# Kurvenanpassung durch Regression (2)

## - polynomiale Regression -

Ac 2017

**Problemstellung:**

Es sind n Wertepaare (Messwerte bzw. Punkte) **Pi (xi ; yi)** gegeben.
Gesucht ist eine ganzrationale Funktion f n-ter Ordnung, die sich den Punkten Pi möglichst gut anpasst ( polynomiale Regression ) .

Beispiel (n=3): $f(x) = ax^3 + bx^2 + cx + d$

Gelöst wird für jede der ganzrationalen Funktionen das lineare Gleichungssystem

$$\boxed{\frac{dS}{da}=0 \quad \wedge \quad \frac{dS}{db}=0 \quad \wedge \quad \frac{dS}{dc}=0 \quad \wedge \; \ldots \qquad mit \qquad S=\sum_{i=1}^{n}(y_i - f(x_i))^2}$$

① Quadratische Regression: $f(x) = ax^2 + bx + c$

Die oben definierte Summe der Quadrate ist : $\boxed{S=\sum_{i=1}^{n}(y_i-(ax_i^2+bx_i+c))^2}$

Für die partiellen Ableitungen ergibt sich:

$$\left| \begin{aligned} \frac{dS}{da} &= \sum_{i=1}^{n} 2(y_i-(ax_i^2+bx_i+c))\cdot(-x_i^2) \\ \frac{dS}{db} &= \sum_{i=1}^{n} 2(y_i-(ax_i^2+bx_i+c))\cdot(-x_i) \\ \frac{dS}{dc} &= \sum_{i=1}^{n} 2(y_i-(ax_i^2+bx_i+c))\cdot(-1) \end{aligned} \right|$$

Setzt man diese gleich 0, so erhält man ein lineares Gleichungssystem (LGS):

$$\left| \begin{aligned} 0 &= \sum_{i=1}^{n}(-x_i^2y_i+ax_i^4+bx_i^3+cx_i^2)) \\ 0 &= \sum_{i=1}^{n}(-x_iy_i+ax_i^3+bx_i^2+cx_i) \\ 0 &= \sum_{i=1}^{n}(-y_i+ax_i^2+bx_i+c) \end{aligned} \right| \Leftrightarrow$$

**LGS für QuadReg**

$$\left| \begin{aligned} a\cdot\sum_{i=1}^{n}x_i^4+b\cdot\sum_{i=1}^{n}x_i^3+c\cdot\sum_{i=1}^{n}x_i^2 &= \sum_{i=1}^{n}x_i^2y_i \\ a\cdot\sum_{i=1}^{n}x_i^3+b\cdot\sum_{i=1}^{n}x_i^2+c\cdot\sum_{i=1}^{n}x_i &= \sum_{i=1}^{n}x_iy_i \\ a\cdot\sum_{i=1}^{n}x_i^2+b\cdot\sum_{i=1}^{n}x_i+c\cdot n &= \sum_{i=1}^{n}y_i \end{aligned} \right|$$

Dies löst man z.B. mit dem Gaußschen Eliminationsverfahren !

Beispiel: (1 ; 5) (2 ; 20,1) (3 ; 44,9) (4 ; 80,2)

sum(xi)=10 sum(xi²)=30 sum(xi³)=100 sum(xi^4)=354
sum(yi) =150,2 sum(xiyi)=500,7 sum(xi²yi)=1772,7 n=4
LGS: 354a + 100b + 30c = 1772,7 100a + 30b + 10c = 500,7 30a + 10b + 4c = 150,2
Lösung des LGS: a = 5,05 b = -0,21 c = 0,2 ==> **f(x) = 5,05x² -0,21x + 0,2**

## ② Kubische Regression: f(x) = ax³ + bx² + cx + d

Die oben definierte Summe der Quadrate ist :

$$\boxed{S=\sum_{i=1}^{n}(y_i-(ax_i^3+bx_i^2+cx_i+d))^2}$$

Für die partiellen Ableitungen ergibt sich:

$$\left|\begin{aligned}
\frac{dS}{da}&=\sum_{i=1}^{n}2(y_i-(ax_i^3+bx_i^2+cx_i+d))\cdot(-x_i^3)\\
\frac{dS}{db}&=\sum_{i=1}^{n}2(y_i-(ax_i^3+bx_i^2+cx_i+d))\cdot(-x_i^2)\\
\frac{dS}{dc}&=\sum_{i=1}^{n}2(y_i-(ax_i^3+bx_i^2+cx_i+d))\cdot(-x_i)\\
\frac{dS}{dd}&=\sum_{i=1}^{n}2(y_i-(ax_i^3+bx_i^2+cx_i+d))\cdot(-1)
\end{aligned}\right|$$

Setzt man diese gleich 0, so erhält man ein lineares Gleichungssystem (LGS):

$$\left|\begin{aligned}
0&=\sum_{i=1}^{n}(-x_i^3y_i+ax_i^6+bx_i^5+cx_i^4+dx_i^3))\\
0&=\sum_{i=1}^{n}(-x_i^2y_i+ax_i^5+bx_i^4+cx_i^3+dx_i^2))\\
0&=\sum_{i=1}^{n}(-x_iy_i+ax_i^4+bx_i^3+cx_i^2+dx_i))\\
0&=\sum_{i=1}^{n}(-y_i+ax_i^3+bx_i^2+cx_i+d))
\end{aligned}\right| \Leftrightarrow$$

**LGS für KubikReg**

$$\left|\begin{aligned}
a\cdot\sum_{i=1}^{n}x_i^6+b\cdot\sum_{i=1}^{n}x_i^5+c\cdot\sum_{i=1}^{n}x_i^4+d\cdot\sum_{i=1}^{n}x_i^3&=\sum_{i=1}^{n}x_i^3y_i\\
a\cdot\sum_{i=1}^{n}x_i^5+b\cdot\sum_{i=1}^{n}x_i^4+c\cdot\sum_{i=1}^{n}x_i^3+d\cdot\sum_{i=1}^{n}x_i^2&=\sum_{i=1}^{n}x_i^2y_i\\
a\cdot\sum_{i=1}^{n}x_i^4+b\cdot\sum_{i=1}^{n}x_i^3+c\cdot\sum_{i=1}^{n}x_i^2+d\cdot\sum_{i=1}^{n}x_i&=\sum_{i=1}^{n}x_iy_i\\
a\cdot\sum_{i=1}^{n}x_i^3+b\cdot\sum_{i=1}^{n}x_i^2+c\cdot\sum_{i=1}^{n}x_i+d\cdot n&=\sum_{i=1}^{n}y_i
\end{aligned}\right|$$

Auffällig ist die strukturelle Ähnlichkeit mit dem LGS für QuadReg !
Die Lösung erfolgt auch hier mit dem Gaußschen Eliminationsverfahren !

**Verallgemeinerung:** Ganzrat. Fkt. k-ten Grades

**LGS für $f(x)=a_kx^k+a_{k-1}x^{k-1}+\ldots+a_2x^2+a_1x+a_0$**

$$\left|\begin{aligned}
a_k\cdot\sum_{i=1}^{n}x_i^{2k}+a_{k-1}\cdot\sum_{i=1}^{n}x_i^{2k-1}+\ \ldots\ +a_1\cdot\sum_{i=1}^{n}x_i^{2k-(k-1)}+a_0\cdot\sum_{i=1}^{n}x_i^{2k-k}&=\sum_{i=1}^{n}x_i^ky_i\\
a_k\cdot\sum_{i=1}^{n}x_i^{2k-1}+a_{k-1}\cdot\sum_{i=1}^{n}x_i^{2k-2}+\ \ldots\ +a_1\cdot\sum_{i=1}^{n}x_i^{2k-k}+a_0\cdot\sum_{i=1}^{n}x_i^{2k-(k+1)}&=\sum_{i=1}^{n}x_i^ky_i\\
\ldots\\
a_k\cdot\sum_{i=1}^{n}x_i^{2k-k}+a_{k-1}\cdot\sum_{i=1}^{n}x_i^{2k-(k+1)}+\ \ldots\ +a_1\cdot\sum_{i=1}^{n}x_i^{2k-(k+k-1)}+a_0\cdot\sum_{i=1}^{n}x_i^{2k-(k+k)}&=\sum_{i=1}^{n}x_i^ky_i
\end{aligned}\right|$$

Beispiel (Polynom 4.Grades):

Gegebene Daten (Testdaten):

$x_i$ = [1.0, 2.0, 3.0, 4.0, 5.0, 6.0, 7.0]
$y_i$ = [4.0, 10.0, 45.0, 157.0, 418.0, 924.0, 1795.0]

Folgende (5 ; 6) - Matrix für das LGS entsteht:

| | | | | | |
|---|---|---|---|---|---|
| 7907396 | 1200304 | 184820 | 29008 | 4676 | 5812550 |
| 1200304 | 184820 | 29008 | 4676 | 784 | 878866 |
| 184820 | 29008 | 4676 | 784 | 140 | 134630 |
| 29008 | 4676 | 784 | 140 | 28 | 20986 |
| 4676 | 784 | 140 | 28 | 7 | 3353 |

Das gesuchte Polynom 4.Grades ist:

$p_4(x) = x^4 - 2x^3 + 1{,}5x^2 + 0{,}5x + 3$