Einführung in Computational Engineering



Grundlagen der Modellierung und Simulation

12. Vorlesung: System Identifikation

20. Januar 2013

Prof. Dr. Jan Peters





Meisenantworten



- Ist MATLAB prüfungsrelevant? Ja, natürlich! Aber maximal wird ein Code-Stück zur Vervollständigung als Aufgabe vorkommen...
- Werden Sie eine Probeklausur hochladen? Hochladen? Nein! Vorrechnen? Ja, in einer der letzten Vorlesungen!
- Lehreevaluationsfragenbögen werden in der Pause verteilt: Bitte ausfüllen und einsammeln!





Grundlagen der Modellierung und Simulation

7. PARAMETERIDENTIFIKATION VON MODELLEN



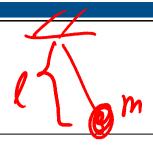
Heutige Lernziele: Kernfragen



- Was ist Systemidentifikation, welches Ziel hat SI und warum ist SI wichtig?
- Welche Probleme macht SI?
- Was unterscheidet SI von der Modellierung?
- Wie geht man bei SI bei einem praktischen Problem vor?
- Was bedeuten Strukturvalidierung und Parameteridentifikation?
- Was sind White-, Grey and Black Box Indentifikation?
- Wie funktioniert lineare Regression? Was sind Basisfunktionen und Parameter?
- Was ist Regularisierung und was hat Ridge-Regression damit zu tun?
- Was ist nichtlineare Regression und warum braucht man sie?



7.1 Systemidentifikation: Einführung





Automatische

Das Ziel ist die Bestimmung der Paramete θ des Modells von dem System

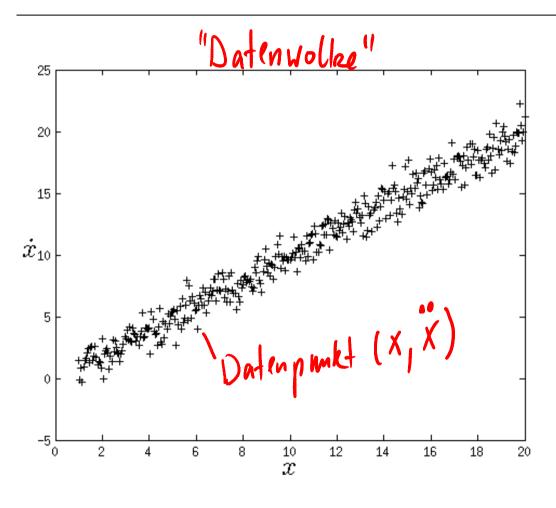
$$\dot{\mathbf{x}} = f(\mathbf{x}, \theta)$$
 Dynamischer System als Modell!

- Nutzen
 - Es können akkuratere Vorhersagen mit der Simulation vorgenommen werden
 - Modellannahmen der Simulation k\u00f6nnen validiert werden



7.1 Systemidentifikation: Beispiel



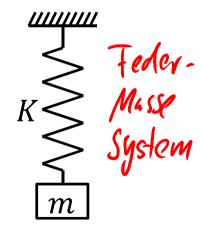


Gleichung

$$\ddot{x} = -\frac{K}{m}x$$

$$\Leftrightarrow \dot{\mathbf{x}} = \begin{bmatrix} \ddot{x} \\ \dot{x} \end{bmatrix} = f(x, \theta) = \begin{bmatrix} -\theta x \\ \dot{x} \end{bmatrix}$$

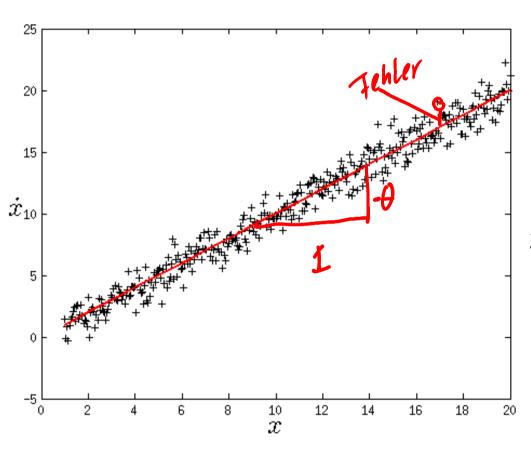
• Wert von θ ?





7.1 Systemidentifikation: Beispiel





- Lineare Struktur: $\ddot{x} = -\theta x$ mit Parameter θ
- Wie kann man die Modellstruktur validieren?
- 2. Wie kann man die Parameter schätzen?



7.1 Systemidentifikation: Probleme



- Reale Systeme sind oft weit von dem mathematischen Modell entfernt
- Parameter sind immer Schätzwerte
- Messwerte sind immer verrauscht

$$\hat{x}_i = \dot{x}_i + \varepsilon_{\dot{x}_i}$$
 bzw. $\hat{x}_i = x_i + \varepsilon_{x_i}$

Messwerte können strukturelle Fehler enthalten, d.h.

$$E\left[\varepsilon_{\chi_i}\right] \neq 0$$

und dann nennt man sie verzerrt (engl. biased)

Verzermy wird

mer midd

tiefer behanded,



7.1 Systemidentifikation: Kontrast



Mnsq abstrakts Modellierung Verständmis	Abyluck des Modelles Mit der Leulität Systemidentifikation
 Bestimmung von Zustands- und Störgrößen 	Basiert direkt auf Experimenten
 Mathematische Beschreibung durch Gleichungen 	 Eingaben und Ausgaben des Systems werden aufgenommen
 System wird nicht genutzt, d.h. keine "reale" Daten 	



7.1 Systemidentifikation: Beispiel aus der Robotik





7.1 Bedeutendste überschneidende Forschungsgebiete



Statistik

- Ökonometrie Machine Leaning: Stedishide Mekada 1 (SoSe)
 Machine Learning ML: Symbolische Methada (Wise)
 Zeitreihenanalyse Robot Learning (Wise)
- Data Mining
- Neuronale Netze
- Chemometrik



7.1 Systemidentifikation: Einführung



- Hauptschritte der Systemidentifikation
 - Strukturidentifikation / Validierung (Stimmt f mit unseren gesammelten Daten überein?)
 - Parameteridentifikation (Stimmt $\underline{\theta}$ mit unseren gesammelten Daten überein?)



7.1 Systemidentifikation:



1. Strukturvalidierung

- Validierung von $\dot{\mathbf{x}} = f(x, \theta)$
- Zielsetzung: Sicherstellen, dass das Modell zu dem jeweiligen
 Zeitschritt t passt

Parameteridentifikation

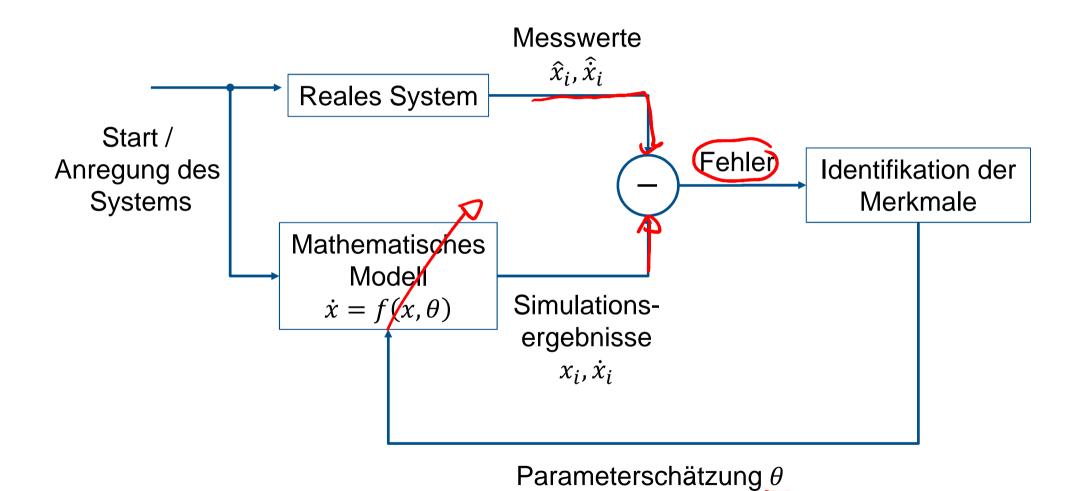
 Struktur ist vorgegeben und Modellparameter θ soll bestimmt werden

 $\hbox{$\stackrel{\bullet}{$}$ Zielsetzung: Finde ein θ so dass $\widehat{\hat{x}}_i = f(\widehat{x}_i, \theta)$ für $$ Messwerte $(\widehat{\hat{x}}_i, \widehat{x})$ stimmt $$ Möjlichst $$ sum (ϕ) and $2ieven. $$$



7.2 Grundidee der Parameteridentifikation

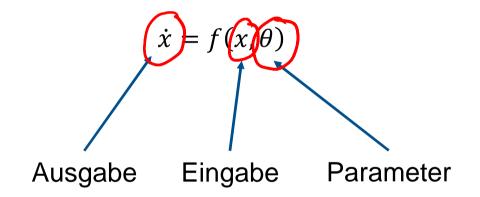








- Schritt 1: Spezifikation
 - Spezifiziere und parametrisiere die Modellklasse









- Schritt 2: Datenerhebung
 - Sammele Datenpaare bestehend aus Eingaben ("sample") und Ausgaben (Datensätze)

$$D = \{(\hat{x_i}, \hat{x_i}) | i \in 1, 2, ..., m\}$$
Anyeben Eingeben

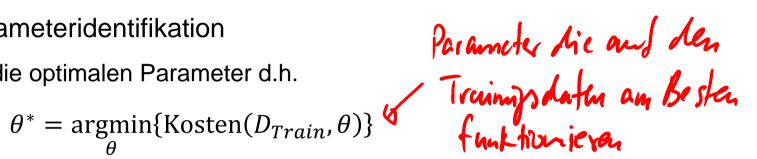
- Unterteile Daten zufällig in
 - Trainingsdatensatz $D_{Train} = randomSelection(D)$
 - Testdatensatz $D_{Test} = D \setminus D_{Train}$





- Schritt 3: Parameteridentifikation
 - Bestimme die optimalen Parameter d.h.

$$\theta^* = \underset{\theta}{\operatorname{argmin}} \{ \operatorname{Kosten}(D_{Train}, \theta) \}$$



- Schritt 4: Strukturvalidierung
 - Überprüfe ob

$$Kosten(D_{Train}, \theta) > Kostenspezifikation$$

Wenn das der Fall ist, dann gehe zurück zu Schritt 1





- Schritt 5: Parametervalidierung
 - Bestimme ob

$$Kosten(D_{Test}, \theta) \leq Kostenspezifikation$$

- Wenn die Ungleichung <u>nicht erfüllt</u> ist, dann gehe zurück zu Schritt 3
- Wenn die Ungleich <u>erfüllt</u> ist, übernimmt man die Parameter



MOODLE FRAGE

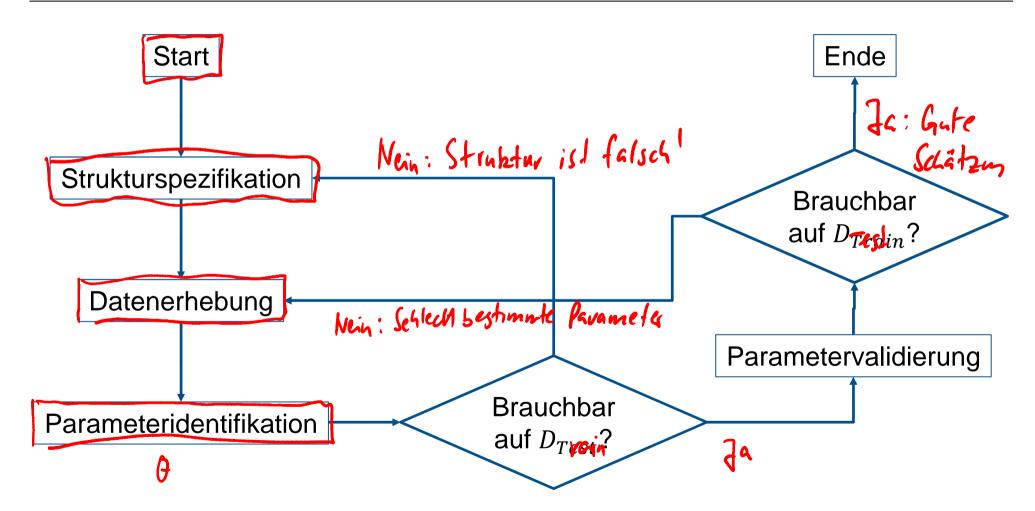


Bitte jetzt auf Moodle Fragen beantworten!



7.2 Vorgehen bei der Systemidentifikation: Überblick







7.3 Arten der Systemidentifikation



- White-Box Identifikation
 - Festes physikalisches Modell
 - Parameter aus den Daten geschätzt
 - Beispiel: Feder-Dämpfer System
- Gray-Box Identifikation
 - Generische, allgemeine Modellklasse die allgemeine Funktionen repräsentieren
 - Parameter bestimmen die Struktur
 - Beispiel: Neuronale Netze



7.3 Arten der Systemidentifikation



- Black-Box Identifikation
 - Modellstruktur und Modellparameter werden aus den Daten geschätzt
 - Ist Bereich der aktuellen Forschung und ist eher selten im Alltag vorzufinden
 - Beispiel: "security pricing" Modelle für den Aktienmarkt; sogenannte Derivate



7.3.1 Systemidentifikation: Einsatzgebiete in der Industrie



- Regelungstechnik und Prozesssteuerung
 - Hochentwickelt
 - Alle Regelstrecken und Prozesse sind verrauscht und verschieden
 - Enormer Kostendruck
 - Es existieren in der Praxis genutzte Toolboxen, z.B. in MATLAB
- Luft- und Raumfahrt
 - White-Box, spezielle Testverfahren



7.3.1 Systemidentifikation: Einsatzgebiete in der Industrie



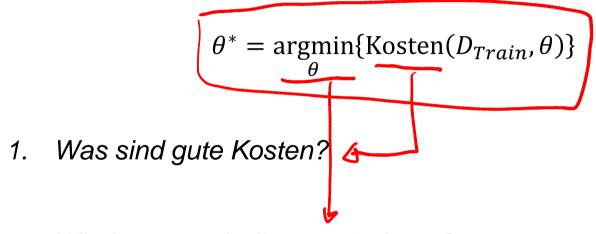
- Auto
 - White-Box, Nutzung zur Kalibrierung, enorme Leistungen zur Validierung
 - Gray-Box für "autonome Automobile"
- Festplatten
 - Geringe cycle-time, Lebenszeitverlängerung
- Embedded Systems
 - short cycle-time



7.4 Parameteridentifikation im Detail



Parameteridentifikation



2. Wie können wir diese optimieren?

Diese zwei Fragen sind *nie* unabhängig voneinander!



7.4.1 Grundidee; Kosten



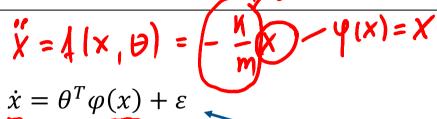
- Fehler in Messwerten $e_i = \hat{x}_i f(\hat{x}_i, \theta)$ soll minimiert werden
- Absoluter Fehler: Kosten = $\sum_{i=1}^{m} ||e_i||$
 - Dünnbesetzte Lösungen
 - Mathematisch aufwendig
- Quadratischer Fehler: Kosten = $\sum_{i=1}^{m} ||e_i||^2$
 - Analytische Lösung und daher einfach zu bestimmen
 - Ist Robust aber hat einen kubischen Datenbedarf

(Mehr in Ljang, L. (1999). System Identification: Theory for the user)









$$\dot{x} = \theta^T \varphi(x) + \varepsilon$$

Ausgaben

Parameter

Rauschen

Eingaben

"Features" Basisfunktionen

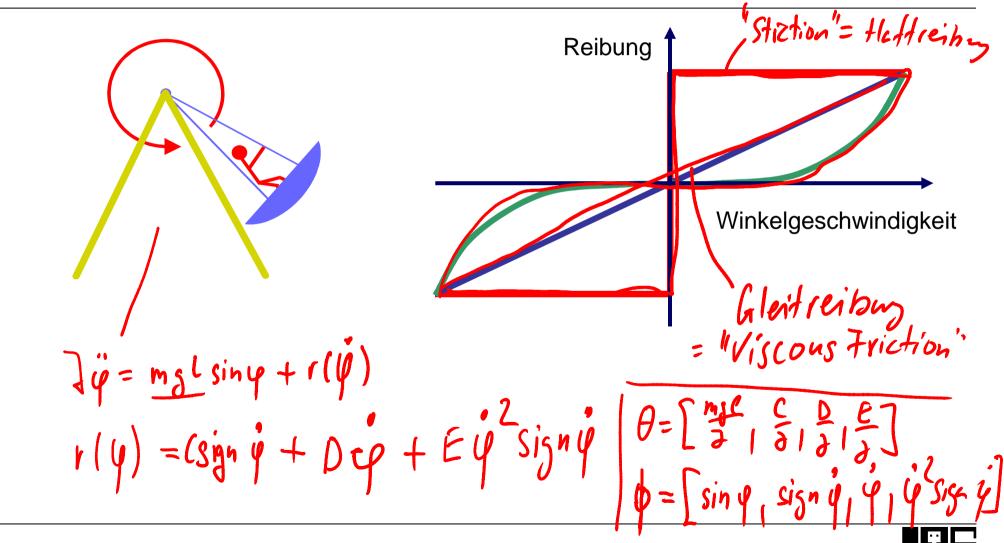
Beispiel:

$$\ddot{x} = -\frac{K}{m} \cdot x \implies \begin{cases} \varphi(x) = x \\ \theta = -\frac{K}{m} \end{cases}$$



7.5 Parameteridentifikation: Basisfunktionen für die Lineare Regression bei der Schiffsschaukel







- Vektorisierung für
 - m Datenpunkte
 - n Parameter

$$\mathbf{y} = \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \vdots \\ \hat{x}_m \end{bmatrix} \quad \mathbf{\Phi} = \begin{bmatrix} \varphi_1(\hat{x}_1) & \cdots & \varphi_n(\hat{x}_1) \\ \varphi_1(\hat{x}_2) & \cdots & \varphi_n(\hat{x}_2) \\ \vdots & \ddots & \vdots \\ \varphi_1(\hat{x}_m) & \cdots & \varphi_n(\hat{x}_m) \end{bmatrix}$$

$$\boldsymbol{\theta} = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_n \end{bmatrix} \qquad \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_m \end{bmatrix} \qquad \text{Madrice mid}$$

• Daher
$$y = \Phi \theta + \varepsilon$$





Das Gleichungssystem

$$y = \Phi\theta + \varepsilon \Rightarrow \theta = \Phi^{-1}(y + \varepsilon)$$

kann nur für $\underline{n} = m$ gelöst werden. Dann wirken sich Fehler aber drastisch aus

■ Der Fall m < n macht keinen Sinn aber was macht man bei m > n?





Man minimiert den Quadratischen Fehler

$$e_i = y_i - \overline{\phi}_i^{\dagger} \theta$$

Kosten
$$(D, \theta) = \frac{1}{2} \sum_{i=1}^{n} \|e_i\|^2 = \frac{1}{2} \sum_{i=1}^{n} e_i^T e_i$$

$$= \frac{1}{2} (y - \Phi \theta)^T (y - \Phi \theta)$$

$$= \frac{1}{2} (y - \Phi \theta)^T (y - \Phi \theta)$$

$$= \frac{1}{2} (y - \Phi \theta)^T (y - \Phi \theta)$$

$$= \frac{1}{2} (y - \Phi \theta)^T (y - \Phi \theta)$$

$$= \frac{1}{2} (y - \Phi \theta)^T (y - \Phi \theta)$$

$$= \frac{1}{2} (y - \Phi \theta)^T (y - \Phi \theta)$$

$$= \frac{1}{2} (y - \Phi \theta)^T (y - \Phi \theta)$$

$$= \frac{1}{2} (y - \Phi \theta)^T (y - \Phi \theta)$$

$$= \frac{1}{2} (y - \Phi \theta)^T (y - \Phi \theta)$$

$$= \frac{1}{2} (y - \Phi \theta)^T (y - \Phi \theta)$$

$$= \frac{1}{2} (y - \Phi \theta)^T (y - \Phi \theta)$$

$$= \frac{1}{2} (y - \Phi \theta)^T (y - \Phi \theta)$$

$$= \frac{1}{2} (y - \Phi \theta)^T (y - \Phi \theta)$$

$$= \frac{1}{2} (y - \Phi \theta)^T (y - \Phi \theta)$$

$$= \frac{1}{2} (y - \Phi \theta)^T (y - \Phi \theta)$$

$$= \frac{1}{2} (y - \Phi \theta)^T (y - \Phi \theta)$$

$$= \frac{1}{2} (y - \Phi \theta)^T (y - \Phi \theta)$$

$$= \frac{1}{2} (y - \Phi \theta)^T (y - \Phi \theta)$$

$$= \frac{1}{2} (y - \Phi \theta)^T (y - \Phi \theta)$$

$$= \frac{1}{2} (y - \Phi \theta)^T (y - \Phi \theta)$$

$$= \frac{1}{2} (y - \Phi \theta)^T (y - \Phi \theta)$$

$$= \frac{1}{2} (y - \Phi \theta)^T (y - \Phi \theta)$$

MOODLE TEST IN ZWEI FOLIEN









Man minimiert den Quadratischen Fehler

Kosten
$$(D, \theta) = \frac{1}{2} \sum_{i=1}^{n} ||e_i||^2 = \frac{1}{2} \sum_{i=1}^{n} e_i^T e_i$$
$$= \frac{1}{2} (y - \Phi \theta)^T (y - \Phi \theta)$$

Daraus folgt

$$\theta = (\Phi^T \Phi)^{-1} \Phi^T y$$



MOODLE FRAGE

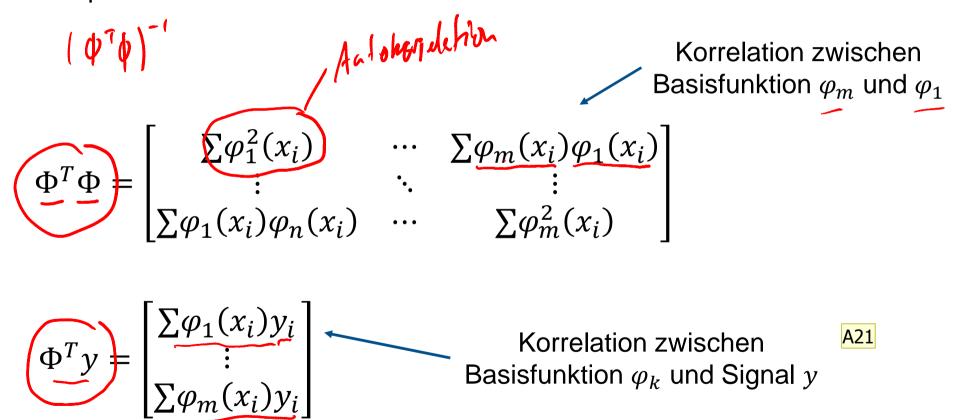


Bitte jetzt auf Moodle Fragen beantworten!





Interpretation als Korrelation





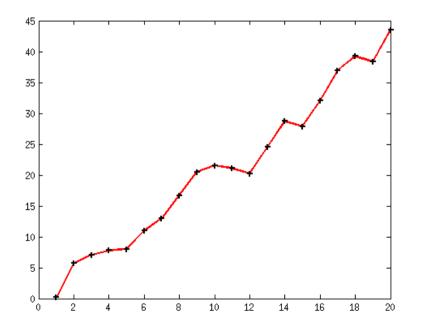
Folie 35

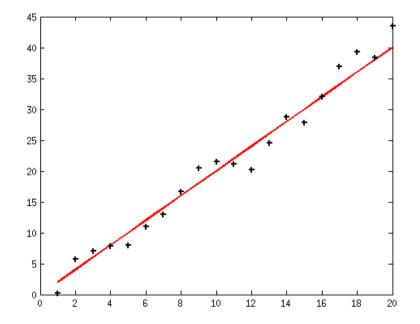
Index auch an y? Andrej; 17.01.2014 A21

7.5 Parameteridentifikation: Lineare Regression



- Angenommen wir haben viele Basisfunktionen und Parameter, d.h. $n \gg I$
- Welche Lösung ist dann besser?



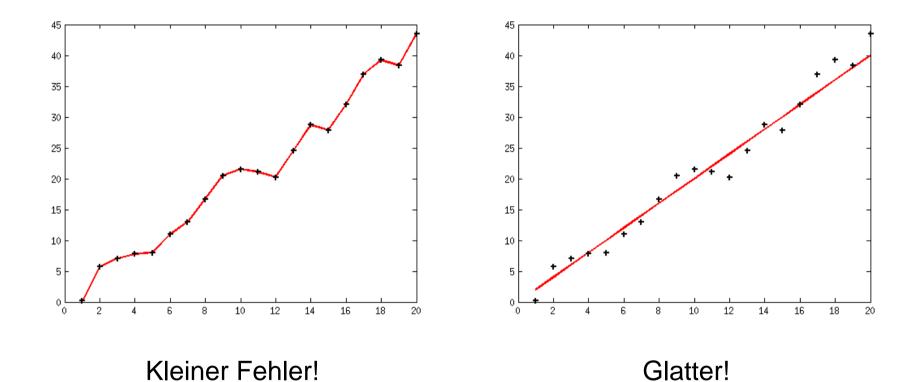




Notizen beachten Andrej; 17.01.2014 A22

7.5 Parameteridentifikation: Lineare Regression





Die bessere Lösung ist nicht die mit dem kleinsten Fehler!



7.6 Regularisierung



Komplexität muss bestraft werden:

Kosten
$$(D, \theta) = \sum_{i=1}^{m} ||e_i||^2 + \text{Komplexität}(\theta)$$

Beispiel für ein Komplexitätsmaß wäre

Komplexität =
$$\frac{1}{2}\theta^T\theta$$

d.h. extrem große Parameter werden bestraft.

• $\lambda \in [10^{-6}, ..., 10^{-8}]$ ist ein "tuning" Parameter



7.6 Regularisierung



Aus

$$Kosten(D, \theta) = \frac{1}{2}(y - \Phi\theta)^{T}(y - \Phi\theta) + \frac{1}{2}\theta^{T}\theta$$

folgt

$$\theta = (\Phi^T \Phi + \lambda I)^{-1} \Phi^T y$$

■ Diese Methode nennt sich Ridge Regression



7.7 Nichtlineare Regression



• Bei dem System $\dot{x} = \theta x$ wissen wir, dass

ssen wir, dass

liner in (

$$x(t_i) = x_i = C \cdot e^{\theta t_i}$$

- Diese Transformation hat unser Modell *nichtlinear* gemacht (in der Zeit nicht im Zustand).
- In der Realität sind alle Prozesse nichtlinear, denn linear ist eine Idealisierung.



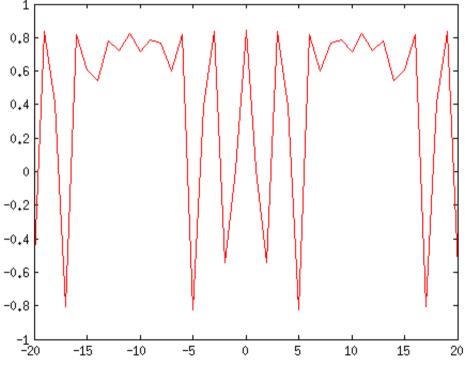
7.7 Nichtlineare Regression



Nichtlineare Regression ist oft schwieriger als lineare
 Regression, besonders wenn es sich um eine nicht konvexe
 Optimierung handelt

Beispiel: Neuronale Netze

 Generell: Immer versuchen nichtlineare auf lineare Regression abzubilden bis auf wenige Parameter





7.7 Nichtlineare Regression: Beispiel



- Gegeben sei $x_i = C \cdot e^{a \cdot t_i} \Leftrightarrow \ln(x_i) = \ln(C) + C \cdot t_i$
- Dabei ist $y_i = \ln(x_i)$, $\varphi_1(x_i) = 1$, $\varphi_2(t_i) = t_i$ linear
- Aus der Lösung gegeben durch

$$\theta = (\Phi^{\mathsf{T}}\Phi + \lambda I)^{-1}\Phi^{\mathsf{T}}y$$

kann dann $C = e^{\theta_1}$ und $a = \theta_2$ bestimmt werden





7.8 Erfolgreiche Validierung der Simulation durch System Identifikation



- Wenn ein Parametersatz θ so berechnet werden kann, dass die Abweichung des Systemverhaltens vom Experiment "klein" ist, dann ist die Stimmigkeit von Simulationsmodell und Parameter sehr hoch
- Zur Kalibrierung der Modellparameter ist es sinnvoll,
 Messwerte von mehreren "repräsentativen" Experimenten zu verwenden



7.8 Misserfolg bei der Parameteridentifikation



- Wenn kein Parametersatz θ gefunden werden kann, für den die Abweichungen "klein" sind:
 - Kostenfunktion oder Optimierungsverfahren ungeeignet
 - Nicht ausreichend relevante experimentellen Daten für die Parameter, um diese daraus eindeutig zu bestimmen
 - Messfehler in den Messdaten zu groß oder Ausreißer in den Messwerten verfälschen die Ergebnisse
 - Das verwendete Modell ist nicht detailliert genug, hat schlechte Basisfunktionen oder enthält nicht alle oder nicht die "richtigen" Effekte





MOODLE FRAGE



Bitte jetzt auf Moodle Fragen beantworten!



Heutige Lernziele: Kernfragen



- Was ist Systemidentifikation, welches Ziel hat SI und warum ist SI wichtig?
- Welche Probleme macht SI?
- Was unterscheidet SI von der Modellierung?
- Wie geht man bei SI bei einem praktischen Problem vor?
- Was bedeuten Strukturvalidierung und Parameteridentifikation?
- Was sind White-, Grey and Black Box Indentifikation?
- Wie funktioniert lineare Regression? Was sind Basisfunktionen und Parameter?
- Was ist Regularisierung und was hat Ridge-Regression damit zu tun?
- Was ist nichtlineare Regression und warum braucht man sie?
- Selbsttest: Können Sie diese Fragen beantworten? Wenn nicht, schnell nochmal das Video anschauen!



Parameteridentifikation und -kalibrierung



"Tuning" der Modellparameter anhand von Messwerten:



experimentelle Messwerte für Zustandstrajektorie (mit Messfehler ε_i) t_i , $j=1,\ldots,n_t$

Optimierungsproblem zur Kalibrierung der Modellparameter



unter der Nebenbedingung, dass

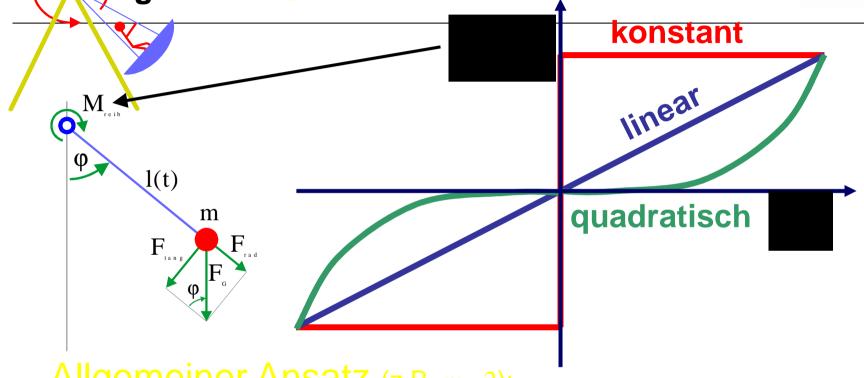


(numerische) Lösung des nichtlinearen Systemmodells



Z.B. Schiffschaukel: verschiedene Reibungsmodelige übliche Varianten:





Allgemeiner Ansatz (z.B. *m* =2):

$$M_{\text{Reib}}(\omega) = \sum_{i=0}^{m} p_i \cdot \text{sgn}(\omega) \cdot |\omega|^i$$



