# Die Methode der kleinsten Quadrate für Linearkombinationen von Funktionen

Zur Anpassung einer Linearkombinationen aus Funktionen

$$\sum_{i}^{p} a_i \cdot f_i(x) \tag{1}$$

an $N$ Datenpunkte $(x_j, y_j)$ kann die analytische Methode der kleinsten Quadrate verwendet werden. Ziel ist es, die Parameter $a_i$ zu bestimmen, bei denen die Summe

$$\sum_{j}^{N} \left( y_j - \sum_{i}^{p} a_i \cdot f_i(x_j) \right)^2 \tag{2}$$

minimal ist, der Funktionsgraph also minimal von den Datenpunkten abweicht.

## Bestimmung der Parameter

1. Zuerst wird die sogenannte Designmatrix $\boldsymbol{A}$ aufgestellt. Sie enthält die Funktionswerte für jede Funktion $f_i$ ausgewertet an den gemessenen $x_j$:

$$\boldsymbol{A} = \begin{pmatrix} f_1(x_1) & f_2(x_1) & \cdots & f_p(x_1) \\ f_1(x_2) & f_2(x_2) & \cdots & f_p(x_2) \\ \vdots & & & \vdots \\ f_1(x_N) & f_2(x_N) & \cdots & f_p(x_N) \end{pmatrix} . \tag{3}$$

2. Als nächstes definieren wir den Spalten-Vektor

$$\vec{y} = (y_1, y_2, ..., y_N)^\top , \tag{4}$$

   der die $y$-Koordinaten unserer Messwerte enthält.

3. Die Kovarianzmatrix der Messwerte sei $\boldsymbol{W}$ und $\boldsymbol{Z} = \boldsymbol{W}^{-1}$.

4. der Parametervektor $\vec{a}$ ergibt sich dann zu:

$$\vec{a} = \left(\boldsymbol{A}^\top \boldsymbol{Z} \boldsymbol{A}\right)^{-1} \cdot \boldsymbol{A}^\top \boldsymbol{Z} \cdot \vec{y} \tag{5}$$

5. Die Kovarianzmatrix ergibt sich zu:

$$\boldsymbol{V} = \left(\boldsymbol{A}^\top \boldsymbol{Z} \boldsymbol{A}\right)^{-1} \tag{6}$$

Wer sich für die Herleitung interessiert: Blobel-Lohrmann: „Statistische und numerische Methoden der Datenanalyse“ bzw. SMD-Vorlesung. Alternativ auch bei Wikipedia: `https://en.wikipedia.org/wiki/Linear_least_squares_(mathematics)`

## Aufgaben

1. Implementiert eine Funktion, die eine Liste von Funktionen und die $x$- und $y$-Werte übergeben bekommt und die lineare Methode der kleinsten Quadrate anwendet. Die Funktion soll den Parametervektor als Array von korrelierten `ufloat`s zurückgeben.

2. Testet eure Funktion indem ihr einen Fit der Form

   $$\Psi(x) = a_1 \cos(x) + a_2 \sin(x) \tag{7}$$

   an die Daten in der Datei `daten.txt` durchführt. Die Unsicherheiten der $y_i$ sollen als $\sigma = 0{,}1$ angenommen werden.

## Tipps

- Matrix-Multiplikation ziwschen numpy arrays mit dem `@` Operator
- Zum invertieren einer Matrix kann `numpy.linalg.inv` genutzt werden.