8. Tag der Mathematik 13. Juli 2013

Statistik: eine Reise von Planetenbahnen zu Data Mining und künstlicher Intelligenz

PD Dr. Robert Hable
Mathematisches Institut
Universität Bayreuth


GN4480100S8
DEUTSCHE BUNDESBANK
Banknote
10
ZEHN DEUTSCHE MARK
1777–1855 Carl Friedr. Gauß
Deutsche Bundesbank
Frankfurt am Main
1. September 1999
GN4480100S8

Ceres:
Zwergplanet im Asteroidengürtel zwischen Mars und Juppiter

Ceres

1. Januar 1801: Ceres erstmals entdeckt von Giuseppe Piazzi

Beobachtung von Ceres ab ca. 12. Februar 1801 wegen der Sonne nicht mehr möglich

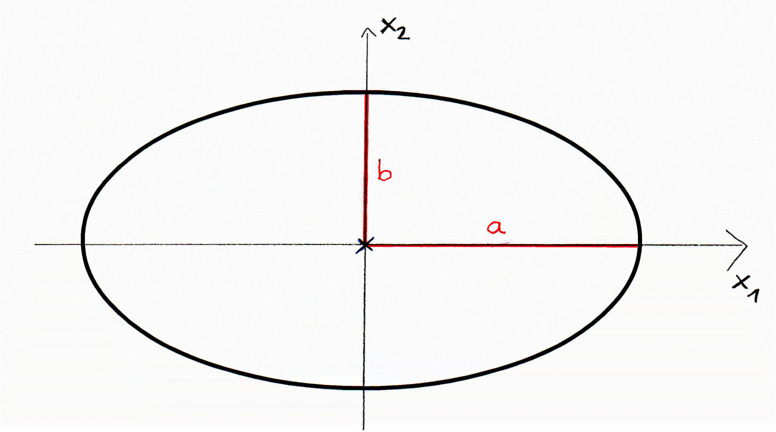
Auch als Ceres eigentlich wieder sichtbar sein müßte, wird er monatelang nicht wiedergefunden

November 1801: Gauß veröffentlicht Artikel mit Berechnung von Ceres Umlaufbahn

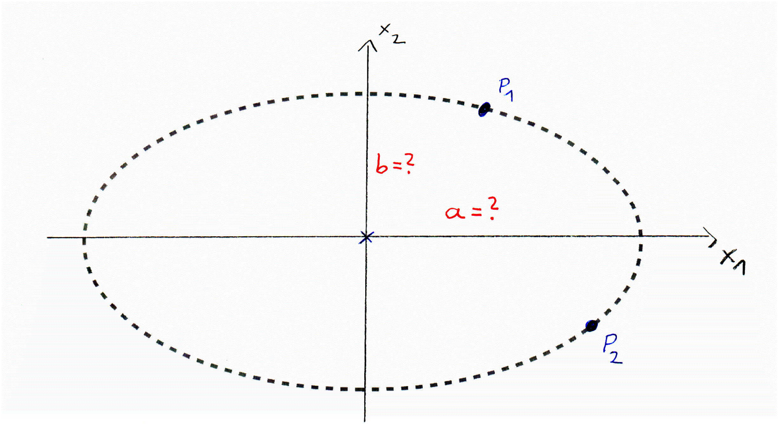
7. Dezember 1801: Franz Xaver von Zach findet Ceres an der von Gauß vorhergesagten Stelle


EUROPA
Deutschland
55
ASTRONOMIE
400 Jahre Keplersche Gesetze
2009

Vereinfachtes Beispiel:



Vereinfachtes Beispiel:



z.B. $P_1 = (1, 2)$ und $P_2 = (3, -1)$

Berechnung von a und b durch P_1 und P_2 :

$P_1 = (1, 2)$

$$1 \stackrel{!}{=} \frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = \frac{1}{a^2} + \frac{4}{b^2} \qquad \text{(I)}$$

$P_2 = (3, -1)$

$$1 \stackrel{!}{=} \frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = \frac{9}{a^2} + \frac{1}{b^2} \qquad \text{(II)}$$

$\longrightarrow$ zwei Gleichungen mit zwei Unbekannten
$\longrightarrow$ Auflösen nach a und b

Berechnung von a und b durch P_1 und P_2 :

(I) $1 = \frac{1}{a^2} + \frac{4}{b^2} \quad \Longrightarrow \quad a^2 = \frac{b^2}{b^2-4}$

(II) $1 = \frac{9}{a^2} + \frac{1}{b^2} = \frac{9b^2 - 36}{b^2} + \frac{1}{b^2}$

Also $\quad b^2 = 9b^2 - 36 + 1$

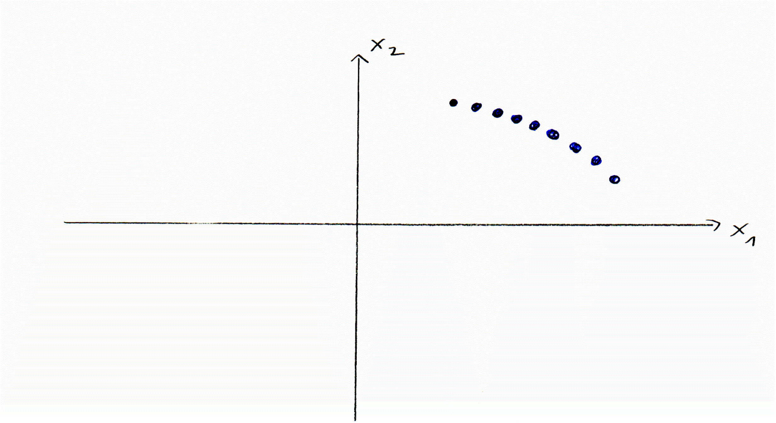
$\Longleftrightarrow \quad 8b^2 = 35$

$\Longleftrightarrow \quad b^2 = \frac{35}{8}$

Und $\quad a^2 = \frac{b^2}{b^2-4} = \ldots = \frac{35}{3}$

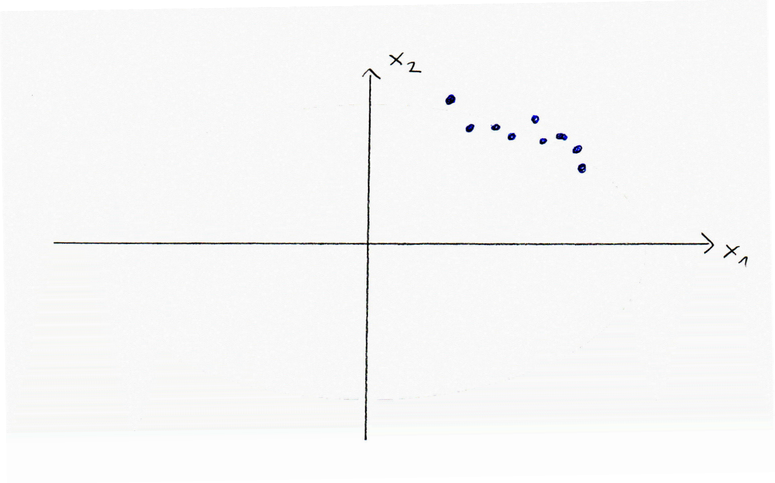
Rekonstruktion der Ellipse

Beobachtungen ohne Messfehler



Rekonstruktion der Ellipse ?

Beobachtungen mit Messfehler



Rekonstruktion der Ellipse ?

Punkte mit Messfehlern: $P_1, P_2, \ldots, P_n$

Idee:

Setze

$$f_{a,b}(P_i) = \frac{{x_{1,i}}^2}{a^2} + \frac{{x_{2,i}}^2}{b^2}$$

Im Idealfall (ohne Messfehler) lägen alle P_i auf einer Ellipse.
Somit gäbe es ein a_0 und ein b_0, so dass

$$f_{a_0,b_0}(P_i) = 1 \qquad \text{für alle Punkte } P_i$$

und somit auch

$$\Big(f_{a_0,b_0}(P_1) - 1\Big)^2 + \ldots + \Big(f_{a_0,b_0}(P_n) - 1\Big)^2 = 0$$

Rekonstruktion der Ellipse ?

Setze

$$R(a, b) = \Big(f_{a,b}(P_1) - 1\Big)^2 + \ldots + \Big(f_{a,b}(P_n) - 1\Big)^2$$

für alle $a > 0$ und $b > 0$.

Die Funktion $(a, b) \mapsto R(a, b)$ hat also folgende Eigenschaften:

- $R(a, b) \geq 0$ für alle $a > 0$ und $b > 0$
- Falls alle P_i auf einer Ellipse lägen, dann gäbe es ein a_0 und ein b_0, so dass

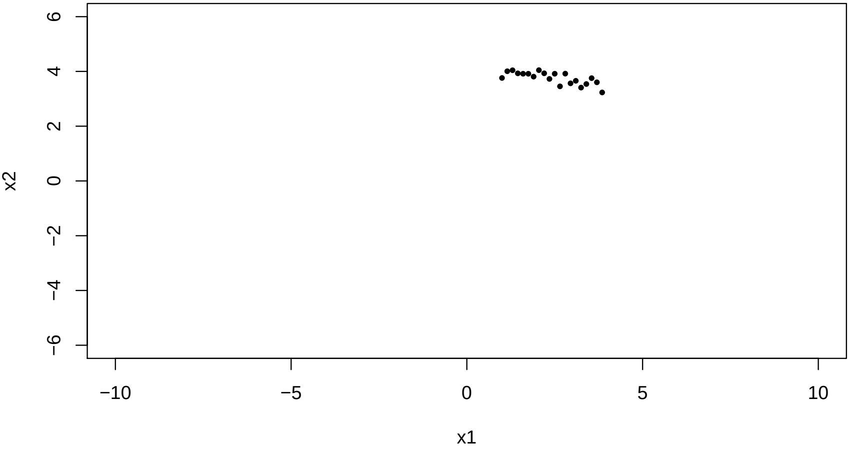
$$R(a_0, b_0) = 0$$

Idee: Wähle meine **Schätzungen** $\hat{a}$ und $\hat{b}$ so, dass

$$R(\hat{a}, \hat{b}) = \text{minimal}$$

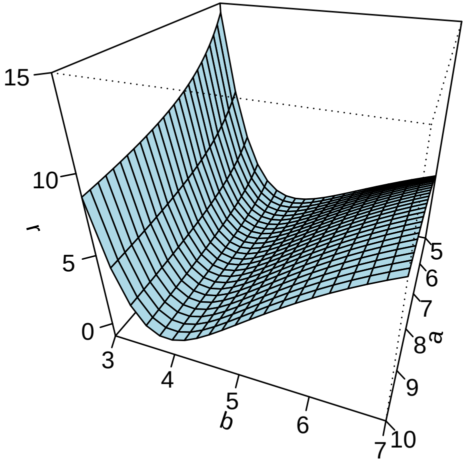
Beispiel:

20 Punkte mit Messfehlern: $P_1, P_2, \ldots, P_{20}$



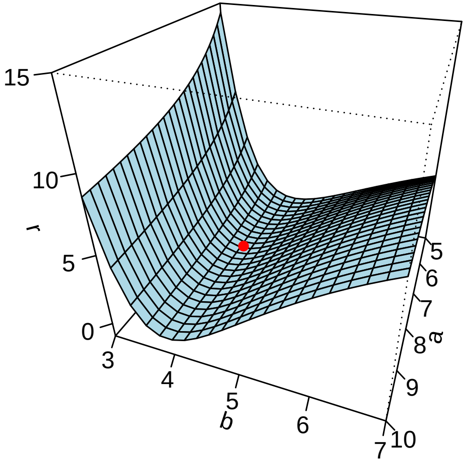
Beispiel:

Bild der Funktion $(a, b) \mapsto R(a, b)$



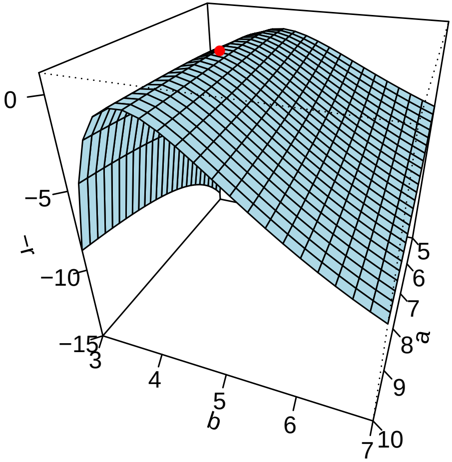
Beispiel:

Bild der Funktion $(a, b) \mapsto R(a, b)$



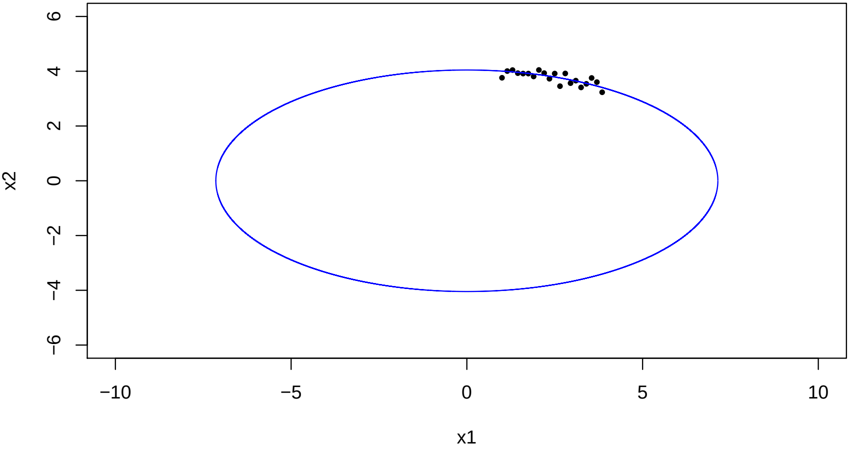
Beispiel:

Bild der Funktion $(a, b) \mapsto -R(a, b)$



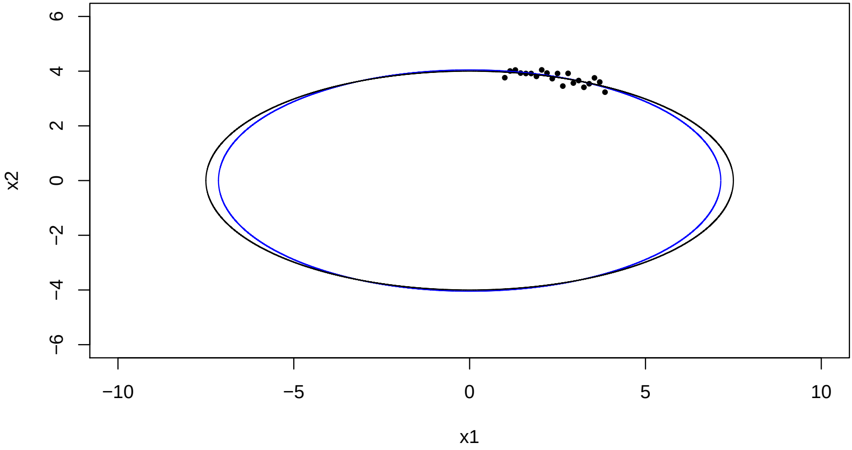
Beispiel:

20 Punkte mit Messfehlern und
die geschätzte Ellipse



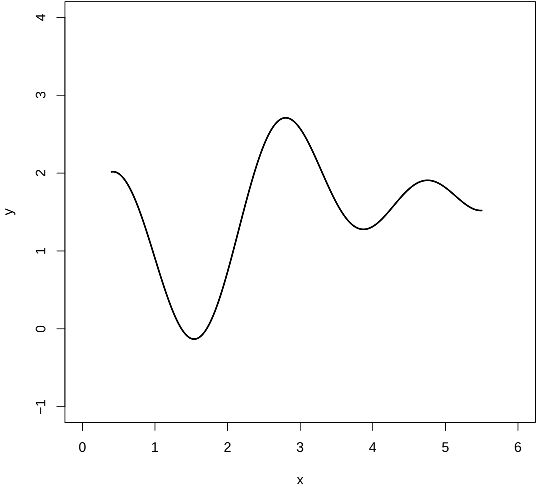
Beispiel:

20 Punkte mit Messfehlern,
die geschätzte und die **wahre** Ellipse



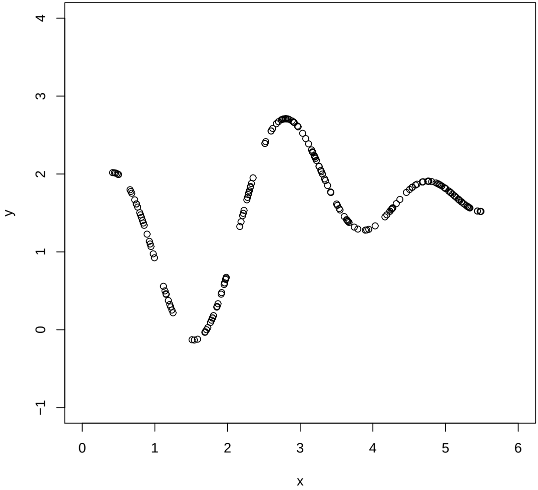
Regression: Zusammenhang $x \rightsquigarrow y$

Funktion



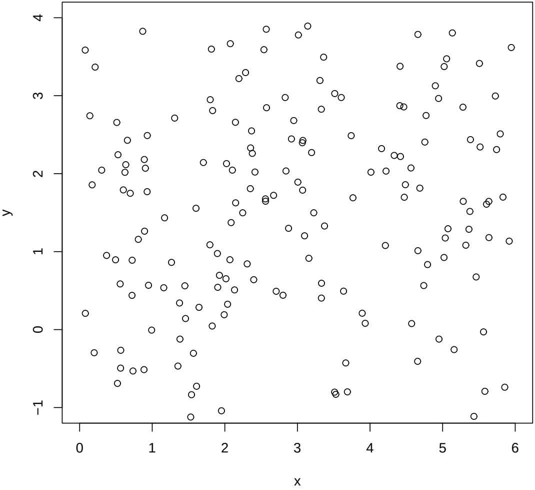
Regression: Zusammenhang $x \rightsquigarrow y$

Daten ohne Streuung



Regression: Zusammenhang $x \rightsquigarrow y$

Daten mit Streuung



Wärmeausdehnungskoeffizient von Eisen

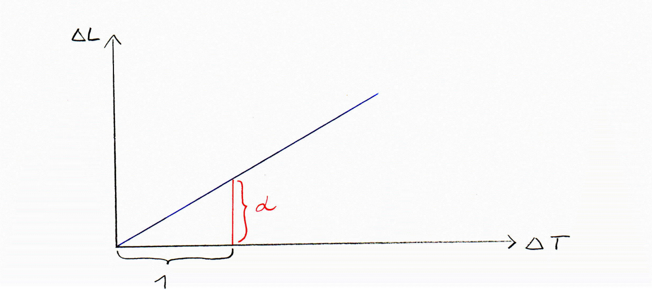
Wärmeausdehnungskoeffizient von Eisen

Eisenstange (1m) bei $20^{o}C$

Bei Temperaturerhöhung dehnt sich Eisenstange aus gemäß

$$\Delta L \;=\; \alpha \cdot \Delta T$$

α : Wärmeausdehnungskoeffizient von Eisen



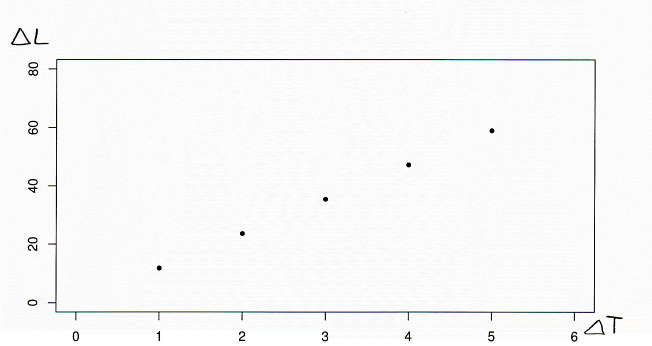
Wärmeausdehnungskoeffizient von Eisen

Eisenstange (1m) bei $20^{\mathbf{o}}C$

Bei Temperaturerhöhung dehnt sich Eisenstange aus gemäß

$$\Delta L \;=\; \alpha \cdot \Delta T$$

α : Wärmeausdehnungskoeffizient von Eisen



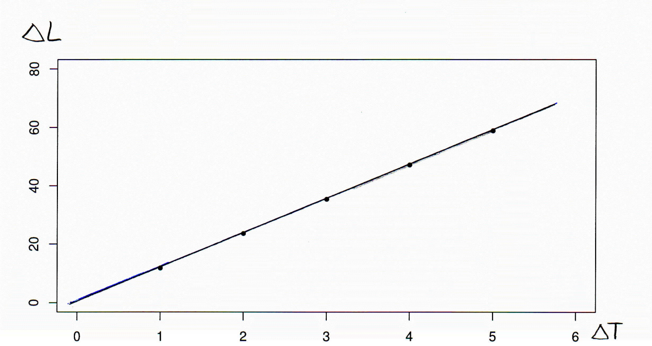
Wärmeausdehnungskoeffizient von Eisen

Eisenstange (1m) bei $20^{\mathbf{o}} C$

Bei Temperaturerhöhung dehnt sich Eisenstange aus gemäß

$$\Delta L \;=\; \alpha \cdot \Delta T$$

α : Wärmeausdehnungskoeffizient von Eisen



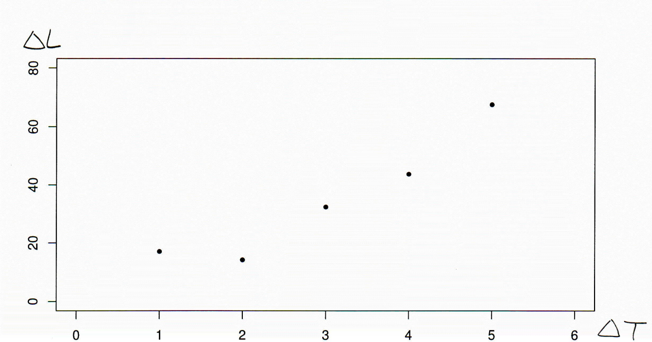
Wärmeausdehnungskoeffizient von Eisen

Eisenstange (1m) bei $20^{\mathbf{o}}C$

Bei Temperaturerhöhung dehnt sich Eisenstange aus gemäß

$$\Delta L = \alpha \cdot \Delta T$$

α : Wärmeausdehnungskoeffizient von Eisen



Wärmeausdehnungskoeffizient von Eisen

Eisenstange (1m) bei $20^{\mathbf{o}}C$

Bei Temperaturerhöhung dehnt sich Eisenstange aus gemäß

$$\Delta L \;=\; \alpha \cdot \Delta T$$

Messergebnisse:

ΔT	ΔL
1	17,3
2	14,1
3	32,4
4	43,2
5	67,0

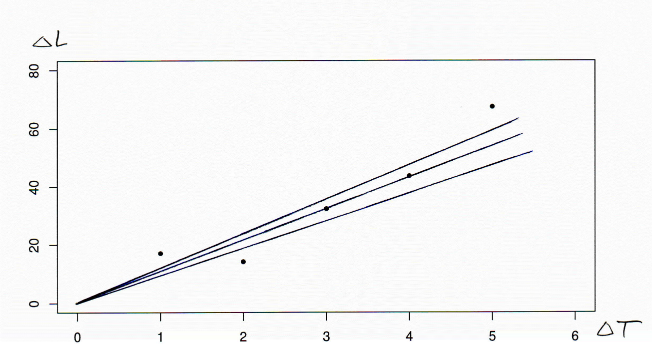
Wärmeausdehnungskoeffizient von Eisen

Eisenstange (1m) bei $20^{\mathbf{o}}C$

Bei Temperaturerhöhung dehnt sich Eisenstange aus gemäß

$$\Delta L \;=\; \alpha \cdot \Delta T$$

α : Wärmeausdehnungskoeffizient von Eisen



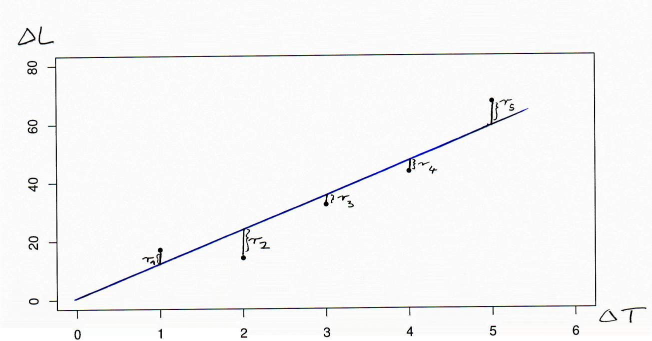
Wärmeausdehnungskoeffizient von Eisen

Eisenstange (1m) bei $20^{\circ} C$

Bei Temperaturerhöhung dehnt sich Eisenstange aus gemäß

$$\Delta L \;=\; \alpha \cdot \Delta T$$

α : Wärmeausdehnungskoeffizient von Eisen



Wärmeausdehnungskoeffizient von Eisen

Messergebnisse:

ΔT	ΔL	r_i
1	17,3	$17,3 - \alpha \cdot 1$
2	14,1	$14,1 - \alpha \cdot 2$
3	32,4	$32,4 - \alpha \cdot 3$
4	43,2	$43,2 - \alpha \cdot 4$
5	67,0	$67,0 - \alpha \cdot 5$

Setze

$$\begin{aligned} R(\alpha) &= r_1^2 + r_2^2 + \ldots + r_5^2 = \\ &= \big(17,3 - \alpha \cdot 1\big)^2 + \big(14,1 - \alpha \cdot 2\big)^2 + \ldots + \big(67,0 - \alpha \cdot 5\big)^2 \end{aligned}$$

Im Idealfall (ohne Messfehler) lägen alle Punkte auf einer Geraden, d.h., es gäbe ein α_0, so dass

$$R(\alpha_0) = 0$$

Wärmeausdehnungskoeffizient von Eisen

Messergebnisse:

ΔT	ΔL	r_i
1	17,3	$17,3 - \alpha \cdot 1$
2	14,1	$14,1 - \alpha \cdot 2$
3	32,4	$32,4 - \alpha \cdot 3$
4	43,2	$43,2 - \alpha \cdot 4$
5	67,0	$67,0 - \alpha \cdot 5$

Setze

$$
\begin{aligned}
R(\alpha) &= r_1^2 + r_2^2 + \ldots + r_5^2 = \\
&= \left(17,3 - \alpha \cdot 1\right)^2 + \left(14,1 - \alpha \cdot 2\right)^2 + \ldots + \left(67,0 - \alpha \cdot 5\right)^2
\end{aligned}
$$

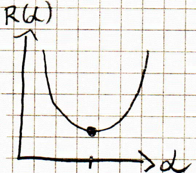
Idee: Wähle meine Schätzung $\hat{\alpha}$ so, dass

$$R(\hat{\alpha}) = \text{minimal}$$

Wärmeausdehnungskoeffizient von Eisen

Berechnung

$$R(\alpha) = (17{,}3 - 1\cdot\alpha)^2 + (14{,}1 - 2\cdot\alpha)^2 +$$
$$+ (32{,}4 - 3\cdot\alpha)^2 + (43{,}2 - 4\cdot\alpha)^2 +$$
$$+ (67{,}0 - 5\cdot\alpha)^2 =$$
$$= \ldots = 55\alpha^2 - 1301\alpha + 7903{,}1$$



$\hat{\alpha}$ = Minimalstelle von $\alpha \mapsto R(\alpha)$

$\hat{\alpha} \approx 11{,}8$

(Quadratische Ergänzung oder Ableitung gleich 0 setzen)

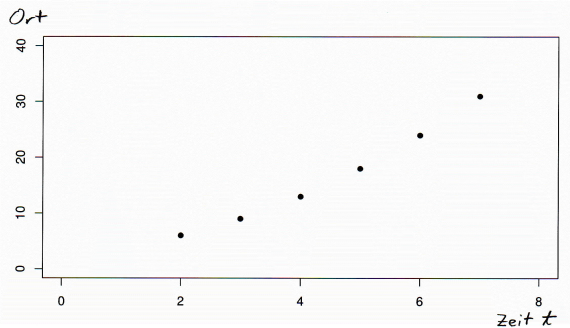
Gleichmäßig beschleunigte Bewegung

$$\mathrm{Ort}(t) \;=\; a \cdot \frac{t^2}{2} + v \cdot t + s \qquad \text{wobei} \quad t \;=\; \text{Zeit},$$

a : Beschleunigung

v : Anfangsgeschwindigkeit

s : Anfangsposition



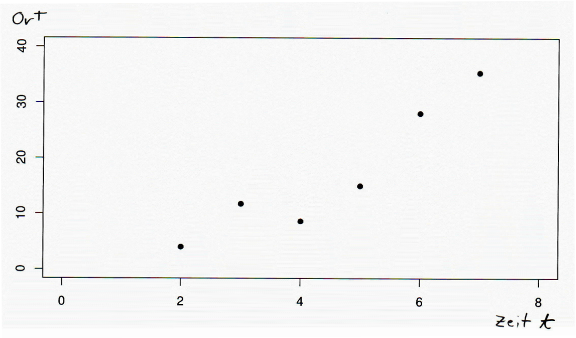
Gleichmäßig beschleunigte Bewegung

$$\mathrm{Ort}(t) = a \cdot \frac{t^2}{2} + v \cdot t + s \qquad \text{wobei} \quad t = \text{Zeit},$$

a : Beschleunigung

v : Anfangsgeschwindigkeit

s : Anfangsposition



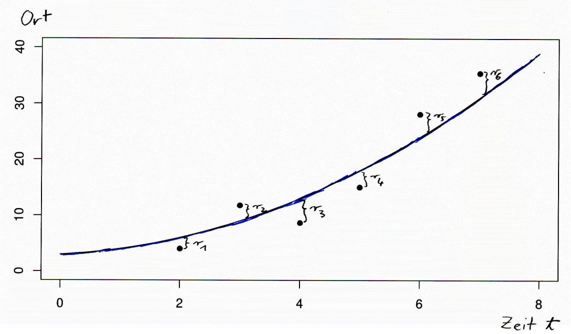
Gleichmäßig beschleunigte Bewegung

$$\mathrm{Ort}(t) = a \cdot \frac{t^2}{2} + v \cdot t + s \qquad \text{wobei} \quad t = \text{Zeit},$$

a : Beschleunigung

v : Anfangsgeschwindigkeit

s : Anfangsposition



Gleichmäßig beschleunigte Bewegung

Messergebnisse:

t	Ort(t)	r_i
2	4,0	$4{,}0 - 2a - 2v - s$
3	11,8	$11{,}8 - 4{,}5a - 3v - s$
4	8,6	$8{,}6 - 8a - 4v - s$
5	15,0	$15{,}0 - 12{,}5a - 5v - s$
6	28,1	$28{,}1 - 18a - 6v - s$
7	35,4	$35{,}4 - 24{,}5a - 7v - s$

Gleichmäßig beschleunigte Bewegung

Messergebnisse:

t	$\text{Ort}(t)$	r_i
2	4,0	$4{,}0 - 2a - 2v - s$
3	11,8	$11{,}8 - 4{,}5a - 3v - s$
4	8,6	$8{,}6 - 8a - 4v - s$
5	15,0	$15{,}0 - 12{,}5a - 5v - s$
6	28,1	$28{,}1 - 18a - 6v - s$
7	35,4	$35{,}4 - 24{,}5a - 7v - s$

Beispiel:

$$r_1 = \text{Ort}(t_1) - \left(a \cdot \frac{t_1^2}{2} + v \cdot t_1 + s\right) =$$

$$\stackrel{t_1=2}{=} 4{,}0 - 2a - 2v - s$$

Gleichmäßig beschleunigte Bewegung

Messergebnisse:

t	Ort(t)	r_i
2	4,0	$4{,}0 - 2a - 2v - s$
3	11,8	$11{,}8 - 4{,}5a - 3v - s$
4	8,6	$8{,}6 - 8a - 4v - s$
5	15,0	$15{,}0 - 12{,}5a - 5v - s$
6	28,1	$28{,}1 - 18a - 6v - s$
7	35,4	$35{,}4 - 24{,}5a - 7v - s$

Setze

$$
\begin{aligned}
R(a, v, s) &= r_1^2 + r_2^2 + \ldots + r_6^2 = \\
&= \left(4{,}0 - 2a - 2v - s\right)^2 + \ldots + \left(35{,}4 - 24{,}5a - 7v - s\right)^2
\end{aligned}
$$

Im Idealfall (ohne Messfehler) lägen alle Punkte auf einer Parabel, d.h., es gäbe ein a_0, ein v_0 und ein s_0, so dass

$$R(a_0, v_0, s_0) = 0$$

Gleichmäßig beschleunigte Bewegung

Messergebnisse:

t	Ort(t)	r_i
2	4,0	$4{,}0 - 2a - 2v - s$
3	11,8	$11{,}8 - 4{,}5a - 3v - s$
4	8,6	$8{,}6 - 8a - 4v - s$
5	15,0	$15{,}0 - 12{,}5a - 5v - s$
6	28,1	$28{,}1 - 18a - 6v - s$
7	35,4	$35{,}4 - 24{,}5a - 7v - s$

Setze

$$\begin{aligned} R(a, v, s) &= r_1^2 + r_2^2 + \ldots + r_6^2 = \\ &= \left(4{,}0 - 2a - 2v - s\right)^2 + \ldots + \left(35{,}4 - 24{,}5a - 7v - s\right)^2 \end{aligned}$$

Idee: Wähle meine Schätzungen $\hat{a}$, $\hat{v}$ und $\hat{s}$ so, dass

$$R(\hat{a}, \hat{v}, \hat{s}) = \text{minimal}$$

Gleichmäßig beschleunigte Bewegung

Mathematisches Problem:

Minimiere

$$R(a, v, s) \;=\; \big(4{,}0 - 2a - 2v - s\big)^2 + \ldots + \big(35{,}4 - 24{,}5a - 7v - s\big)^2$$

über alle möglichen a, v und s.

$\longrightarrow$ Differentialrechnung für Funktionen mehrerer Veränderlicher, partielle Ableitungen **(Analysis II, 2. Semester)**

$\longrightarrow$ Gleichungssystem in Matrizenschreibweise

$$X^{\mathsf{T}} X \theta \;=\; X^{\mathsf{T}} y,$$

auflösen nach $\theta = (a, v, s)$
(Lineare Algebra I, 1. Semester)

Gleichmäßig beschleunigte Bewegung

t	Ort(t)	r_i
2	4,0	$4{,}0 - 2a - 2v - s$
3	11,8	$11{,}8 - 4{,}5a - 3v - s$
4	8,6	$8{,}6 - 8a - 4v - s$
5	15,0	$15{,}0 - 12{,}5a - 5v - s$
6	28,1	$28{,}1 - 18a - 6v - s$
7	35,4	$35{,}4 - 24{,}5a - 7v - s$

Setze

$$\text{Matrix } X = \begin{pmatrix} 2 & 2 & 1 \\ 4{,}5 & 3 & 1 \\ 8 & 4 & 1 \\ 12{,}5 & 5 & 1 \\ 18 & 6 & 1 \\ 24{,}5 & 7 & 1 \end{pmatrix} \quad \text{und} \quad y = \begin{pmatrix} 4{,}0 \\ 11{,}8 \\ 8{,}6 \\ 15{,}0 \\ 28{,}1 \\ 35{,}4 \end{pmatrix}$$

Gleichmäßig beschleunigte Bewegung

Mit

$$\text{Matrix } X = \begin{pmatrix} 2 & 2 & 1 \\ 4{,}5 & 3 & 1 \\ 8 & 4 & 1 \\ 12{,}5 & 5 & 1 \\ 18 & 6 & 1 \\ 24{,}5 & 7 & 1 \end{pmatrix} \quad \text{und} \quad y = \begin{pmatrix} 4{,}0 \\ 11{,}8 \\ 8{,}6 \\ 15{,}0 \\ 28{,}1 \\ 35{,}4 \end{pmatrix}$$

ergibt sich für Gleichung $X^{\mathsf{T}} X \theta = X^{\mathsf{T}} y$ hier konkret

$$\begin{pmatrix} 1168{,}75 & 391{,}5 & 69{,}5 \\ 391{,}50 & 139{,}0 & 27{,}0 \\ 69{,}50 & 27{,}0 & 6{,}0 \end{pmatrix} \cdot \begin{pmatrix} a \\ v \\ s \end{pmatrix} = \begin{pmatrix} 1690{,}5 \\ 569{,}2 \\ 102{,}9 \end{pmatrix}$$

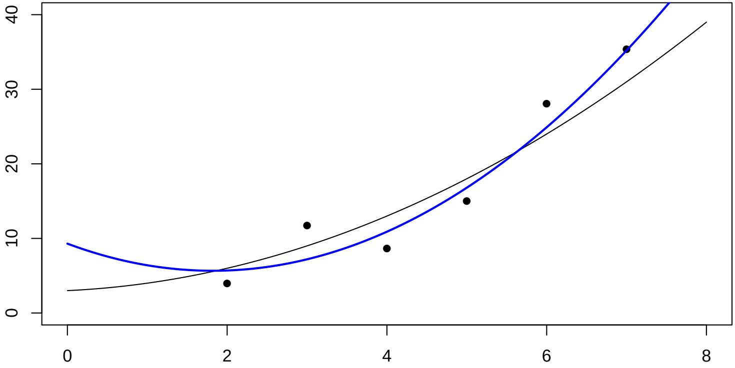
Gleichmäßig beschleunigte Bewegung

Lösung:

```
t <- c(2,3,4,5,6,7)
X <- cbind(t^2/2,t,1)
X <- matrix(X,ncol=3,dimnames = NULL)
y <- c(4.0, 11.8,  8.6, 15.0, 28.1, 35.4)
theta <- solve(t(X)%*%X,t(X)%*%y)
round(theta,digit=1)
      [,1]
[1,]   2.2
[2,]  -4.0
[3,]   9.3
```

(Computer-Implementierung: Numerik, 3. Semester)

Gleichmäßig beschleunigte Bewegung



Kleinste-Quadrate-Regression

Zusammenhang

$$\mathbf{x} \rightsquigarrow \mathbf{y}$$

Kleinste-Quadrate-Regression:

Suche $\hat{\theta}$, so dass

$$R(\theta) = \Big(y_1 - f_\theta(x_1)\Big)^2 + \ldots + \Big(y_n - f_\theta(x_n)\Big)^2$$

minimal in θ ist. **(Statistik, 4. Semester)**

Kleinste-Quadrate-Regression

Zusammenhang

$$\mathbf{x} \rightsquigarrow \mathbf{y}$$

Kleinste-Quadrate-Regression:

Suche $\hat{\theta}$, so dass

$$R(\theta) \;=\; \Big(y_1 - f_\theta(x_1)\Big)^2 + \;\ldots\; + \Big(y_n - f_\theta(x_n)\Big)^2$$

minimal in θ ist. **(Statistik, 4. Semester)**

Alternative:

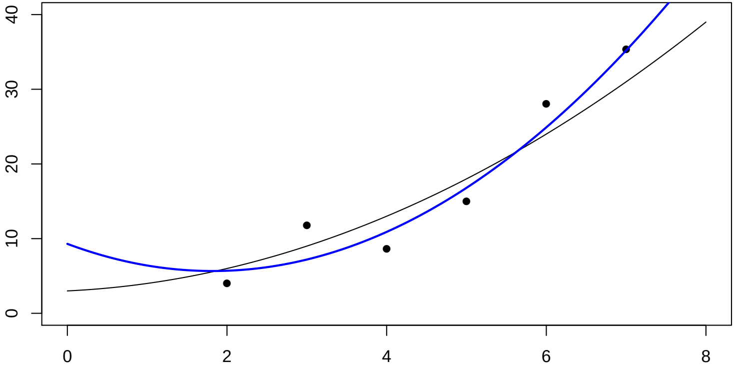
Suche $\hat{\theta}$, so dass

$$R(\theta) \;=\; \Big|y_1 - f_\theta(x_1)\Big| + \;\ldots\; + \Big|y_n - f_\theta(x_n)\Big|$$

minimal in θ ist. **(Nichtlin. Optimierung, 5./6. Semester)**

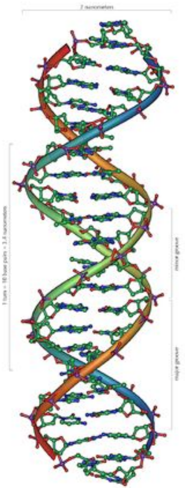
Vorheriges Beispiel:

$n = 6$ Beobachtungen, $p = 3$ zu schätzende Parameter



Ideal: Anzahl der Parameter p sehr klein im Vergleich zur Anzahl der Beobachtungen n

Gen-Daten



Beispiel: Einfluss von Genen auf Krankheiten

$\longrightarrow$ hochdimensionale Modelle

z.B.

$n = 49$ Beobachtungen

$p = 7129$ Parameter

$\longrightarrow$ bisherige Methode nicht mehr anwendbar

LASSO

Least Absolute Shrinkage and Selection Operator

Kleinste Quadrate:

Suche $\hat{\theta}$, so dass

$$\Big(y_1 - f_\theta(x_1)\Big)^2 + \ldots + \Big(y_n - f_\theta(x_n)\Big)^2$$

minimal in $\theta = (\theta_1, \ldots, \theta_p)$ ist.

LASSO:

Suche $\hat{\theta}$, so dass

$$\Big(y_1 - f_\theta(x_1)\Big)^2 + \ldots + \Big(y_n - f_\theta(x_n)\Big)^2 + \lambda\Big(\big|\theta_1\big| + \ldots + \big|\theta_p\big|\Big)$$

minimal in $\theta = (\theta_1, \ldots, \theta_p)$ ist.

Nichtparametrische Regression

Bisherige Beispiele

- Ellipse:

$$f_{a,b}(x_1, x_2) = \frac{x_1^2}{a^2} + \frac{x_2^2}{b^2}$$

- Wärmeausdehnung

$$f_{\alpha}(\Delta T) = \alpha \cdot \Delta T$$

- Gleichmäßig beschleunigte Bewegung

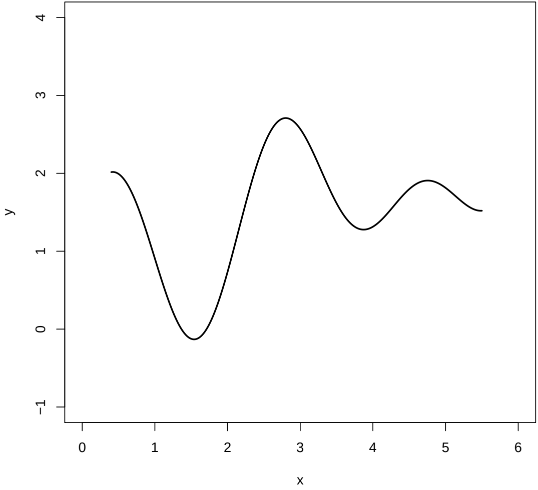
$$f_{a,v,s}(t) = a \cdot \frac{t^2}{2} + v \cdot t + s$$

Also: Funktion (d.h. Zusammenhang zwischen x und y) bereits bekannt bis auf (wenige) Parameter

Jetzt: Funktion vollkommen unbekannt

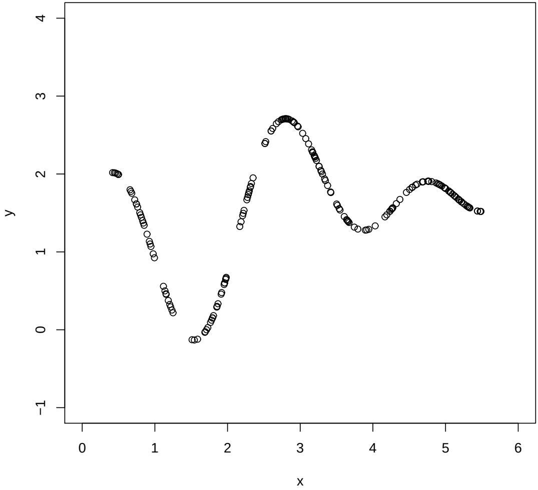
Nichtparametrische Regression

Funktion



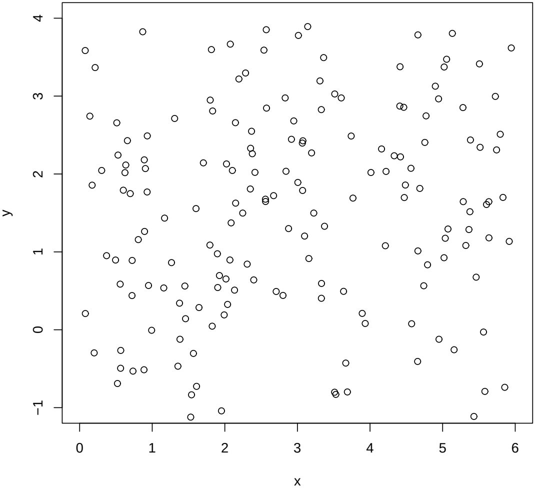
Nichtparametrische Regression

Daten ohne Streuung



Nichtparametrische Regression

Daten mit Streuung



Parametrische Regression, z.B. $f_\theta(x) = \theta_1 + \theta_2 x + \theta_3 x^2$

Suche $\hat{\theta} \in \mathbb{R}^p$, so dass

$$\Big(y_1 - f_\theta(x_1)\Big)^2 + \ldots + \Big(y_n - f_\theta(x_n)\Big)^2$$

minimal in $\theta = (\theta_1, \ldots, \theta_p)$ ist.

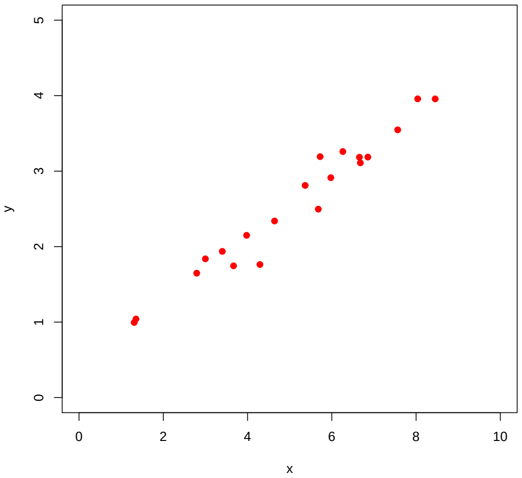
Nichtparametrische Regression

Suche $\hat{f} \in H$, so dass

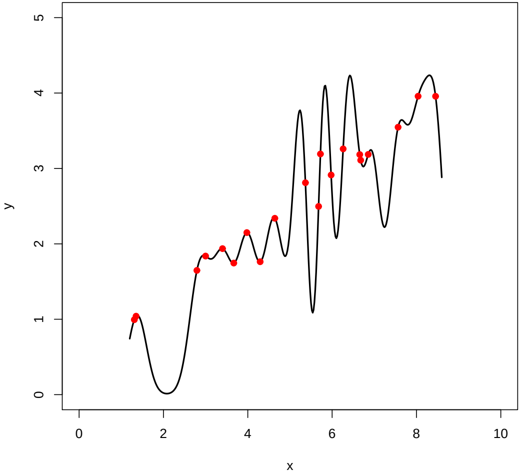
$$\Big(y_1 - f(x_1)\Big)^2 + \ldots + \Big(y_n - f(x_n)\Big)^2$$

minimal in f ist.

Overfitting



Overfitting



Parametrische Regression, z.B. $f_\theta(x) = \theta_1 + \theta_2 x + \theta_3 x^2$

Suche $\hat{\theta} \in \mathbb{R}^p$, so dass

$$\Big(y_1 - f_\theta(x_1)\Big)^2 + \ldots + \Big(y_n - f_\theta(x_n)\Big)^2$$

minimal in $\theta = (\theta_1, \ldots, \theta_p)$ ist.

Nichtparametrische Regression

Suche $\hat{f} \in H$, so dass

$$\Big(y_1 - f(x_1)\Big)^2 + \ldots + \Big(y_n - f(x_n)\Big)^2$$

minimal in f ist.

Parametrische Regression, z.B. $f_\theta(x) = \theta_1 + \theta_2 x + \theta_3 x^2$

Suche $\hat{\theta} \in \mathbb{R}^p$, so dass

$$\Big(y_1 - f_\theta(x_1)\Big)^2 + \ldots + \Big(y_n - f_\theta(x_n)\Big)^2$$

minimal in $\theta = (\theta_1, \ldots, \theta_p)$ ist.

Nichtparametrische Regression

Suche $\hat{f} \in H$, so dass

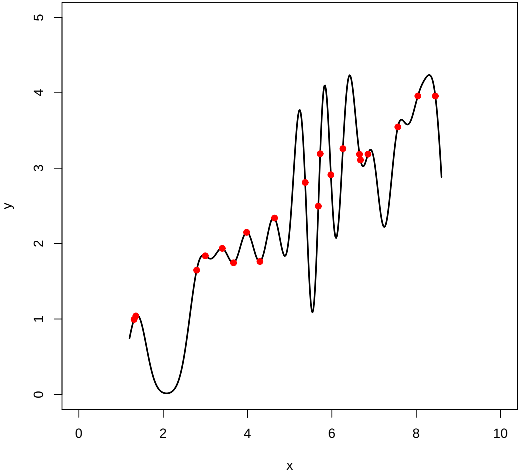
$$\Big(y_1 - f(x_1)\Big)^2 + \ldots + \Big(y_n - f(x_n)\Big)^2 + \text{Komplexität}(f)$$

minimal in f ist.

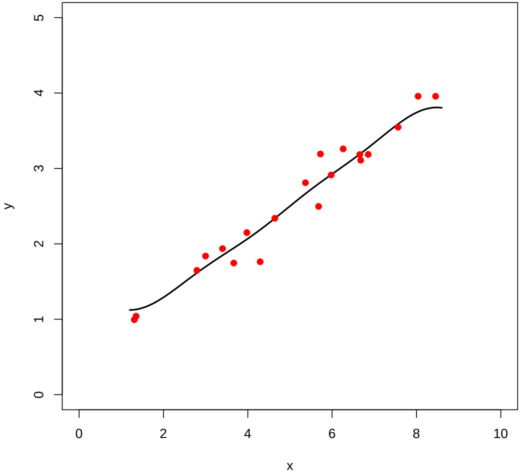
$\longrightarrow$ **LS-Support Vector Regression**

(Support Vector Machines, 5./6. Semester)

Overfitting

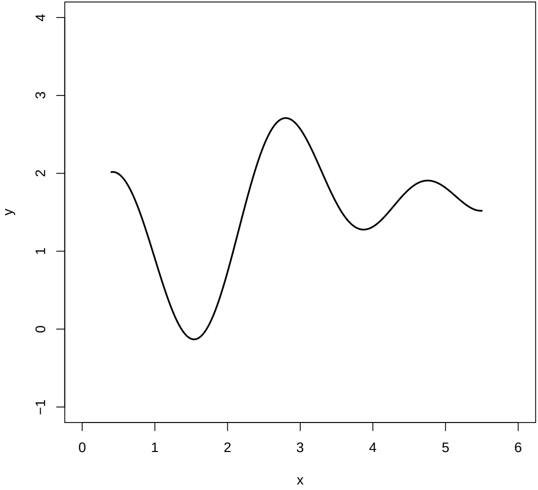


Overfitting



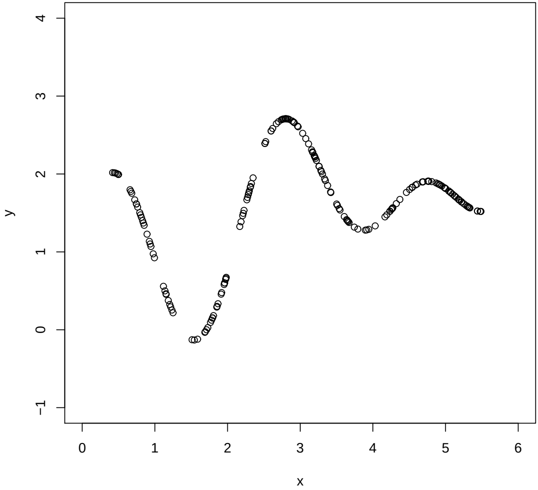
Nichtparametrische Regression

Funktion



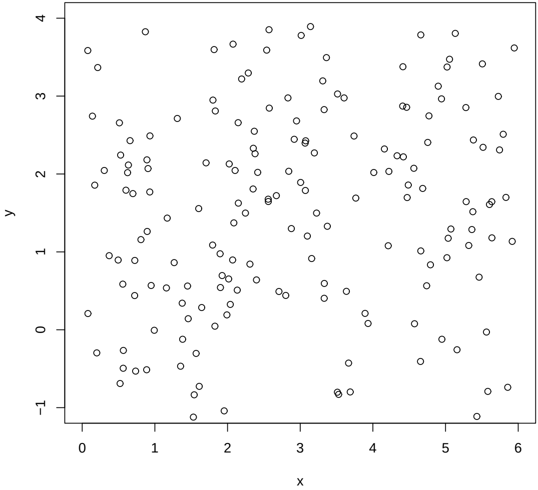
Nichtparametrische Regression

Daten ohne Streuung



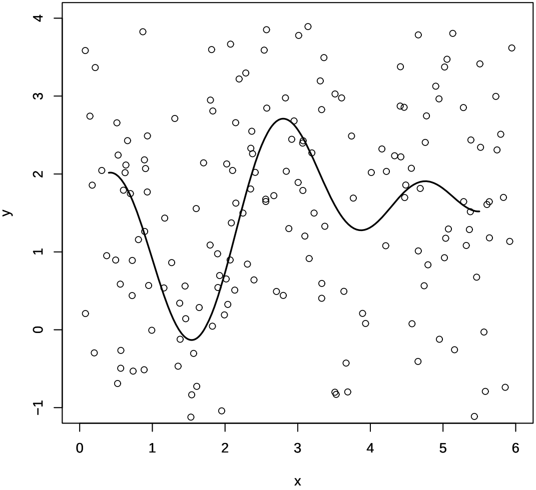
Nichtparametrische Regression

Daten mit Streuung



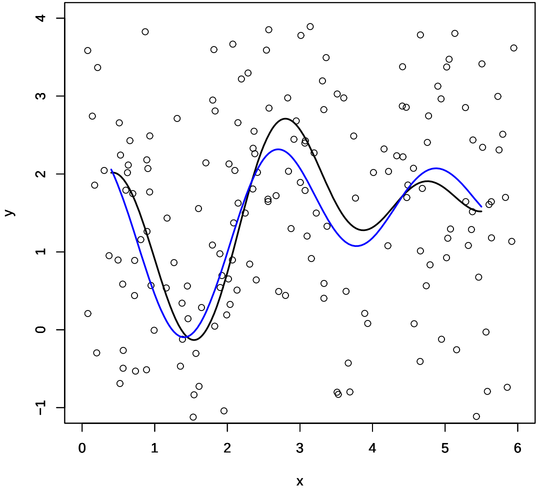
Nichtparametrische Regression

Daten mit Streuung



Nichtparametrische Regression

Daten mit Streuung



Klassifikation: Kreditausfall

- zwei Möglichkeiten

  $y_i = -1$: Kunde i zahlt Kredit zurück

  $y_i = +1$: Kunde i kann Kredit nicht zurückzahlen
- Informationen über den Kunden:

  $x_i = (x_{1,i}, x_{2,i}, \ldots, x_{d,i})$

  (Bsp.: Einkommen, Familienstand, Alter, Wohngegend, etc.)

Ziel: Prognose, ob Kunde mit Eigenschaften

$$x = \ldots$$

den Kredit zurückzahlen wird.

Mögliches Verfahren:

Suche $\hat{f} \in H$, so dass

$$\Big(y_1 - f(x_1)\Big)^2 + \ldots + \Big(y_n - f(x_n)\Big)^2 + \text{Komplexität}(f)$$

minimal in f ist. $\longrightarrow$ **LS-Support Vector Machine**

Klassifikation: Spam-Filter

- zwei Möglichkeiten

  $y_i = -1$: Email i ist keine Spam-Email

  $y_i = +1$: Email i ist eine Spam-Email

- Informationen über die Email:

  $x_i = (x_{1,i}, x_{2,i}, \ldots, x_{d,i})$

  (Bsp.: Betreff, bestimmte Schlagworte, Dateianhänge, etc.)

Ziel: Einschätzung, ob Email mit Eigenschaften

$$x = \ldots$$

eine Spam-Email ist.

Mögliches Verfahren:

Suche $\hat{f} \in H$, so dass

$$\Big(y_1 - f(x_1)\Big)^2 + \ldots + \Big(y_n - f(x_n)\Big)^2 + \text{Komplexität}(f)$$

minimal in f ist. $\longrightarrow$ **LS-Support Vector Machine**

Data Mining, Maschinelles Lernen

- Datensatz

$$(x_1, y_1), \ldots, (x_n, y_n)$$

- statistisches Verfahren
(z.B.: Kleinste-Quadrate,
Support Vector Machines)
- Berechnung der Schätzung
- Untersuchung von Zusammenhängen zwischen x und y

- Trainigsdaten

$$(x_1, y_1), \ldots, (x_n, y_n)$$

- Lern-Algorithmus
(z.B.: Kleinste-Quadrate,
Support Vector Machines)
- Trainieren des Algorithmus / Lernen
- Mustererkennung

(im Wesentlichen hier nur andere Bezeichnungen)

Klassifikation: Spam-Filter

- zwei Möglichkeiten

  $y_i = -1$: Email i ist keine Spam-Email

  $y_i = +1$: Email i ist eine Spam-Email
- Informationen über die Email:

  $x_i = (x_{1,i}, x_{2,i}, \ldots, x_{d,i})$

  (Bsp.: Betreff, bestimmte Schlagworte, Dateianhänge, etc.)

Ziel: Einschätzung, ob Email mit Eigenschaften

$$x = \ldots$$

eine Spam-Email ist.

Mögliches Verfahren:

Suche $\hat{f} \in H$, so dass

$$\Big(y_1 - f(x_1)\Big)^2 + \ldots + \Big(y_n - f(x_n)\Big)^2 + \text{Komplexität}(f)$$

minimal in f ist. $\longrightarrow$ **LS-Support Vector Machine**

Algorithmen

Bsp.: Lösungsformel für **quadratische Gleichung** $ax^2 + bx + c = 0$

```
quadratische.gleichung <- function(a,b,c){
    if(a == 0){
       print(paste("Die Loesung ist ",-c/b,".",sep=""))
    }
    if(a != 0){
       Diskriminante <- b^2 - 4*a*c
       if(Diskriminante < 0){
         print("Es gibt keine reelle Loesung.")
       }
       if(Diskriminante == 0){
         print(paste("Die Loesung ist ",-b/(2*a),".",sep=""))
       }
       if(Diskriminante > 0){
         print(paste("Die beiden Loesungen sind ",
                     (-b-sqrt(Diskriminante))/(2*a),
                     " und ",
                     (-b+sqrt(Diskriminante))/(2*a),".",
                     sep=""
                    )
              )
       }
    }
}
```

Künstliche Intelligenz ?

Computer: Abarbeiten von Algorithmen

- Computer arbeiten extrem schnell und präzise Anweisungen ab.
- Anweisungen müssen allerdings genauestens und absolut fehlerfrei formuliert werden.

Schachprogramme:

“Von der aktuellen Stellung (Spielsituation) ausgehend, führt das Programm eine Iterative Tiefensuche durch. In jeder Iteration führt es verschiedene Zugfolgen der Reihe nach aus, bewertet die erreichten Stellungen (. . .) und damit auch die Züge.”

aus dem Wikipedia-Artikel “Schachprogramm”

Zusammenfassung

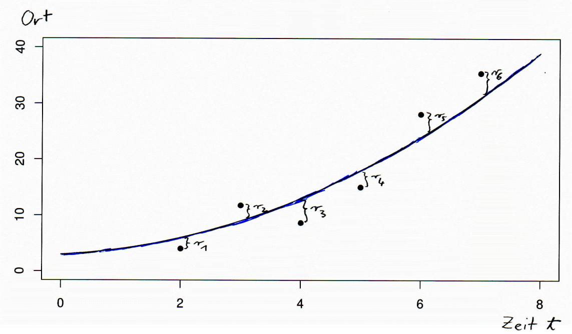
Gleichmäßig beschleunigte Bewegung

$$\mathrm{Ort}(t) \;=\; a \cdot \frac{t^2}{2} + v \cdot t + s \qquad \text{wobei} \quad t \;=\; \text{Zeit},$$

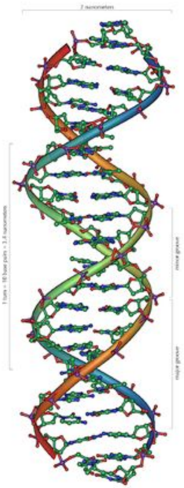
a : Beschleunigung

v : Anfangsgeschwindigkeit

s : Anfangsposition



Gen-Daten



Beispiel: Einfluss von Genen auf Krankheiten

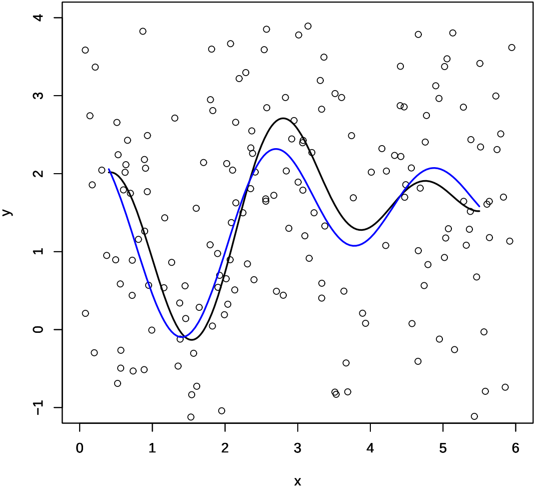
$\longrightarrow$ hochdimensionale Modelle

z.B.

$n = 49$ Beobachtungen
$p = 7129$ Parameter

Nichtparametrische Regression

Daten mit Streuung



Klassifikation: Spam-Filter

- zwei Möglichkeiten

  $y_i = -1$: Email i ist keine Spam-Email

  $y_i = +1$: Email i ist eine Spam-Email

- Informationen über die Email:

  $x_i = (x_{1,i}, x_{2,i}, \ldots, x_{d,i})$

  (Bsp.: Betreff, bestimmte Schlagworte, Dateianhänge, etc.)

Ziel: Einschätzung, ob Email mit Eigenschaften

$$x = \ldots$$

eine Spam-Email ist.

Mögliches Verfahren:

Suche $\hat{f} \in H$, so dass

$$\Big(y_1 - f(x_1)\Big)^2 + \ldots + \Big(y_n - f(x_n)\Big)^2 + \text{Komplexität}(f)$$

minimal in f ist. $\longrightarrow$ **LS-Support Vector Machine**

Vielen Dank für Ihre Aufmerksamkeit

... und weiterhin einen spannenden Tag der Mathematik an der Uni Bayreuth