# Ausgleichsprobleme

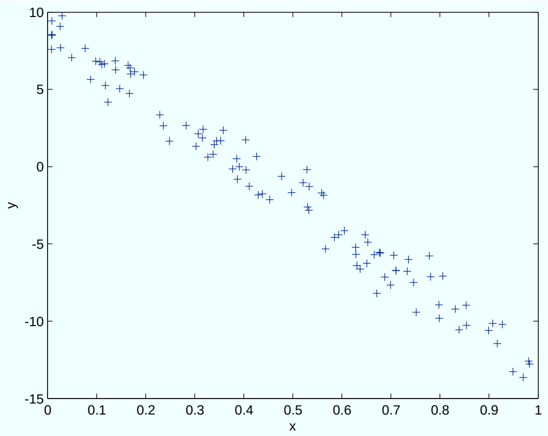
Numerische Mathematik 1
WS 2011/12

# Lineare Regression

Es sei eine Menge von Datenpunkten

$$(x_1, y_1), \dots, (x_n, y_n) \in \mathbb{R}^2$$

gegeben, die mit Fehlern behaftet sind.

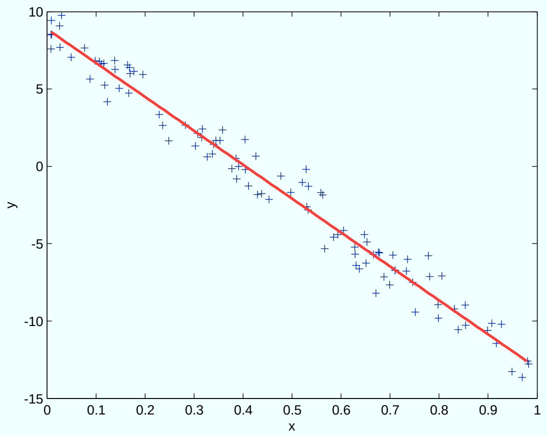


# Lineare Regression

Ziel: Eine Gerade $g : \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}$ finden

$$g(x; a_1, a_0) := a_1 x + a_0$$

die diese Datenpunkte möglichst gut beschreibt.

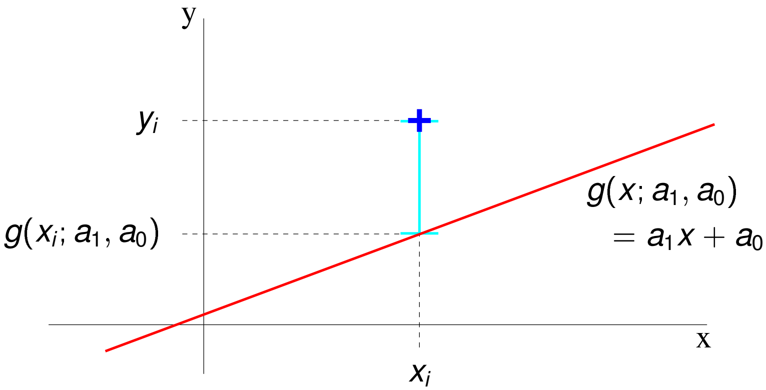


# Lineare Regression

Für gegebene $a_1, a_0 \in \mathbb{R}$ kann der Fehler im $i$-ten Datenpunkt durch

$$R_i(a_1, a_0) := |g(x_i; a_1, a_0) - y_i| = |a_1 x_i + a_0 - y_i|$$

gemessen werden.



# Lineare Regression

Für gegebene $a_1, a_0 \in \mathbb{R}$ kann der Fehler im allen $n$ Datenpunkten durch die Funktion $c : \mathbb{R}^2 \to \mathbb{R}$ mit

$$c(a_1, a_0) = \left\| \begin{bmatrix} R_1(a_1, a_0) \\ \vdots \\ R_n(a_1, a_0) \end{bmatrix} \right\|_2 = \left\| \begin{bmatrix} g(x_1; a_1, a_0) - y_1 \\ \vdots \\ g(x_n; a_1, a_0) - y_n \end{bmatrix} \right\|_2$$

gemessen werden.

# Lineare Regression

Wegen der Form von $g(x; a_1, a_0) = a_1 x + a_0$ ist

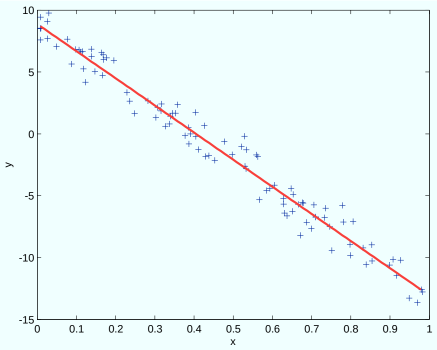
$$
\begin{aligned}
c(a_1, a_0) &:= \left\| \begin{bmatrix} g(x_1; a_1, a_0) - y_1 \\ \vdots \\ g(x_n; a_1, a_0) - y_n \end{bmatrix} \right\|_2 = \left\| \begin{bmatrix} a_1 x_1 + a_0 - y_1 \\ \vdots \\ a_1 x_n + a_0 - y_n \end{bmatrix} \right\|_2 \\
&= \left\| \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} a_1 + \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} a_0 - \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \right\|_2 \\
&= \left\| \underbrace{\begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}}_{=: M \in \mathbb{R}^{n,2}} \underbrace{\begin{bmatrix} a_0 \\ a_1 \end{bmatrix}}_{=: a} - \underbrace{\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}}_{=: y} \right\|_2 = \|Ma - y\|_2
\end{aligned}
$$

# Lineare Regression

Das Problem

$$\min_{a \in \mathbb{R}^2} c(a) = \min_{a \in \mathbb{R}^2} \|Ma - b\|_2$$

ist also ein lineares Ausgleichsproblem, wobei die Anzahl der Zeilen von $M \in \mathbb{R}^{n,2}$ der Anzahl der Datenpunkte (**+**) entspricht.



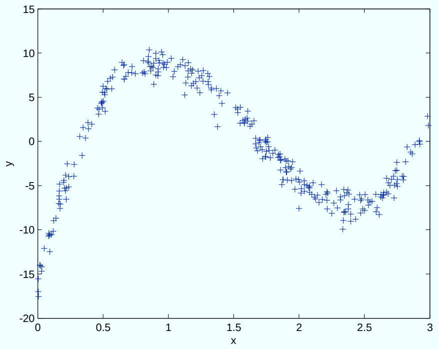
# Polynomieller Ausgleich

Ebenso kann man eine Menge von Datenpunkten

$$(x_1, y_1), \ldots, (x_n, y_n) \in \mathbb{R}^2$$

durch ein Polynom vom Grad $m \in \mathbb{N}$ beschreiben, d.h. durch eine Funktion $g : \mathbb{R} \times \mathbb{R}^{m+1} \to \mathbb{R}$ von der Form

$$g(x; a_m, \ldots, a_0) := a_m x^m + \ldots + a_1 x + a_0.$$



# Polynomieller Ausgleich

Hier kann man also den folgenden Ausdruck minimieren:

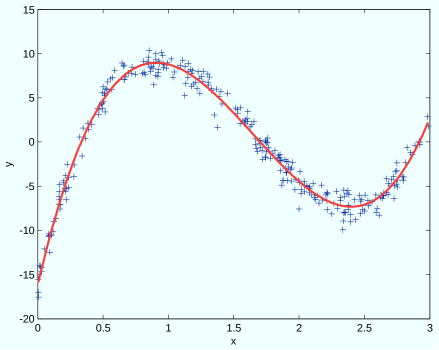
$$
\begin{aligned}
& \left\| \begin{bmatrix} g(x_1; a_m, \ldots, a_0) - y_1 \\ \vdots \\ g(x_n; a_m, \ldots, a_0) - y_n \end{bmatrix} \right\|_2 \\
= & \left\| \begin{bmatrix} a_m x_1^m + \ldots + a_1 x_1 + a_0 - y_1 \\ \vdots \\ a_m x_n^m + \ldots + a_1 x_n + a_0 - y_n \end{bmatrix} \right\|_2 \\
= & \left\| a_m \begin{bmatrix} x_1^m \\ \vdots \\ x_n^m \end{bmatrix} + \ldots + a_1 \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + a_0 \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} - \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \right\|_2 \\
= & \left\| \begin{bmatrix} 1 & x_1 & \ldots & x_1^m \\ \vdots & \vdots & & \vdots \\ 1 & x_n & \ldots & x_n^m \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_m \end{bmatrix} - \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \right\|_2
\end{aligned}
$$

# Polynomieller Ausgleich

Das Problem des Ausgleichs durch ein Polynom vom Grad $m$ führt also auf ein lineares Ausgleichsproblem der Form

$$\min_{a \in \mathbb{R}^{m+1}} \|Ma - b\|_2,$$

wobei die Anzahl der Zeilen von $M \in \mathbb{R}^{n,m+1}$ der Anzahl der Datenpunkte (+) entspricht.



# Nichtlineares Ausgleichsproblem

Ebenso kann man eine Menge von Datenpunkten

$$(x_1, y_1), \ldots, (x_n, y_n) \in \mathbb{R}^p \times \mathbb{R}^q$$

durch eine Funktion beschreiben, welche von $m \in \mathbb{N}$ Parametern abhängt, d.h. durch eine Funktion der Gestalt

$$g : \mathbb{R}^p \times \mathbb{R}^m \to \mathbb{R}^q.$$

Dies führt auf das Problem

$$\begin{aligned} \min_{a \in \mathbb{R}^m} \sum_{i=1}^{n} \|g(x_i; a) - y_i\|_2^2 &= \min_{a \in \mathbb{R}^m} \left\| \begin{bmatrix} g(x_1; a) - y_1 \\ \vdots \\ g(x_n; a) - y_n \end{bmatrix} \right\|_2^2 \\ &=: \min_{a \in \mathbb{R}^m} \|R(a)\|_2^2 , \end{aligned}$$

wobei $R : \mathbb{R}^m \to \mathbb{R}^{nq}$.

# Nichtlineares Problem

Nun ist $R : \mathbb{R}^m \to \mathbb{R}^{nq}$ aber im allgemeinen nicht affin-linear, d.h.

$$R(a) \neq Ma - b,$$

wobei $M \in \mathbb{R}^{nq,m}$ und $b \in \mathbb{R}^{nq}$.

Also haben wir keine Methode das Problem

$$\min_{a \in \mathbb{R}^m} \|R(a)\|,$$

zu lösen.

Idee: Benutze einen iterativen Prozess, wobei in jedem Schritt ein lineares Problem gelöst wird (wie beim Newton-Verfahren).

# Linearisierung

Ziel: Mit $R : \mathbb{R}^m \to \mathbb{R}^{nq}$ wollen wir

$$\min_{a \in \mathbb{R}^m} \|R(a)\| = \|R(a^*)\|,$$

lösen.

Taylor-Entwicklung erster Ordnung an der Stelle $a_k \in \mathbb{R}^m$ lässt hoffen, dass

$$R(a) \approx R(a_k) + \underbrace{DR(a_k)}_{\in \mathbb{R}^{nq,m}}(a - a_k),$$

und somit, dass

$$\min_{a \in \mathbb{R}^m} \|R(a)\| \approx \min_{a \in \mathbb{R}^m} \|R(a_k) + DR(a_k)(a - a_k)\|,$$

wobei letzters ein lineares Ausgleichsproblem ist.

# Gauß-Newton-Verfahren

Wähle $a_0 \in \mathbb{R}^m$
für $k = 0, 1, 2, \ldots$
berechne die Lösung

$$\min_{s_k \in \mathbb{R}^m} \| R(a_k) + DR(a_k) s_k \|,$$

$$a_{k+1} := a_k + s_k$$

# Beispiel

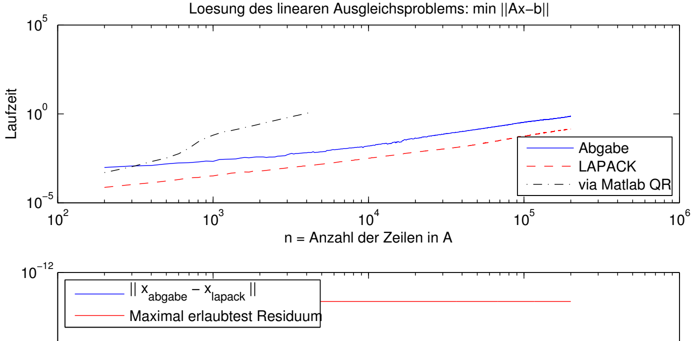
Wir versuchen nun die asymptotische Komplexität

$$\mathcal{O}(n^\alpha), \qquad \text{mit } \alpha \in \mathbb{R}$$

zu bestimmen, indem wir einen Ausgleich der Laufzeit-Daten (obere Graphik) gegen die Funktion

$$Cn^\alpha, \qquad \text{mit } C, \alpha \in \mathbb{R}$$

durchführen.



# Beispiel

Seien

$$(n_1, z_1), \dots, (n_\ell, z_\ell) \in \mathbb{R} \times \mathbb{R}$$

die gemessenen Zeiten für verschiedene Problemgrößen.

Mit $g : \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}$ gegeben durch

$$g(n; C, \alpha) := Cn^\alpha,$$

wollen wir also

$$\begin{aligned}
\min_{(C,\alpha)\in\mathbb{R}^2} \left\| \begin{bmatrix} g(n_1; C, \alpha) - z_1 \\ \vdots \\ g(n_\ell; C, \alpha) - z_\ell \end{bmatrix} \right\| &= \min_{(C,\alpha)\in\mathbb{R}^2} \left\| \begin{bmatrix} Cn_1^\alpha - z_1 \\ \vdots \\ Cn_\ell^\alpha - z_\ell \end{bmatrix} \right\| \\
&=: \min_{(C,\alpha)\in\mathbb{R}^2} \|R(C, \alpha)\|
\end{aligned}$$

lösen.

# Beispiel

Da die Ableitung von $R : \mathbb{R}^2 \to \mathbb{R}^\ell$

$$R(C,\alpha) = \begin{bmatrix} Cn_1^\alpha - z_1 \\ \vdots \\ Cn_\ell^\alpha - z_\ell \end{bmatrix}$$

durch

$$DR(C,\alpha) = \begin{bmatrix} n_1^\alpha & C\log(n_1)n_1^\alpha \\ \vdots & \vdots \\ n_\ell^\alpha & C\log(n_\ell)n_\ell^\alpha \end{bmatrix}$$

gegeben ist muss man im $k$-ten Schritt

$$\min_{(D,\beta)\in\mathbb{R}^2} \left\| \begin{bmatrix} C_k n_1^{\alpha_k} - z_1 \\ \vdots \\ C_k n_\ell^{\alpha_k} - z_\ell \end{bmatrix} + \begin{bmatrix} n_1^{\alpha_k} & C_k\log(n_1)n_1^{\alpha_k} \\ \vdots & \vdots \\ n_\ell^{\alpha_k} & C_k\log(n_\ell)n_\ell^{\alpha_k} \end{bmatrix} \begin{bmatrix} D \\ \beta \end{bmatrix} \right\|$$

lösen und dann $(C_{k+1},\alpha_{k+1}) := (C_k,\alpha_k) + (D,\beta)$ setzen.