungs- und Regelungstechnik, TU Wien

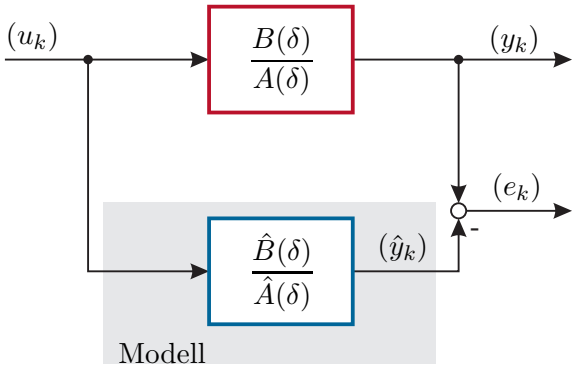


Abbildung 1.9: Zum Ausgangsfehler.

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

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

©W. Kemmetmüller, A. Kugi, Institut für Automatisierungs- und Regelungstechnik, TU Wien

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

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

©W. Kemmetmüller, A. Kugi, Institut für Automatisierungs- und Regelungstechnik, TU Wien

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.

©W. Kemmetmüller, A. Kugi, Institut für Automatisierungs- und Regelungstechnik, TU Wien

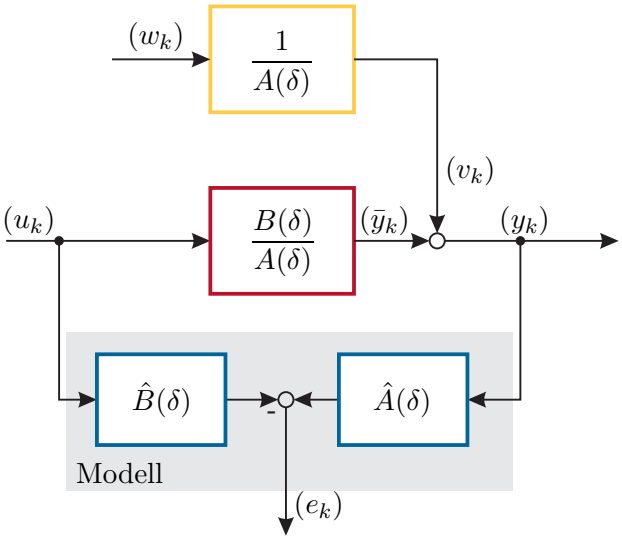


Abbildung 1.10: Zur Least-Squares Identifikation mit stochastischer Störung.

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

©W. Kemmetmüller, A. Kugi, Institut für Automatisierungs- und Regelungstechnik, TU Wien

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

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

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

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)

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

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

$$\vdots \quad \vdots$$

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

©W. Kemmetmüller, A. Kugi, Institut für Automatisierungs- und Regelungstechnik, TU Wien

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

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

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

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

©W. Kemmetmüller, A. Kugi, Institut für Automatisierungs- und Regelungstechnik, TU Wien

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

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

©W. Kemmetmüller, A. Kugi, Institut für Automatisierungs- und Regelungstechnik, TU Wien

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.

©W. Kemmetmüller, A. Kugi, Institut für Automatisierungs- und Regelungstechnik, TU Wien

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

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

©W. Kemmetmüller, A. Kugi, Institut für Automatisierungs- und Regelungstechnik, TU Wien

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.

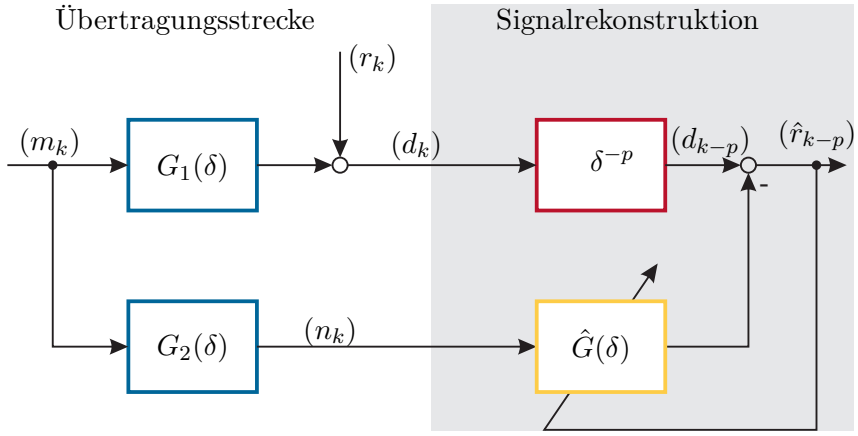


Abbildung 1.11: Zur adaptiven Signalfilterung mit dem LMS-Algorithmus.

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

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

©W. Kemmetmüller, A. Kugi, Institut für Automatisierungs- und Regelungstechnik, TU Wien

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?

©W. Kemmetmüller, A. Kugi, Institut für Automatisierungs- und Regelungstechnik, TU Wien

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.

©W. Kemmetmüller, A. Kugi, Institut für Automatisierungs- und Regelungstechnik, TU Wien