

# Zusammenfassung Modellbildung und Systemidentifikation

## Grundlagen

### Norm

Motivation: Maß für Fehler ist notwendig

Eigenschaften:

- Abbildung von Vektor auf reelle Zahl
- $\|x\| = 0 \Leftrightarrow x = 0$
- $\|\alpha \cdot x\| = |\alpha| \cdot \|x\|$
- $\|x + y\| \leq \|x\| + \|y\|$

Beispiele:

Euklidische Norm (2-Norm):  $\|x\|_2 = \sqrt{\sum_{i=1}^n x_i^2}$

Summennorm (Betragsnorm):  $\|x\|_1 = \sqrt{\sum_{i=1}^n |x_i|}$

Maximumsnorm :  $\|x\|_\infty = \max_i |x_i|$

### Matrixnormen

Matrixnorm heißt verträglich mit Vektornorm wenn:

$$\|A \cdot x\| \leq \|A\| \cdot \|x\|$$

(A: Matrix, V: Vektor)

Durch Vektornorm induzierte Matrixnorm:

$$\|A\| := \max_{x \neq 0} \frac{\|A \cdot x\|}{\|x\|}$$

(A: Matrix, V: Vektor)

Bsp: Spektralnrm (induziert durch 2-Norm):

$$\|A\|_2 = \max_{\|x\|_2=1} \|A \cdot x\| = \sqrt{\{\lambda(A^T \cdot A)\}}$$

wobei  $\lambda$  die Singulärwerte der Matrix  $A$  sind.

Bsp für nicht induzierte Norm: Frobenius-Norm:

$$\|A\|_F = \sqrt{\text{trace}(A^T \cdot A)}$$

## Dynamische Systeme

- mind 1 Eingang, mind 1 Ausgang, enthält “Gedächtnis” (=Energiespeicher)
- **linear**, wenn Superpositionsprinzip gilt:

$$G(\alpha_1 u_1 + \alpha_2 u_2) = \alpha_1 \cdot G \circ u_1 + \alpha_2 \cdot G \circ u_2 \text{ mit } \alpha_1, \alpha_2 \in \mathbb{R}$$

- **zeitinvariant**, wenn beliebige zeitl. Verschiebung von Eingangssignal entsprechende zeitl. Verschiebung am Ausgang bewirkt
- nicht-parametrische Modellierung: Sprungantwort, Impulsantwort, Frequenzgang

## Parametrische Modelle

physikalische Modellierung => DLG => Laplace-Transformation => Zustandsraummodell mit Systemmatrizen

Abtastung liefert diskretes Modell:

$$\begin{aligned} x[k+1] &= Ax[k] + Bu[k] \\ y[k] &= Cx[k] \end{aligned}$$

Berechnung der Systemmatrizen aus zeitkontinuierlichen Modell:

$$\begin{aligned} A_D &= e^{A^C \cdot \Delta t} \\ B_D &= B^C \cdot \int_{\tau}^{\Delta t} e^{A^C \cdot \tau} d\tau \\ C_D &= C^C \end{aligned}$$

## Modellarten

$u(z)$  : Systemeingang,  $v(z)$  : Fehler ### Klassifikation nach Störung

### AR (autoregressives System)

$$y = \frac{1}{C(z)} \cdot v$$

### MA (Moving Average)

$$y = D(z) \cdot v$$

### ARMA

$$y = \frac{D(z)}{C(z)} \cdot v$$

### ARX (autoregressive with external input)

$$y = \frac{B(z)}{A(z)}u + \frac{1}{A(z)}v$$

### AR MAX (autoregressive with external input)

$$y = \frac{B(z)}{A(z)}u + \frac{D(z)}{A(z)}v$$

### Klassifikation nach determ. Eingang (Ausgangsfehler-Modell):

#### Fir

$$y = B(z)u + v$$

#### OE (output error)

$$y = \frac{B(z)}{A(z)}u + v$$

**BJ (Box-Jenkins)**

$$y(z) = \frac{B(z)}{A(z)} \cdot u + \frac{D(z)}{C(z)} \cdot v$$

## MkQ für Statische Systeme

### Parameterlineare Modelle

**Prinzip:** Kostenfunktion  $\epsilon^T \cdot \epsilon$  definieren und minimieren (Variante: Gewichtete Kostenfunktion)

mit  $\epsilon = \text{Messwert} - \text{Modell}$

$$y = \phi \cdot p$$

Mit Ableitung ergibt sich Lösung :

$$p = (\phi^T \cdot \phi)^{-1} \cdot \phi^T \cdot y = \phi^+ \cdot y$$

Singulärwertzerlegung (SVD) kann für einfache Berechnung von  $\phi^+$  genutzt werden:

$$\phi = U \cdot \Sigma \cdot V^T \rightarrow \phi^+ = V \cdot \Sigma^T \cdot U^T$$

### ParameterNICHTlineare Modelle

Ansatz wie bei parameterlinearen Modellen.

**Problem:** nichtlineare Gleichungen

**Lösung:** Linearisierung der Fehlergleichung

Verfahren:

- Gauß-Newton-Verfahren (geg. mit Dämpfungsfaktor)
- Gradientenverfahren (line search)
- Levenberg-Marquardt-Algorithmus

## MkQ für Dynamische Systeme

### Dynamisch zeitdiskrete Systeme

Dynamische Modelle = ARX (autoregressive) Modelle

### Dynamisch zeitkontinuierliche Systeme

Ausgangspunkt: DGL

Problem: Ableitungen beschaffen

Lösung:

- a) Finite Differenzen (Vorwärts/Rückwärtsdifferenz) Störanfällig, Messrauschen wird verstärkt

b) Filterung von Ein- und Ausgangssignalen

Idee: Ausnutzen von Eigenschaften des Faltungsoperators

$$d/dt(x(t) * g(t)) = x(t) * d/dt(g(t)) \quad (g(t): \text{Impulsantwort})$$

Zustandsvariablenfilter:

Ansatz:

$$F(s) = \frac{f_0}{f_0 + f_1 s + \dots + s^n}$$

Adaptives Zustandsvariablenfilter:

z.B. Butterworth-Filter

**Wann sind physikalische Parameter vollständig identifizierbar?**

- $np = n + m + 1$
- Jacobi-Matrix  $\delta f / \delta p$  ist regulär

## Rekursive MkQ

Iterationsvorschrift -> siehe Skript

**Bestimmung der Startwerte**

**Nutzung der nicht-rekursiven MkQ**

**Wahl von Standardwerten**

Startwertwahl von  $a_0 = 0$   $P_0 = 1/\alpha I$  ( $I$ : Einheitsmatrix). Dies führt für große  $\alpha$  zu  $P_k \approx \phi_k^T \phi_k$

**Rekursive MkQ mit exponentiell nachlassendem Gedächtnis**

**Rechentechnische Umsetzung der MkQ**

TODO

## Identifikation nicht-linearer Systeme

**Hammerstein-Modell:** nicht-linear statisches System + dynamisch lineares System

Einfacher Ansatz für nicht-Linearität:  $\tilde{u}[k] = r_0 + r_1 \cdot u[k] + \dots r_p \cdot u[k]^p$

Ergibt lineares Modell mit mehreren Eingängen, darstellbar in der Form  $y[k] = \phi a$

## 2.7 Modifikationen der MkQ

### 2.7.1 Totale MkQ (orthogonale Regression)

Minimierung des Fehlers der Ausgangsdaten  $F$  und des Fehlers der Eingangsdaten  $\epsilon$ :

$$y + \epsilon = (\Phi + F)a$$

$$\Rightarrow [(\Phi y) + (F\epsilon)] \begin{pmatrix} a \\ -1 \end{pmatrix}$$

$\Rightarrow$  Minimierung von  $(F\epsilon)$  im Sinne der Frobeniusnorm.

### Einschub: Singulärwertzerlegung

Mann kann Matrizen unter gewissen Voraussetzungen folgendermaßen zerlegen:

$$C = U\Sigma V^T$$

TODO

### 2.7.2 Methode der Hilfsvariablen

Anwendung: bei verzerrten Schätzern (ARX-Modell nicht perfekt)

Prinzip: Multiplikation der Modellfehlergleichung mit sog. Hilfsvariablen:  
 $W^T \epsilon = W^T y - W^T \Phi a$

$W$  ist so wählen, dass Spalten unkorreliert mit  $\epsilon$  sind.

$\Rightarrow$  Lösung der modifizierten Normalengleichung:  $a = (W^T \Phi)^{-1} W^T y$

Wahl von  $W$ :

1. Schätzen von Parametervektor mit MkQ:  $\hat{a} = \Phi^+ y$
2. Simulation des Modells  $\hat{y} = \Phi \hat{a}$
3.  $W = \dots$  (siehe Skript)
4. Schätzung mittels Hilfsvariablen

Iteratives Wiederholen von 2-4 beseitigt Bias von MkQ Schätzer



### 3. Subspace-based State-Space System Identification (4 SID)

#### 3.1 Grundgleichungen, Zustandsraummodelle

Zustandsraummodell:

$x[k+1] = Ax[k] + Bu[k]$  - Folgezustand abhängig von aktuellem Zustand + Eingang

$y[k] = Cx[k] + Du[k]$  - Ausgang abhängig von Zustand ü Eingang

Bekannt: N Messdatenpaare  $u[k], y[k]$

Problem: Weder Zustandsfolge  $x[k]$  noch Zustandsdimension  $n$  bekannt

$$\begin{pmatrix} y[0] \\ y[1] \\ \vdots \\ y[k-1] \end{pmatrix} = \underbrace{\begin{pmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{k-1} \end{pmatrix}}_{\substack{Q_{B,k} \\ \text{Beobachtbarkeitsmatrix}}} x[0] + \underbrace{\begin{pmatrix} D & 0 & \dots & 0 \\ CB & D & \dots & 0 \\ \dots & \dots & D & 0 \\ CA^{k-2}B & CA^{k-3}B & CB & D \end{pmatrix}}_{H_k} \begin{pmatrix} u[0] \\ u[1] \\ \vdots \\ u[k-1] \end{pmatrix}$$

Zusammenfassung in Blockmatrizen:

$$Y = \begin{pmatrix} y[0] & \dots & y[N-2k] \\ \vdots & \dots & \vdots \\ y[k-1] & \dots & y[N-k-1] \\ \vdots & \vdots & \vdots \\ y[k] & \dots & y[N-k] \\ \vdots & \vdots & \vdots \\ y[2k-1] & \dots & y[N-1] \end{pmatrix} = \begin{pmatrix} Y_f \\ Y_p \end{pmatrix}$$

Analog für:  $U = \begin{pmatrix} U_f \\ U_p \end{pmatrix}$

#### Subspace-Gleichungen

$$Y_p = Q_{B,k}X_p + H_K U_p$$

$$Y_f = Q_{B,k}X_f + H_K U_f$$

$$X_f = A^k x_p + Q_{S,k} U_p$$

mit  $Q_{S,k} = (A^{k-1}B \ A^{k-2}B \ \dots \ AB \ B)$  (erweiterte Steuerbarkeitsmatrix)

Durch Umformen/Einsetzen ergibt sich:

$X_f = \dots$  (nur abhängig von Vergangenheit)

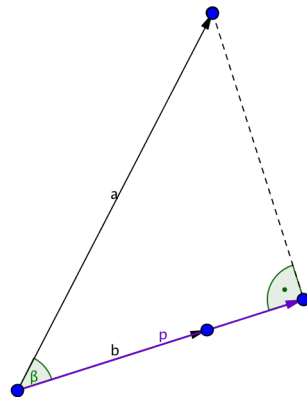
$$Y_f = Q_{B-k} L_{P,k} (U_p y_p)^T + H_k U_f$$

=> Für nächsten Ausgang Wissen der zukünftigen Eingabe erforderlich

## 3.2 Grundlagen: Projektion

### 3.2.1 Orthogonale Projektion

(0.5, 4.75)



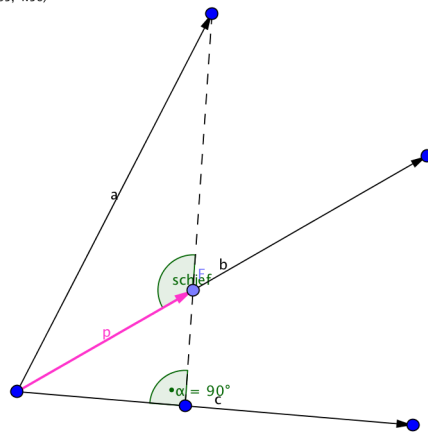
(5.58, 0.58)

Im zweidimensionalen lässt sich der Projektor  $p$  folgendermaßen bestimmen:

$$p = a \cos \alpha \frac{|a|}{|b|} = a \frac{b^T b}{b b^T}$$

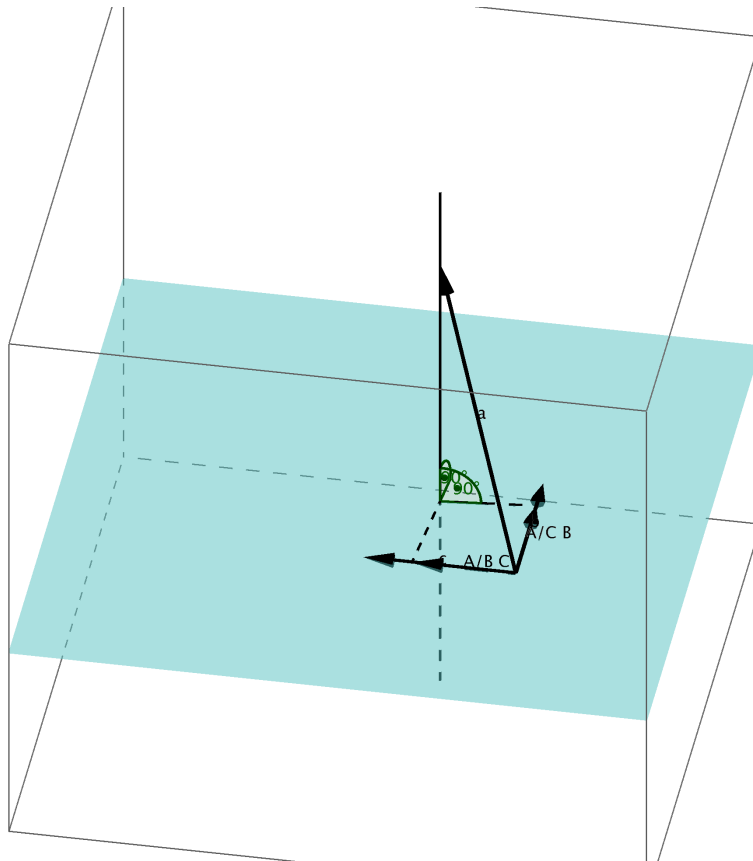
### 3.2.2 Schiefe Projektion

(0.65, 4.56)



(5.63, 0.47)

## Allgemeine schiefe Projektion



Vorgehen:

- Senkrechte Projektion in die von B und C aufgespannte Ebene
- Schiefe Projektion der Ebene

Definition: Allgemeine schiefe Projektion entlang des Orthogonalkomplements von C auf B:

$$A/_CB := A \cdot \Pi$$

mit:

$$\Pi = (B^T \quad C^T) \begin{pmatrix} BB^T & BC^T \\ CB^T & CC^T \end{pmatrix}^+ \begin{pmatrix} B \\ 0 \end{pmatrix}$$

#### Ablaufschema 4 SID

- Messdaten  $u[i]$ ,  $y[i]$  aufnehmen, in Hankelmatrizen  $U$ ,  $Y$  anordnen
- Schiefen Prädiktor  $P$  berechnen
- SVD von  $P$  (Schätzung der Systemordnung (Länge Zustandsvektor); Schätzung für Beobachtbarkeitsmatrix  $Q_{B,k}$ )
- Berechnen von  $A$ ,  $C$
- Berechnen von  $B$ ,  $D$

## 4. Kalman-Filter

Ausgangspunkt: System mit Zustandsraummodell

$$\begin{aligned}x[k+1] &= Ax[k] + Bu[k] + v[k] \\ y[k] &= Cx[k] + e[k]\end{aligned}$$

### Annahmen

- Systemmatrizen A,B,C bekannt
- Eingang u[k], Ausgang y[k] bekannt
- Systemrauschen v[k] und Messrauschen e[k]: unkorrelierte, mittelwertfreie Rauschprozesse

### Ziel

Schätzung  $\hat{x}[k]$  des Zustandsvektors  $x[k]$

### Ansatz: Filterstruktur

$$\hat{x}[k+1] = \underbrace{A\hat{x}[k] + Bu[k]}_{\substack{\text{Prädiktionsterm} \\ \text{"a priori Schätzung"}}} + \underbrace{K[k]}_{\substack{\text{Kalman-Matrix} \\ \text{"Korrekturmatrix"}}} \underbrace{(y[k] - \hat{x}[k])}_{\text{Korrekturterm}}$$

- Dieser Schätzer ist besonders gut, da erwartungstreu
- Kalman-Matrix so wählen, dass Kovarianz  $P[k]$  von Schätzfehler  $\tilde{x}[k+1] = \hat{x}[k+1] - x[k+1]$  minimiert wird

$$\begin{aligned}K[k] &= P[k]C^T(Y + CP[k]C^T)^{-1} \\ P[k+1] &= AP[k]A^T + V - K[k]CP[k]\end{aligned}$$

- V: Kovarianz des Systemrauschens
- Y: Kovarianz des Messrauschens (frei wählbar)

**Ergebnis:** erwartungstreuer Schätzer mit kleinster Varianz

### Filteralgorithmus

1. Init:  $\hat{x}[0]$  &  $P[0]$  (Anfangszustand aus phys. Vorwissen wählen oder 0 setzen)
2. Prädiktion/Zeitupdate: Schätzung des Zustands auf Basis der Messwerte bis Zeitpunkt k

$$\hat{x}^-[k+1] = A\hat{x}[k] + Bu[k]$$

3. Korrektur/Messupdate: Berechnung der neuen Kalman-Matrix, Korrektur der Zustandsschätzung anhand des neuen Messwertes  $y[k+1] \rightarrow$  a posteriori Schätzung
- Schritte 2. und 3. iterativ für alle Messwerte wiederholen bis Schätzung des internen Zustands konvergiert

## Kalman-Filter als Parameterschätzer

# 5. Identifikation nichtparametrischer Modelle

## 5.1 Frequenzgang mit period. Anregung

### Variante 1: Anregung mit harmon. Eingangssignal (Sinus)

- Nach Einschwingen: Amplitude + Phase messen
- Wiederholung für versch. Frequenzen

Problem: reine Sinusschwingungen schwierig zu erzeugen

### Variante 2: Anregung mit Trapez- oder Rechtecksignal

- Beginn bei hohen Frequenzen
- Bei kleineren Anregungsfrequenzen: Berücksichtigung der höheren harmonischen notwendig
- Aus vorherigen Messungen ist Übertragungsverhalten für hohe Frequenzen bekannt  $\rightarrow$  Signalanteile können subtrahiert werden; somit wird Grundschwingung isoliert

Nachteil: zeitaufwendig, da warten auf Einschwingen

Ausweg: Signale mit mehreren Frequenzanteilen

Allgemeine Nachteile:

- nur stabile Systeme
- keine passive Messungen
- nur kleine Störsignale

## 5.2 Korrelationsanalyse

### 5.2.1 Schätzung der Korrelationsfunktion

Def. Kreuzkorrelationsfolge:  $R_{uy}[j] = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=0}^{N-1} u[k-j]y[k]$

Def. Autokorrelationsfolge:  $R_u[j] = R_{uu}[j] = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=0}^{N-1} u[k-j]u[k]$

Eigenschaften:

- $R_{uy}[j] = R_{yu}[-j]$
- $R_u[j] = R_u[-j]$

Problem: Messung nur über endlichen Zeithorizont => Schätzung

Schätzung der Autokorrelationsfolge:  $\hat{R}_u[j] = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=0}^{N-|j|-1} u[k-|j|]u[k]$

Bemerkung: Schätzung  $\hat{R}_u[j]$  ist nicht erwartungstreu:  $E\{\hat{R}_u[j]\} = (1 - \frac{|j|}{N})R_u[j]$

Andere Möglichkeit:  $\hat{R}'_u[j] = \lim_{N \rightarrow \infty} \frac{1}{N-|j|} \sum_{k=0}^{N-|j|-1} u[k-|j|]u[k]$

Dieser Schätzer ist erwartungstreu, weist aber eine um den Faktor  $N/(N-|j|)$  größere Varianz auf -> Praktisch wird ersterer verwendet

### 5.2.2 Schätzung der Gewichtsfolge

Für lineare zeitdiskrete Systeme sind Ein- und Ausgangssignal mittels Faltungssumme verknüpft:

$$y[k] = g * u = \sum_{l=0}^{\infty} g[k-l]u[l] = \sum_{l=0}^{\infty} g[l]u[k-l]$$

Für Kreuzkorrelationsfolge  $R_{uy}[j]$  gilt:

$$R_{uy}[j] = \dots (\text{siehe Skript}) = g * R_u = \sum_{l=0}^{\infty} g[l]R_u[j-l]$$

Annahme:  $R_{uy}[j]$  und  $R_u[j]$  bekannt für  $j$  mit  $-P \leq j \leq M$

=> Gleichungssystem aufstellen (siehe Skript)

Lösung mittels MkQ liefert:

$$\hat{g} = \hat{R}_u^+ \hat{R}_{uy}$$

Falls Eingangssignal weißes Rauschen mit Autokorrelationsfunktion

...

Dann folgt:

$$\hat{R}_{uy} = \hat{\sigma}^2 g[i]$$

=> Schätzung für Gewichtsfolge:  $\hat{g}[i] = \frac{1}{\hat{\sigma}^2} \hat{R}_{uy}$

## **Parameterschätzung für Differenzengleichungen**



## 6. Statistische Parameteridentifikation

### 6.1 Maximum-Likelihood-Methode

#### 6.1.1 Grundgedanke

**bisher:** keine Annahme für Verteilungsfunktion der betrachteten Fehlersignale

**jetzt:** stochastische (meist normalverteilte) Fehlersignale mit bekannter Verteilung

Sei  $X$  eine Zufallsvariable mit Wahrscheinlichkeitsdichte  $p(x, a)$  abhängig von Parameter  $a$ . Wahrscheinlichkeit für Auftreten der Stichprobe  $x_1, \dots, x_N$ :

$$L(a) = P(x_1, a) \cdot P(x_2, a) \cdot \dots \cdot P(x_N, a)$$

Diese Funktion wird als **Likelihood-Funktion** bezeichnet.

**Maximum-Likelihood-Schätzung:**  $\hat{a} = \underset{a}{\operatorname{argmax}} L$

Einfacher zu analysieren: **Log-Likelihood-Funktion**

$$\ln(L(a)) = \ln(P(x_1, a)) + \ln(P(x_2, a)) + \dots + \ln(P(x_N, a))$$

Ableitung liefert notwendige Bedingung für Maximum:

$$\frac{\partial(L)}{\partial a} = \sum_{i=1}^N \frac{\partial \ln p(x_i, a)}{\partial a} \stackrel{!}{=} 0$$

#### 6.1.2 Maximum-Likelihood-Schätzer für statische Systeme

- statischer Prozess mit Eingang  $u$ , Ausgang  $y$
- Likelihood-Funktion als bedingte Wahrscheinlichkeit  $p(y|u, a)$

TODO

## Zusammenfassende Fragen

- Ansatz Herleitung der MkQ; Was wird Minimiert?
- Gegeben: Messdaten Eingang + Ausgang
  - Welches Vorgehen?
  - Welcher Ansatz setzt was voraus?
  - linear  $\leftrightarrow$  nicht-linear unterscheiden
  - parameterlinear  $\leftrightarrow$  nicht-parameterlinear unterscheiden
- statisch  $\leftrightarrow$  dynamische Systeme
- Umrechnung kontinuierliche  $\leftrightarrow$  diskrete Systeme
  - Warum notwendig?
  - Wie zurück rechnen?
- 4SID nur knapp
  - Zustandsraum
  - Vorteile (Mehrgrößen und Dimension)
  - Nachteil (Verzerrter Schätzer wie MkQ; Zusammenhang physikalische Systeme, schlechte Abbildung)
- Kalman-Filter (Struktur, nicht unbedingt Details)
- Rekursive MkQ
  - keine Herleitung
  - Was bringt diese Methode?
  - Matrix-Inversion gespart
  - Rechenaufwand gering
  - online implementierbar
  - Wieviele/ welche Daten müssen gespeichert werden? Vergleich mit MkQ
- Nicht-lineare Systeme (Hammerstein-Modell)
- Frequenzgang-Messung (aufwendig, zeitintensiv, nicht-parametrisches Modell nicht besonders nützlich)
- Erzeugung Pseudo-Rausch-Binär-Signal

## Weniger relevant

- Matrixzerlegung
- Projektion
- Methode der Hilfsvariablen; totale MkQ (jeweils nur Grundkonzept wissen)