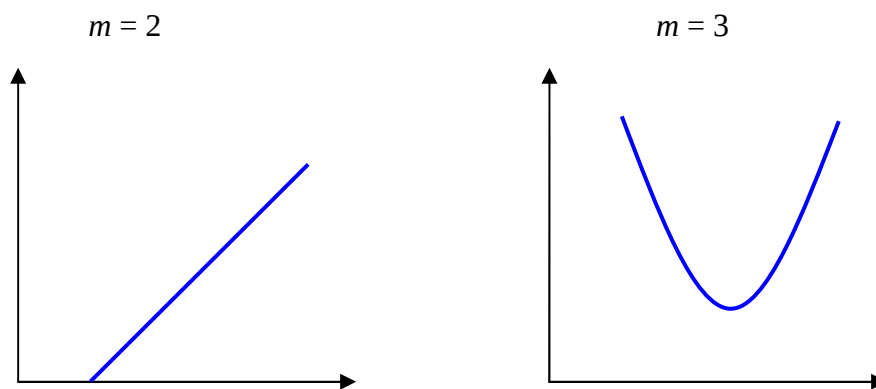


§ 33 Methode der kleinsten Quadrate

33.1 Problemstellung

In einem Experiment interessiert man sich für die Beziehung zwischen zwei Variablen x und y . Hierzu hat man viele Wertepaare $(x_1, y_1), \dots, (x_n, y_n)$ gemessen. Die Messungen können Fehler und statistische Schwankungen enthalten. Man möchte nun die Beziehung zwischen x und y durch ein einfaches Polynom $y = \sum_{k=1}^m a_k x^{k-1}$ beschreiben (z.B. $m = 2$, Gerade; $m = 3$, Parabel) und sucht die optimalen Koeffizienten $a_1, a_2, \dots, a_m$.



33.2 Methode der kleinsten Quadrate

Jede der n Messungen (x_i, y_i) beschreibt eine lineare Gleichung in den Unbekannten $a_1, a_2, \dots, a_m$:

$$\begin{aligned} a_1 + a_2 x_1 + a_3 x_1^2 + \dots + a_m x_1^{m-1} &= y_1 \\ &\vdots \\ a_1 + a_2 x_n + a_3 x_n^2 + \dots + a_m x_n^{m-1} &= y_n \end{aligned}$$

Im Allgemeinen hat man sehr viel mehr Gleichungen als Unbekannte ($n \gg m$) und das lineare Gleichungssystem

$$\underbrace{\begin{pmatrix} 1x_1 \dots x_1^{m-1} \\ \vdots \\ 1x_n \dots x_n^{m-1} \end{pmatrix}}_M \cdot \underbrace{\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}}_a = \underbrace{\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}}_y$$

ist nicht konsistent (einige Gleichungen widersprechen sich).

Beispielsweise kann man nicht erwarten, dass 50 Messpunkte auf einer Geraden liegen.

Da das Gleichungssystem $Ma = y$ nicht exakt lösbar ist, sucht man stattdessen eine Lösung a^* , die den quadratischen Fehler

$$|Ma - y|^2 = \left(\sum_{k=1}^m a_k x_1^{k-1} - y_1 \right)^2 + \dots + \left(\sum_{k=1}^m a_k x_n^{k-1} - y_n \right)^2$$

minimiert. (**Ausgleichskurve, Regressionskurve**)

Wir suchen also das Minimum der Funktion

$$\begin{aligned} f(a_1, \dots, a_n) &= (Ma - y)^T \cdot (Ma - y) \\ &= a^T M^T Ma - \underbrace{a^T M^T y}_{a^T M^T y} + y^T y \\ &= a^T M^T Ma - 2a^T M^T y + y^T y \end{aligned}$$

Notwendige Bedingung:

$$0 = \nabla_A f = \begin{pmatrix} \frac{\partial f}{\partial a_1} \\ \vdots \\ \frac{\partial f}{\partial a_n} \end{pmatrix} = M^T Ma + \underbrace{a^T M^T M}_{M^T Ma} - 2M^T y = 2M^T Ma - 2M^T y$$

Die Lösung a^* löst also die so genannte **Normalengleichung**

$$M^T Ma = M^T y \quad \text{mit } n \text{ Gleichungen und } m \text{ Unbekannten.}$$

Wenn $M^T M$ invertierbar ist, gilt also

$$a^* = (M^T M)^{-1} \cdot M^T y.$$

Man nennt

$$M^+ := (M^T M)^{-1} \cdot M^T$$

die **Pseudoinverse (Moore-Penrose-Inverse)** der (nicht invertierbaren!) $n \times m$ Matrix M .

Bemerkungen:

(a) Man kann zeigen, dass a^* tatsächlich ein Minimum von $f(a)$ ist.

(b) $M^T M$ ist invertierbar, falls m der n Gleichungen des Systems $Ma=y$ linear unabhängig sind, d.h. $\text{rang}(M) = m$.

Damit haben wir am Beispiel der Ausgleichsrechnung ein allgemeines Verfahren hergeleitet, um ein überbestimmtes (und i.A. inkonsistentes) Gleichungssystem zu „lösen“:

33.3 Pseudolösung überbestimmter Gleichungssysteme

Sei $n > m$, $A \in \text{Mat}_{n \times m}(\mathbb{R})$, $\text{rang}(A) = m$ und $b \in \mathbb{R}^n$.

Falls das überbestimmte lineare Gleichungssystem $Ax=b$ inkonsistent ist, ist es nicht lösbar.

Es existiert jedoch eine eindeutige **Pseudolösung** x^* , die den quadratischen Fehler

$|Ax - b|^2$ minimiert. Sie ist gegeben durch $x^* = A^+b$ mit der Pseudoinversen

$$A^+ := (A^T A)^{-1} \cdot A^T$$

Bemerkung:

Pseudolösungen spielen auch in der Informatik eine wichtige Rolle. Überbestimmte Gleichungssysteme treten z.B. bei der Suche in Internetdatenbanken auf.

33.4 Beispiel

Man bestimme mit der Methode der kleinsten Quadrate die Näherungsgerade der Punkte (0,1); (1,3); (2,4); (3,4).

Lösung:

Wir suchen die Koeffizienten a_1, a_2 der Geradengleichung $y = a_1 + a_2 x$.

Hierzu haben wir 4 Bestimmungsgleichungen:

$$1 = a_1 + a_2 \cdot 0$$

$$3 = a_1 + a_2 \cdot 1$$

$$4 = a_1 + a_2 \cdot 2$$

$$4 = a_1 + a_2 \cdot 3$$

Das überbestimmte Gleichungssystem lautet $Ma=y$ mit

$$M = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix} \quad a = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \quad y = \begin{pmatrix} 1 \\ 3 \\ 4 \\ 4 \end{pmatrix}.$$

$$M^T M = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 4 & 6 \\ 6 & 14 \end{pmatrix}$$

$$(M^T M)^{-1} = \dots = \frac{1}{10} \begin{pmatrix} 7 & -3 \\ -3 & 2 \end{pmatrix}$$

Damit lautet die Pseudolösung von $Ma=y$:

$$a^* = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = (M^T M)^{-1} \cdot M^T y = \frac{1}{10} \begin{pmatrix} 7 & -3 \\ -3 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 3 \\ 4 \\ 4 \end{pmatrix} = \begin{pmatrix} 1,5 \\ 1 \end{pmatrix}$$

Die Ausgleichsgerade lautet also: $y = 1,5 + x$

33.5 Robuste Statistik

Einige statistische Operationen wie die Mittelwertbildung durch arithmetisches Mitteln sind sehr anfällig gegenüber Ausreißern. Es gibt jedoch kaum Verfahren der robusten Statistik, die kaum empfindlich gegenüber Ausreißern sind.

Beim Beispiel der Mittelung ersetzt man hierbei das arithmetische Mittel durch den so genannten **Median**. Bei $2n+1$ Messwerten $x_0, \dots, x_{2n}$ ordnet man hierzu die Werte in aufsteigender Reihenfolge. Der Median ist damit das $(n+1)$ -te Element in dieser Reihe. Er ist robust gegenüber Ausreißern.