Die Methode der kleinsten Quadrate (Carl Friedrich Gauß, 1801)

Gegeben: Ein als bekannt vorausgesetzter Zusammenhang $\underline{\mathbf{y}} = \underline{\mathbf{A}} \cdot \underline{\mathbf{x}}$, mit:

$\underline{y}$: (m x 1)-Vektor mit gemessenen Werten y_i ($1 \le i \le m$)

$\underline{A}$: (m x n)-Matrix mit als bekannt angenommenen Werten a_{ij} ($1 \le i \le m$; $1 \le j \le n$; $m > n$)

$\underline{x}$: (n x 1)-Vektor mit den gesuchten Werten x_j ($1 \le j \le n$)

$\underline{A}$ und $\underline{y}$ seien potentiell fehlerbehaftet (Rundungen, Meßfehler, falsche Annahmen etc.), so daß die n $\underline{x}$-Werte nicht allen m Gleichungen gleichzeitig genügen können.

Gesucht: $\underline{x}$.

Da zur Ermittlung der n Elemente von $\underline{x}$ m Gleichungen verfügbar sind ($m > n$), soll durch Einbeziehung aller Gleichungen versucht werden, den Fehler in $\underline{x}$ zu minimieren.

Beispiel: Ermittlung der 3 Mischungsfaktoren $\underline{x}$ (3x1), die aus den RGB-Werten einer Farbbild-Spalte $\underline{A}$ (z.B.: 1024x3) die SW-Spalte $\underline{y}$ (z.B.: 1024x1) ergeben haben.

Gauß'scher Ansatz:

Der Vektor $\underline{x}$ soll so bestimmt werden, daß eine damit durchgeführte Proberechnung ($\underline{y}' = \underline{A} \cdot \underline{x}$) Werte y_i' liefert, die „möglichst nahe" bei den gemessenen Werten y_i liegen – konkreter: Der Mittelwert (bzw. die Summe) der quadrierten Abweichungen (Differenzen) zwischen errechneten $\underline{y}'$- und gemessenen $\underline{y}$-Elementen soll minimal werden. Damit sich positive und negative Abweichungen nicht gegenseitig aufheben, wird das Quadrat der Differenzen betrachtet. (Triviales Beispiel: Es gibt genau eine Gerade, bei der die Summe ihrer quadrierten Abstände von zwei gemessenen Punkten verschwindet – aber unendlich viele, deren Abstände sich zu Null summieren.)

Die Minimierungsforderung lautet also :

$$S = (a_{11} x_1 + ...+ a_{1n} x_n - y_1)^2 + ... + (a_{m1} x_1 + ... + a_{mn} x_n - y_m)^2 = \min !$$
$$= (\underline{A} \cdot \underline{x} - \underline{y})^T \cdot (\underline{A} \cdot \underline{x} - \underline{y})$$

Für die gesuchten x_i-Werte ergeben sich daraus die Teilforderunen:

$\partial S / \partial x_i = 0!$ und $\partial^2 S / \partial x_i^2 > 0$!

Diese Forderungen lauten ausgeschrieben :

$$\partial S / \partial x_1 = 2\, a_{11} (a_{11} x_1 + ...+ a_{1n} x_n - y_1) + ... + 2\, a_{m1} (a_{m1} x_1 + ... + a_{mn} x_n - y_m) = 0$$
(...)
$$\partial S / \partial x_n = 2\, a_{1n} (a_{11} x_1 + ...+ a_{1n} x_n - y_1) + ... + 2\, a_{mn} (a_{m1} x_1 + ... + a_{mn} x_n - y_m) = 0$$

und lassen sich (nach Wegkürzen der 2) als Matrizengleichung schreiben :

$\underline{0} = \underline{A}^T \cdot (\underline{A} \cdot \underline{x} - \underline{y}) \Rightarrow \underline{A}^T \cdot \underline{A} \cdot \underline{x} = \underline{A}^T \cdot \underline{y}$,

woraus folgt die Gaußsche Optimierungsformel: $\boxed{\underline{\mathbf{x}} = (\underline{\mathbf{A}}^T \cdot \underline{\mathbf{A}})^{-1} \cdot \underline{\mathbf{A}}^T \cdot \underline{\mathbf{y}}}$

Es handelt sich um ein Minimum, denn es gilt immer:

$\partial^2 S / \partial x_i^2 = 2\, a_{1i}^2 + (\ldots) + 2\, a_{mi}^2 > 0.$