Analysis III

Winter 2017/2018

Jörn Behrens
mit
Kai Rothe

Extrema mit Nebenbedingungen
Ausgleichsrechnung

Buch Kapitel 5.14-5.15

Erinnerung: Extremalaufgaben (ohne Nebenbedingungen)

Definition: (lokale/relative Extrema)
Sei $f : D \to \mathbb{R}$, $D \subset \mathbb{R}^n$, $\mathbf{x}_0 \in D$ mit Umgebung U von $\mathbf{x}_0$.

- Falls
$$f(\mathbf{x}) \leq f(\mathbf{x}_0) \quad \text{für alle } \mathbf{x} \in U \cap D, \mathbf{x} \neq \mathbf{x}_0,$$
so ist $\mathbf{x}_0$ lokale oder relative Maximalstelle von f und f besitzt ein lokales oder relatives Maximum.
- Gilt in der obigen Ungleichung $f(\mathbf{x}) < f(\mathbf{x}_0)$, so heißt $\mathbf{x}_0$ echte lokale Maximalstelle von f.
- Gilt entsprechend $f(\mathbf{x}) \geq f(\mathbf{x}_0)$ bzw. $f(\mathbf{x}) > f(\mathbf{x}_0)$, so heißt $\mathbf{x}_0$ (echte) lokale Minimalstelle von f und die Funktion nimmt ein Minimum an.
- Allgemein spricht man von Extremalstellen bzw. Extrema.

Satz: (Hinreichende Bedingung für Extrema)
Sei $f : D \to \mathbb{R}$, $D \subset \mathbb{R}^n$ zweimal stetig diff'bar.
Dann ist $\mathbf{x}_0 \in \mathring{D}$ mit $f'(\mathbf{x}_0) = \mathbf{0}$

1. echte lokale Maximalstelle, falls $(\mathbf{z} \cdot \nabla)^2 f(\mathbf{x}_0) < 0$,
2. echte lokale Minimalstelle, falls $(\mathbf{z} \cdot \nabla)^2 f(\mathbf{x}_0) > 0$,

für alle $\mathbf{z} \in \mathbb{R}^n$, $\mathbf{z} \neq \mathbf{0}$.

Satz: (Notwendige Bedingung für Extrema)
Ist $\mathbf{x}_0 \in \mathring{D}$ lokale Extremalstelle einer partiell diff'baren Funktion $f : D \to \mathbb{R}$, $D \subset \mathbb{R}^n$, so gilt
$$f'(\mathbf{x}_0) = \mathbf{0}.$$

Bemerkungen:

- $f'(\mathbf{x}_0) = 0$ bedeutet, dass alle partiellen Ableitungen von f verschwinden.
- $\mathring{D}$ bezeichnet die inneren Punkte von D.

Definition: (lokale/relative Extrema)
Sei $f : D \to \mathbb{R}$, $D \subset \mathbb{R}^n$, $\mathbf{x}_0 \in D$ mit Umgebung U von $\mathbf{x}_0$.

- Falls
$$f(\mathbf{x}) \leq f(\mathbf{x}_0) \quad \text{für alle } \mathbf{x} \in U \cap D, \mathbf{x} \neq \mathbf{x}_0,$$
so ist $\mathbf{x}_0$ lokale oder relative Maximalstelle von f und f besitzt ein lokales oder relatives Maximum.

- Gilt in der obigen Ungleichung $f(\mathbf{x}) < f(\mathbf{x}_0)$, so heißt $\mathbf{x}_0$ echte lokale Maximalstelle von f.

- Gilt entsprechend $f(\mathbf{x}) \geq f(\mathbf{x}_0)$ bzw. $f(\mathbf{x}) > f(\mathbf{x}_0)$, so heißt $\mathbf{x}_0$ (echte) lokale Minimalstelle von f und die Funktion nimmt ein Minimum an.

- Allgemein spricht man von Extremalstellen bzw. Extrema.

ng für Extrema)
mal stetig diff'bar

-unktion nimmt ein Minimum an.

lstellen bzw. Extrema.

Satz: (Notwendige Bedingung für Extrema)
Ist $\mathbf{x}_0 \in \mathring{D}$ lokale Extremalstelle einer partiell diff'baren Funktion $f : D \to \mathbb{R}$, $D \subset \mathbb{R}^n$, so gilt

$$f'(\mathbf{x}_0) = \mathbf{0}.$$

Bemerkungen:

- $f'(\mathbf{x}_0) = \mathbf{0}$ bedeutet, dass alle partiellen Ableitungen von f verschwinden.
- $\mathring{D}$ bezeichnet die inneren Punkte von D.

lokale Minimalstelle von f und die

- Allgemein spricht man von Extrem

Satz: (Hinreichende Bedingung für Extrema)
Sei $f : D \to \mathbb{R}$, $D \subset \mathbb{R}^n$ zweimal stetig diff'bar.
Dann ist $\mathbf{x}_0 \in \mathring{D}$ mit $f'(\mathbf{x}_0) = \mathbf{0}$

1. echte lokale Maximalstelle, falls $(\mathbf{z} \cdot \nabla)^2 f(\mathbf{x}_0) < 0$,

2. echte lokale Minimalstelle, falls $(\mathbf{z} \cdot \nabla)^2 f(\mathbf{x}_0) > 0$,

für alle $\mathbf{z} \in \mathbb{R}^n$, $\mathbf{z} \neq \mathbf{0}$.

Extremalaufgaben (*mit* Nebenbedingungen)

Definition: (Maxima-/Minimalstelle)
Wie schon zuvor definieren wir: Eine lokale Maximalstelle x_0 von $f|_M$ ist ein Punkt in M, zu dem eine Umgebung $U \subseteq D$ existiert, so dass

$$f(x) \le f(x_0) \quad \forall x \in U \cap D.$$

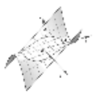
Analog ist eine Minimalstelle definiert.

Satz: (Notwendige Extremalbedingung bei m Nebenbedingungen)
Seien $f : D \to \mathbb{R}$ und $\mathbf{h} : D \to \mathbb{R}^m$ stetig diff'bar auf offener Menge $D \subseteq \mathbb{R}^n$, $m < n$, wobei die Jacobi-Matrix $\mathbf{h}'(\mathbf{x})$ für jedes $\mathbf{x} \in D$ den Rang m hat.

Ist dann $\mathbf{x}_0 \in D$ eine lokale Extremalstelle unter Nebenbedingung $\mathbf{h}(\mathbf{x}) = 0$, so existiert dazu eine $(1 \times m)$-Matrix $\mathbf{L} = (\lambda_1, \ldots, \lambda_m)$ (Zeilenvektor) mit

$$f'(\mathbf{x}_0) - \mathbf{L}\mathbf{h}'(\mathbf{x}_0) = 0 \quad \text{und} \quad \mathbf{h}(\mathbf{x}_0) = 0$$

Die reellen $\lambda_1, \ldots, \lambda_m$ heißen Lagrange Multiplikatoren.



1

Satz: (Notwendige Extremalbedingung bei einer Nebenbedingung)
Seien $f : D \to \mathbb{R}$ und $g : D \to \mathbb{R}$ stetig partiell diff'bar auf offener Menge $D \subseteq \mathbb{R}^n$.
Es gelte $\nabla g(\mathbf{x}) \neq 0$ für jedes $\mathbf{x} \in D$.

Ist dann $\mathbf{x}_0 \in D$ eine lokale Extremalstelle von f unter der Nebenbedingung $g(\mathbf{x}) = 0$, so gilt

$$\nabla f(\mathbf{x}_0) + \lambda \nabla g(\mathbf{x}_0) = 0$$

mit $\lambda \in \mathbb{R}$ Lagrange-Multiplikator.

Idee: Mit den folgenden Vereinbarungen

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad \mathbf{h} = \begin{pmatrix} h_1 \\ \vdots \\ h_m \end{pmatrix}, \quad \text{und} \quad \mathbf{L} = (\lambda_1, \ldots, \lambda_m).$$

Schreibt man

$$\frac{\partial f}{\partial x_j}(\mathbf{x}) - \sum_{k=1}^{m} \lambda_k \frac{\partial h_k}{\partial x_j} = 0, \quad \forall j = 1, \ldots, n$$

und

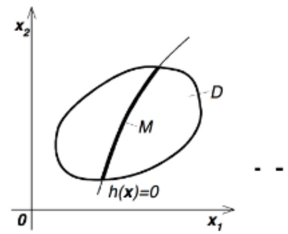
$$h_k(\mathbf{x}) = 0 \quad \forall j = 1, \ldots, m.$$

Also gibt es $n + m$ Gleichungen für $n + m$ Unbekannte $x_1, \ldots, x_n, \lambda_1, \ldots, \lambda_m$.
Die Lösungen $\mathbf{x} = (x_1, \ldots, x_n)^T$ heißen stationäre oder kritische Punkte und sie sind die Kandidaten für Extremalstellen.

Vorbemerkung: (Motivation)

- Häufig sind Extremwerte einer Funktion f gesucht, die weiteren Bedingungen genügen, Nebenbedingungen der Form

$$\mathbf{h}(\mathbf{x}) = 0.$$



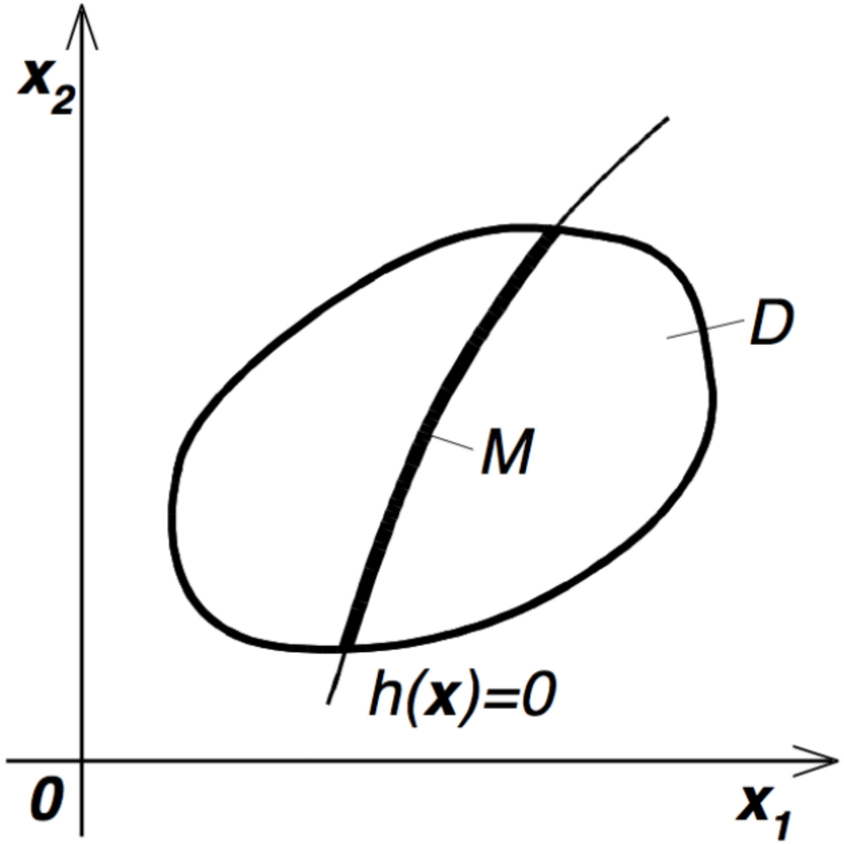
Beispiel: (Graphisch)
Menge $M = \{\mathbf{x} \in D : \mathbf{h}(\mathbf{x}) = 0\}$ für $n = 2, m = 1$.

- Man nimmt an:
  - $f : D \to \mathbb{R}$ mit $D \subset \mathbb{R}^n$,
  - $\mathbf{h} : D \to \mathbb{R}^m$ mit $m < n$,
  - Gesucht sind dann Extremalstellen der Einschränkung $f|_M$ von f auf

$$M := \{\mathbf{x} \in D : \mathbf{h}(\mathbf{x}) = 0\} \subset D.$$

- Für $m < n$ ist $\mathbf{h}(\mathbf{x}) = 0$ ein unterbestimmtes Gleichungssystem. M wird i.A. eine Mannigfaltigkeit mit $(n - m)$ freien Parametern sein.

- Wäre $n \leq m$ bestünde M i.A. nur aus einem Punkt oder wäre leer, also wäre eine Extremalwertsuche sinnlos.



Beispiel: (Graphisch)
Menge $M = \{\mathbf{x} \in D : \mathbf{h}(\mathbf{x}) = 0\}$ für $n = 2, m = 1$.

Beispiele: (Praktisch/Anschaulich)

- Finde maximale Leistung eines Motors innerhalb eines vorgegebenen Drehzahlbereiches.
- Finde kürzeste Strecke einer Logistik-Kette unter der Bedingung, dass eine bestimmte Anzahl Punkte versorgt werden.
- Finde energieeffizientesten Betriebszustand einer Verbrennung unter der Bedingung, dass eine vorgegebene Temperatur erreicht wird.

Definition: (Maximal-/Minimalstelle)
Wie schon zuvor definieren wir: Eine lokale Maximalstelle $\mathbf{x}_0$ von $f|_M$ ist ein Punkt in M, zu dem eine Umgebung $U \subset D$ existiert, so dass

$$f(\mathbf{x}) \leq f(\mathbf{x}_0) \quad \forall \mathbf{x} \in U \cap D.$$

Analog ist eine Minimalstelle definiert.

Satz: (Notwendige Extremalbedingung bei m Nebenbedingungen)
Seien $f : D \to \mathbb{R}$ und $\mathbf{h} : D \to \mathbb{R}^m$ stetig diff'bar auf offener Menge $D \subset \mathbb{R}^n$, $m < n$, wobei die Jacobi-Matrix $\mathbf{h}'(\mathbf{x})$ für jedes $\mathbf{x} \in D$ den Rang m hat.

Ist dann $\mathbf{x}_0 \in D$ eine lokale Extremalstelle unter Nebenbedingung $\mathbf{h}(\mathbf{x}) = \mathbf{0}$, so existiert dazu eine $(1 \times m)$-Matrix $\mathbf{L} = (\lambda_1, \ldots, \lambda_m)$ (Zeilenvektor) mit

$$f'(\mathbf{x}_0) + \mathbf{L}\mathbf{h}'(\mathbf{x}_0) = \mathbf{0} \quad \text{und} \quad \mathbf{h}(\mathbf{x}_0) = \mathbf{0}.$$

Die reellen $\lambda_1, \ldots, \lambda_m$ heißen Lagrange-Multiplikatoren.

Bemerkungen:

- Der Satz formuliert *notwendige Bedingung*: Falls $\mathbf{x}_0$ Extremum von f unter Nebenbedingung $\mathbf{h}(\mathbf{x}) = \mathbf{0}$ ist, so erfüllt $\mathbf{x}_0$ die Gleichung

$$F'(\mathbf{x}_0) + \mathbf{L}\mathbf{h}'(\mathbf{x}_0) = \mathbf{0} \quad \text{und} \quad \mathbf{h}(\mathbf{x}_0) = \mathbf{0}.$$

- Frage: Können wir Extremalstellen durch Lösung der Gleichung erhalten?

Idee: Mit den folgenden Vereinbarungen

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad \mathbf{h} = \begin{pmatrix} h_1 \\ \vdots \\ h_m \end{pmatrix}, \quad \text{und} \quad \mathbf{L} = (\lambda_1, \dots, \lambda_m),$$

Schreibt man

$$\frac{\partial f}{\partial x_j}(\mathbf{x}) + \sum_{k=1}^{m} \lambda_k \frac{\partial h_k}{\partial x_j} = 0, \quad \forall j = 1, \dots, n$$

und

$$h_k(\mathbf{x}) = 0 \quad \forall j = 1, \dots, m.$$

Also gibt es $n + m$ Gleichungen für $n + m$ Unbekannte $x_1, \dots, x_n, \lambda_1, \dots, \lambda_m$. Die Lösungen $\mathbf{x} = (x_1, \dots, x_n)^\top$ heißen stationäre oder kritische Punkte und sie sind die Kandidaten für Extremalstellen.

Satz: (Notwendige Extremalbedingung bei einer Nebenbedingung)
Seien $f : D \to \mathbb{R}$ und $g : D \to \mathbb{R}$ stetig partiell diff'bar auf offener Menge $D \subset \mathbb{R}^n$.
Es gelte $\nabla g(\mathbf{x}) \neq \mathbf{0}$ für jedes $\mathbf{x} \in D$.

Ist dann $\mathbf{x}_0 \in D$ eine lokale Extremalstelle von f unter der Nebenbedingung $g(\mathbf{x}) = \mathbf{0}$, so gilt

$$\nabla f(\mathbf{x}_0) + \lambda \nabla g(\mathbf{x}_0) = \mathbf{0}$$

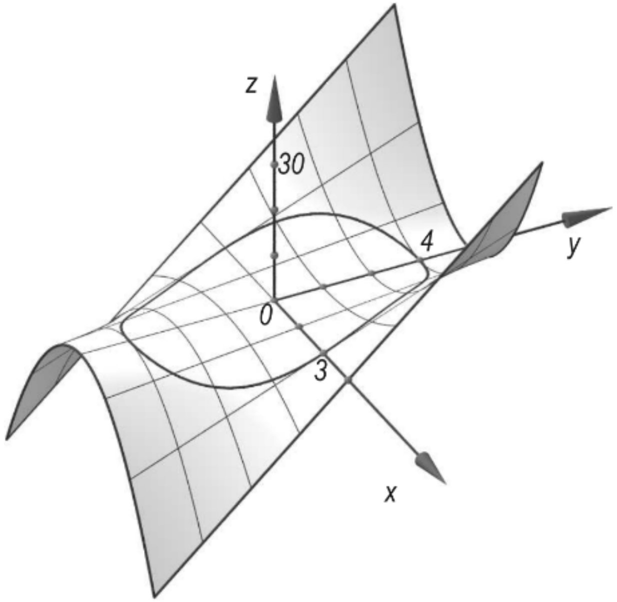
mit $\lambda \in \mathbb{R}$ Lagrange-Multiplikator.

Beispiel:
Betrachte

$$f(x, y) = x^2 y, \quad (x, y) \in \mathbb{R}^2$$

Gesucht sind Extremalstellen unter der Nebenbedingung

$$g(x, y) = \frac{x^2}{4} + \frac{y^2}{9} - 1 = 0.$$



1

Alternative: (Lagrange-Form)
Statt zwei Gleichungen

$$\frac{\partial f}{\partial x_j}(\mathbf{x}) + \sum_{k=1}^{m} \lambda_k \frac{\partial h_k}{\partial x_j} = 0, \quad \forall j = 1, \dots, n$$

und

$$h_k(\mathbf{x}) = 0 \quad \forall j = 1, \dots, m.$$

zu lösen, ist es auch möglich, die Lagrange-Form für $n+m$ Veränderliche $x_1, \dots, x_n, \lambda_1, \dots, \lambda_m$ zu formulieren:

$$L(\mathbf{x}, \lambda_1, \dots, \lambda_n) = f(\mathbf{x}) + \sum_{k=1}^{m} \lambda_k h_k(\mathbf{x}).$$

Dann ermittelt man die stationären Punkte durch Lösen von

$$\nabla L(\mathbf{x}, \lambda_1, \dots, \lambda_m) = \mathbf{0}.$$

Ausgleichsrechnung

Motivation: Gesucht ist ein funktionaler Zusammenhang

$$y = f(x_1, \dots, x_n).$$

Gegeben ist aber nicht f, sondern eine Reihe von *Messungen*

$$\begin{pmatrix} y_1 & x_{1_1} & \cdots & x_{1_n} \\ y_2 & x_{2_1} & \cdots & x_{2_n} \\ \vdots & & & \vdots \\ y_m & x_{m_1} & \cdots & x_{m_n} \end{pmatrix} =: (\mathbf{y}, \mathbf{x}_1, \dots, \mathbf{x}_n)$$

Ansatz: (lineare Theorie)
Vorabinformation über die gesuchte Funktion ist wahrscheinlich.
Nimm linearen Zusammenhang an:

$$f(x_1, \dots, x_n) = r_0 + r_1 x_1 + \cdots + r_n x_n$$

Methode: (Methode der kleinsten Fehlerquadrate)
Ziel: Finde $r_0, \dots, r_m$, die den quadratischen Fehler zwischen den gemessenen y_k und der Funktion $f(x_1, \dots, x_n)$ minimieren:

$$\begin{aligned} F(r_0, r_1, \dots, r_m) &= \sum_{k=1}^{m} (y_k - f(x_{k_1}, \dots, x_{k_n}))^2 \\ &= \sum_{k=1}^{m} (y_k - (r_0 + r_1 x_{k_1} + \cdots + r_n x_{k_n}))^2 \end{aligned}$$

Die Aufgabe lautet also (Methode der kleinsten Fehlerquadrate)

$$F(r_0, r_1, \dots, r_m) = \min!$$

Satz: (lineares Ausgleichsproblem)

1. Das lineare Ausgleichsproblem $F(r_0, \dots, r_m) = \min!$ ist immer lösbar.
2. Die Lösungen des Ausgleichsproblems und des zugehörigen Gleichungssystems $A\mathbf{r} = \mathbf{b}$ stimmen überein.
3. Ist der Rang der Matrix $(\mathbf{1}\ \mathbf{x}_1 \dots \mathbf{x}_n)$ gleich $n+1$, so ist die Ausgleichslösung eindeutig.
4. Falls $m \leq n$ (weniger Messungen als Parameter), so ist A singulär.

Lösungsidee:
Das Minimierungsproblem ist ein *Extremalproblem ohne Nebenbedingung*! Also löse

$$\nabla F = \mathbf{0}$$

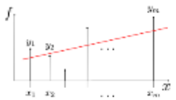
Dies führt auf die Gleichungen

$$\begin{aligned} \frac{\partial F}{\partial r_0} &= 2 \sum_{k=1}^{m} (y_k - (r_0 + r_1 x_{k_1} + \cdots + r_n x_{k_n})) \\ \frac{\partial F}{\partial r_1} &= 2 \sum_{k=1}^{m} (y_k - (r_0 + r_1 x_{k_1} + \cdots + r_n x_{k_n})) x_{k_1} \\ &\vdots \\ \frac{\partial F}{\partial r_n} &= 2 \sum_{k=1}^{m} (y_k - (r_0 + r_1 x_{k_1} + \cdots + r_n x_{k_n})) x_{k_n} \end{aligned}$$

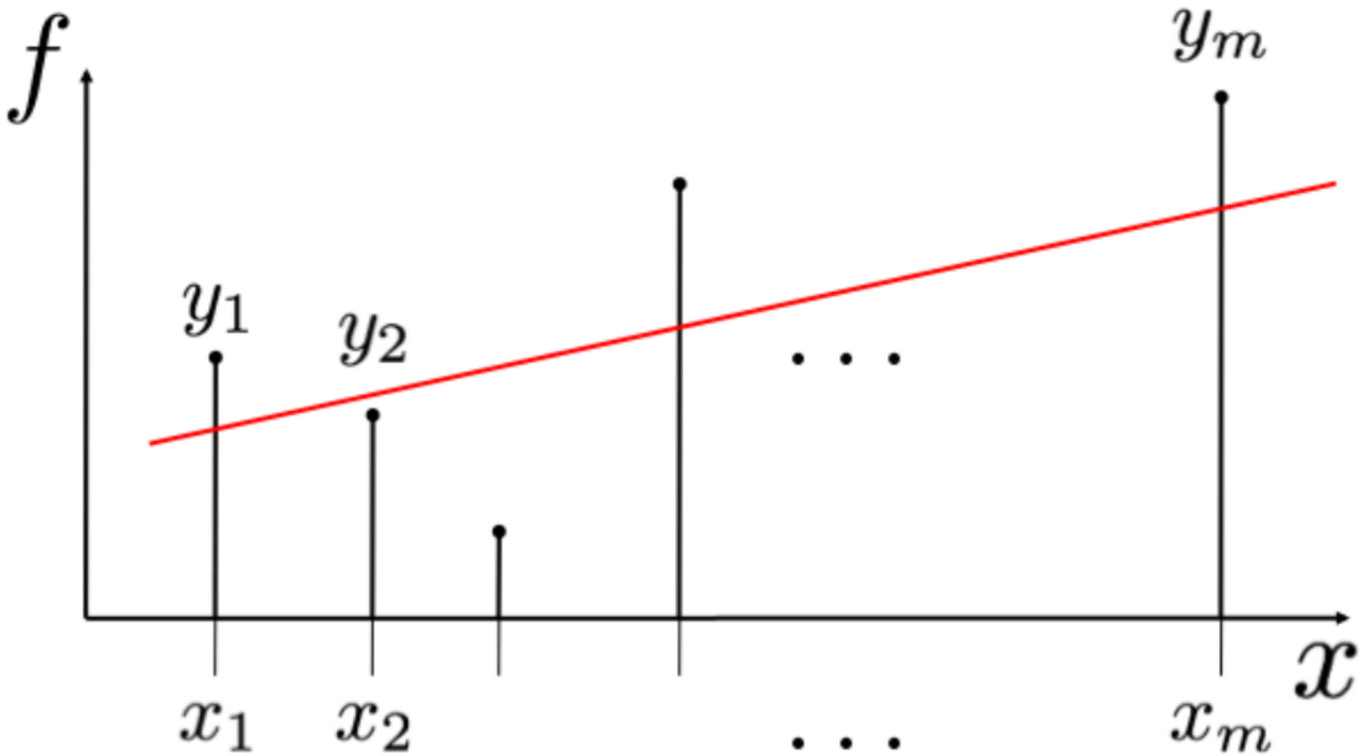
Motivation: Gesucht ist ein funktionaler Zusammenhang

$$y = f(x_1, \dots, x_n).$$

Gegeben ist aber nicht f, sondern eine Reihe von *Messungen*



$$\begin{pmatrix} y_1 & x_{1_1} & \cdots & x_{1_n} \\ y_2 & x_{2_1} & \cdots & x_{2_n} \\ \vdots & & & \vdots \\ y_m & x_{m_1} & \cdots & x_{m_n} \end{pmatrix} =: (\mathbf{y}, \mathbf{x}_1, \dots, \mathbf{x}_n)$$



f
y_1
y_2
y_m
$\cdots$
x
x_1
x_2
$\cdots$
x_m

Ansatz: (lineare Theorie)
Vorabinformation über die gesuchte Funktion ist wahrscheinlich.
Nimm linearen Zusammenhang an:

$$f(x_1, \ldots, x_n) = r_0 + r_1 x_1 + \cdots + r_n x_n$$

Methode: (Methode der kleinsten Fehlerquadrate)
Ziel: Finde $r_0, \ldots, r_m$, die den quadratischen Fehler zwischen den gemessenen y_k und der Funktion $f(x_1, \ldots, x_n)$ minimieren:

$$\begin{aligned} F(r_0, r_1, \ldots, r_m) &= \sum_{k=1}^{m} (y_k - f(x_{k_1}, \ldots, x_{k_n}))^2 \\ &= \sum_{k=1}^{m} (y_k - (r_0 + r_1 x_{k_1} + \cdots + r_n x_{k_n}))^2 \end{aligned}$$

Die Aufgabe lautet also (Methode der kleinsten Fehlerquadrate)

$$F(r_0, r_1, \ldots, r_m) = \min!$$

Lösungsidee:
Das Minimierungsproblem ist ein *Extremalproblem ohne Nebenbedingung*! Also löse

$$\nabla F = \mathbf{0}$$

Dies führt auf die Gleichungen

$$\begin{aligned}
\frac{\partial F}{\partial r_0} &= 2\sum_{k=1}^{m}(y_k - (r_0 + r_1 x_{k_1} + \cdots + r_n x_{k_n})) \\
\frac{\partial F}{\partial r_1} &= 2\sum_{k=1}^{m}(y_k - (r_0 + r_1 x_{k_1} + \cdots + r_n x_{k_n}))x_{k_1} \\
&\vdots \\
\frac{\partial F}{\partial r_n} &= 2\sum_{k=1}^{m}(y_k - (r_0 + r_1 x_{k_1} + \cdots + r_n x_{k_n}))x_{k_n}
\end{aligned}$$

Nach Umformung erhält man

$$A\mathbf{r} = \mathbf{b}$$

Nach Umformung erhält man

$$A\mathbf{r} = \mathbf{b}$$

für $\mathbf{r} = (r_0, \ldots, r_m)^\top$ mit $\mathbf{b} = (b_1, \ldots, b_m)^\top$ einem Vektor aus Beiträgen $b_i = \sum_{k=1}^m y_k x_{k_i}$ und

$$A = \begin{pmatrix} \sum_{k=1}^m 1 & \sum_{k=1}^m x_{k_1} & \cdots & \sum_{k=1}^m x_{k_n} \\ \sum_{k=1}^m x_{k_1} & \sum_{k=1}^m x_{k_1}^2 & \cdots & \sum_{k=1}^m x_{k_1} x_{k_n} \\ \vdots & & \ddots & \vdots \\ \sum_{k=1}^m x_{k_n} & \sum_{k=1}^m x_{k_1} x_{k_n} & \cdots & \sum_{k=1}^m x_{k_n}^2 \end{pmatrix}.$$

Satz: (lineares Ausgleichsproblem)

1. Das lineare Ausgleichsproblem $F(r_0, \dots, r_m) = \min!$ ist immer lösbar.
2. Die Lösungen des Ausgleichsproblems und des zugehörigen Gleichungssystems $A\mathbf{r} = \mathbf{b}$ stimmen überein.
3. Ist der Rang der Matrix $(\mathbf{1}\ \mathbf{x}_1 \dots \mathbf{x}_n)$ gleich $n+1$, so ist die Ausgleichslösung eindeutig.
4. Falls $m \le n$ (weniger Messungen als Parameter), so ist A singulär.

Extremalaufgaben (*mit* Nebenbedingungen)

Ausgleichsrechnung

Analysis III

Erinnerung: Extremalaufgaben (ohne Nebenbedingungen)