

# Die Methode der kleinsten Quadrate für Linearkombinationen von Funktionen

Zur Anpassung einer Linearkombinationen aus Funktionen

$$\sum_i^p a_i \cdot f_i(x) \quad (1)$$

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

$$\sum_j^N \left( y_j - \sum_i^p a_i \cdot f_i(x_j) \right)^2 \quad (2)$$

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

## Bestimmung der Parameter

1. Zuerst wird die sogenannte Designmatrix  $A$  aufgestellt, sie enthält die Funktionswerte für jede Funktion  $f_i$  ausgewertet an den gemessenen  $x_j$ :

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

2. Als nächstes definieren wir den Spalten-Vektor

$$\vec{y} = (y_1, y_2, \dots, y_N)^T, \quad (4)$$

der die  $y$ -Koordinaten unserer Messwerte enthält.

3. Die Kovarianzmatrix der Messwerte sei  $W$  und  $Z = W^{-1}$ .
4. der Parametervektor  $\vec{a}$  ergibt sich dann zu:

$$\vec{a} = (A^T Z A)^{-1} \cdot A^T Z \cdot \vec{y} \quad (5)$$

5. Die Kovarianzmatrix ergibt sich zu:

$$V = (A^T Z A)^{-1} \quad (6)$$

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\\_\(mathematics\)](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 `ufloats` zurückgeben.
2. Testet eure Funktion indem ihr einen Fit der Form

$$\Psi(x) = a_1 \cos(x) + a_2 \sin(x) \quad (7)$$

an die Daten in der Datei `daten.txt` durchführt. Der Fehler der  $y_i$  soll als  $\sigma = 0,1$  angenommen werden.

## Tipps

Hilfreich könnte die NumPy-Datenstruktur `matrix` mit ihren Methoden

- `.T`,
- `.I`,
- `.flat`

sein.