(Polynom)Regression von Datenpunkten

Dipl.- Ing. Björnstjerne Zindler, M.Sc.

www.Zenithpoint.de

Erstellt: 19. Juli 2013 – Letzte Revision: 24. Januar 2021

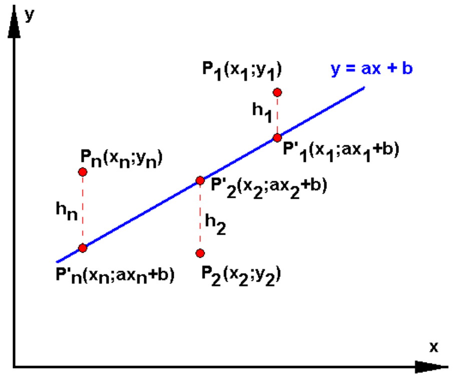
Inhaltsverzeichnis

Literatur

[001] Keine für vorliegenden Text.

1 Die Methode der kleinsten Quadrate anhand eines Beispiels

1.1 Modell zur Berechnung einer Lösung für den linearen Fall



Modell zur Berechnung einer Linearen Regression nach der Methode der kleinsten Quadrate.

Die Methode der kleinsten Quadrate ist ein Verfahren in der Regressionsrechnung. Dabei wird eine Funktion gesucht, welche möglichst nahe der Datenpunkte verläuft. [001]ff.

Das Vorgehen geht dahin hinaus, das eine Mismatch-, Fehl oder auch Straffunktion $F\,(x; a; b)$ gesucht wird, welche die Koeffizienten a und b im Punkt des lokalen (und auch globalen) Minimum liefert.

Der Abstand h als Strecke zwischen den Punkten P und P_i' (parallel zur Ordinate).

$$\begin{aligned}
h_1^2 &= (x_1 - x_1)^2 + (y_1 - (a \cdot x_1 + b))^2 \\
h_2^2 &= (x_2 - x_2)^2 + (y_2 - (a \cdot x_2 + b))^2 \\
&\vdots \\
h_n^2 &= (x_n - x_n)^2 + (y_n - (a \cdot x_n + b))^2
\end{aligned}$$

$\Rightarrow$

$$\begin{aligned}
h_1^2 &= (y_1 - a \cdot x_1 - b)^2 \\
h_2^2 &= (y_2 - a \cdot x_2 - b)^2 \\
&\vdots \\
h_n^2 &= (y_n - a \cdot x_n - b)^2
\end{aligned}$$

Die Fehlfunktion ist definiert durch:

$$F = \sum_{i=1}^{n} h_i^2$$

$\Rightarrow$

$$F = h_1^2 + h_2^2 + \ldots + h_n^2$$

Damit ergibt sich für F:

$$F = \begin{array}{r} (y_1 - a \cdot x_1 - b)^2 \\ +\,(y_2 - a \cdot x_2 - b)^2 \\ \vdots \\ +\,(y_n - a \cdot x_n - b)^2 \end{array}$$

$\Rightarrow$

$$F = \begin{array}{r} y_1^2 - 2a \cdot x_1 y_1 - 2b \cdot y_1 + 2ab \cdot x_1 + a^2 \cdot x_1^2 + b^2 \\ +y_2^2 - 2a \cdot x_2 y_2 - 2b \cdot y_2 + 2ab \cdot x_2 + a^2 \cdot x_2^2 + b^2 \\ \vdots \\ +y_n^2 - 2a \cdot x_n y_n - 2b \cdot y_n + 2ab \cdot x_n + a^2 \cdot x_n^2 + b^2 \end{array}$$

$\Rightarrow$

$$F = \left\{y^2\right\} - 2a \cdot \{xy\} - 2b \cdot \{y\} + 2ab \cdot \{x\} + a^2 \cdot \left\{x^2\right\} + n \cdot b^2$$

1.2 Suche nach einem Extrema

Um das Extrema, das vermutete Minimum zu finden muss $F(a;b)$ partiell differenziert werden (da
[001]ff. jetzt zweidimensional).

$$\frac{\partial}{\partial a}F = 0 \qquad \frac{\partial}{\partial b}F = 0$$

$\Rightarrow$

$$\frac{\partial}{\partial a}F = 0 - 2\cdot\{xy\} - 0 + 2b\cdot\{x\} + 2a\cdot\left\{x^2\right\} + 0 = 0$$

$$\frac{\partial}{\partial b}F = 0 - 0 - 2\cdot\{y\} + 2a\cdot\{x\} + 0 + 2n\cdot b = 0$$

$\Rightarrow$

$$2a\cdot\left\{x^2\right\} + 2b\cdot\{x\} - 2\cdot\{xy\} = 0$$

$$2a\cdot\{x\} + 2b\cdot n - 2\cdot\{y\} = 0$$

Über den Gaußalgorithmus wird das vorliegende Gleichungssystem gelöst.

$$a + b\cdot\frac{\{x\}}{\{x^2\}} - \frac{\{xy\}}{\{x^2\}} = 0$$

$$a + b\cdot\frac{n}{\{x\}} - \frac{\{y\}}{\{x\}} = 0$$

$\Rightarrow$

$$b\cdot\left(\frac{\{x\}}{\{x^2\}} - \frac{n}{\{x\}}\right) - \left(\frac{\{xy\}}{\{x^2\}} - \frac{\{y\}}{\{x\}}\right) = 0$$

$\Rightarrow$

$$b = \frac{\{xy\}\cdot\{x\} - \{y\}\cdot\left\{x^2\right\}}{\{x\}^2 - n\cdot\{x^2\}}$$

Um a zu ermitteln wird nicht rückwärts eingesetzt, sondern folgender Ausdruck genutzt:

$$\{y\} = a\cdot\{x\} + n\cdot b$$

$\Rightarrow$

$$a = \frac{\{y\}}{\{x\}} - \frac{n}{\{x\}}\cdot b$$

$\Rightarrow$

$$a = \frac{\{y\}}{\{x\}} - \frac{n}{\{x\}}\cdot\frac{\{xy\}\cdot\{x\} - \{y\}\cdot\left\{x^2\right\}}{\{x\}^2 - n\cdot\{x^2\}}$$

$\Rightarrow$

$$a = \frac{\{x\}\cdot\{y\} - n\cdot\{xy\}}{\{x\}^2 - n\cdot\{x^2\}}$$

Dabei gilt:

$$\{x\} = \sum_{i=1}^{n} x_i \qquad \{y\} = \sum_{i=1}^{n} y_i$$

$$\left\{x^2\right\} = \sum_{i=1}^{n}(x_i\cdot x_i) \qquad \{xy\} = \sum_{i=1}^{n}(x_i\cdot y_i)$$

1.3 Nachweis des Extrema als Minimum

Damit nachgewiesen ist, dass das gefundene Extrema ein Minimum ist muss gelten: [001]ff.

• $\frac{\partial}{\partial a}F = 0$

Wurde im vorherigen Abschnitt erfüllt.

• $\frac{\partial}{\partial b}F = 0$

Wurde im vorherigen Abschnitt erfüllt.

• $\frac{\partial}{\partial a}\frac{\partial}{\partial a}F > 0$

Nachweis das Extrema ein Minimum ist.

$$\frac{\partial}{\partial a}\frac{\partial}{\partial a}F = 2 \cdot \{x^2\} > 0$$

Ist immer erfüllt für ab zwei unterscheidbare Datenpunkte.

• $\frac{\partial}{\partial b}\frac{\partial}{\partial b}F > 0$

Nachweis das Extrema ein Minimum ist.

$$\frac{\partial}{\partial b}\frac{\partial}{\partial b}F = 2 \cdot n = 0$$

Ist immer erfüllt.

• Eine Diskriminante D ist $D > 0$

$$D = \left(\frac{\partial}{\partial a}\frac{\partial}{\partial a}F\right)^2 \cdot \left(\frac{\partial}{\partial b}\frac{\partial}{\partial b}F\right)^2 - \left(\frac{\partial}{\partial a}\frac{\partial}{\partial b}F\right)^2 > 0$$

Mit:

$$\frac{\partial}{\partial a}\frac{\partial}{\partial b}F = \frac{\partial}{\partial b}\frac{\partial}{\partial a}F = 2\{x\}$$

Gilt:

$$D = 4 \cdot \{x^2\}^2 \cdot 4 \cdot n^2 - 4 \cdot \{x\}^2 > 0$$

$\Rightarrow$

$$n > \frac{1}{2} \cdot \frac{\{x\}}{\{x^2\}}$$

Ein Nachweis obigen Ausdrucks ist zahlentheoretisch möglich. Wenn

$$x \in N$$

ist ein Nachweis schnell gezeigt.

$$\{x\} = \sum_{i=1}^{n} k = \frac{n}{2} \cdot (n+1) \qquad \{x^2\} = \sum_{i=1}^{n} k^2 = \frac{n}{6} \cdot (n+1) \cdot (2n+1)$$

$\Rightarrow$

$$n > \frac{1}{2} \cdot \frac{6 \cdot n \cdot (n+1)}{2 \cdot n \cdot (n+1) \cdot (2n+1)}$$

$\Rightarrow$

$$n^2 + \frac{1}{2} \cdot n - \frac{3}{4} > 0$$

Für alle $n > 1$ ist diese quadratische Ungleichung erfüllt und F befindet sich im Minimum.

1.4 Beschreibung des Problems in Matrizenform

[001]ff. Aus den vorhergehenden Abschnitten ist bekannt.

$$b \cdot n + a \cdot \{x\} = \{y\}$$
$$b \cdot \{x\} + a \cdot \left\{x^2\right\} = \{xy\}$$

Die dazu gehörige Matrize:

$$\begin{pmatrix} n & \{x\} \\ \{x\} & \{x \cdot x\} \end{pmatrix} \cdot \begin{pmatrix} b \\ a \end{pmatrix} = \begin{pmatrix} \{y\} \\ \{x \cdot y\} \end{pmatrix}$$

Lösungsmöglichkeiten sind vorhanden, eine Verallgemeinerung für Regressionspolynome der Form

$$P(x) = a_n \cdot x^n + a_{n-1} \cdot x^{n-1} + \ldots + a_2 \cdot x^2 + a_1 \cdot x + a_0$$

ist hierüber möglich.

2 Regression n- ter Ordnung

Gesucht ist das Regressionspolynom $P(\lambda)$.

$$P(\lambda) = a_n \cdot \lambda^n + a_{n-1} \cdot \lambda^{n-1} + \ldots + a_2 \cdot \lambda^2 + a_1 \cdot \lambda + a_0 \qquad (a_n \neq 0)$$

Wobei n der Grad oder die Ordnung des Polynoms $P(\lambda)$ ist.

2.1 Regression n- ter Ordnung über die Determinanten

[001]ff. Um das Polynom $P(\lambda)$ ermitteln zu können, wird zuerst eine quadratische Matrix aufgestellt:

$$M = \begin{pmatrix} n & \{x\} & \{x^2\} & \cdots & \{x^n\} \\ \{x\} & \{x^2\} & \{x^3\} & \cdots & \{x^{n+1}\} \\ \{x^2\} & \{x^3\} & \{x^4\} & \cdots & \{x^{n+2}\} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \{x^n\} & \{x^{n+1}\} & \{x^{n+2}\} & \cdots & \{x^{2\cdot n}\} \end{pmatrix}$$

Zusätzlich wird eine Spaltenmatrix benötigt.

$$V = \begin{pmatrix} \{y\} \\ \{x\cdot y\} \\ \{x^2\cdot y\} \\ \vdots \\ \{x^n\cdot y\} \end{pmatrix}$$

Nächster Schritt, jedem Koeffizienten a_i wird eine quadratische Matrix zugeordnet. Diese Matrix ist grundsätzlich identisch zu M. Unterschied, die $i+1$ Spalte wird mit den Elementen von V ersetzt. So gilt beispielshalber:

$$a_0 \leftrightarrow \begin{pmatrix} \{y\} & \{x\} & \{x^2\} & \cdots & \{x^n\} \\ \{x\cdot y\} & \{x^2\} & \{x^3\} & \cdots & \{x^{n+1}\} \\ \{x^2\cdot y\} & \{x^3\} & \{x^4\} & \cdots & \{x^{n+2}\} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \{x^n\cdot y\} & \{x^{n+1}\} & \{x^{n+2}\} & \cdots & \{x^{2\cdot n}\} \end{pmatrix}$$

$$a_1 \leftrightarrow \begin{pmatrix} n & \{y\} & \{x^2\} & \cdots & \{x^n\} \\ \{x\} & \{x\cdot y\} & \{x^3\} & \cdots & \{x^{n+1}\} \\ \{x^2\} & \{x^2\cdot y\} & \{x^4\} & \cdots & \{x^{n+2}\} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \{x^n\} & \{x^n\cdot y\} & \{x^{n+2}\} & \cdots & \{x^{2\cdot n}\} \end{pmatrix}$$

$$a_2 \leftrightarrow \begin{pmatrix} n & \{x\} & \{y\} & \cdots & \{x^n\} \\ \{x\} & \{x^2\} & \{x\cdot y\} & \cdots & \{x^{n+1}\} \\ \{x^2\} & \{x^3\} & \{x^2\cdot y\} & \cdots & \{x^{n+2}\} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \{x^n\} & \{x^{n+1}\} & \{x^n\cdot y\} & \cdots & \{x^{2\cdot n}\} \end{pmatrix}$$

$$\vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots$$

$$a_n \leftrightarrow \begin{pmatrix} n & \{x\} & \{x^2\} & \cdots & \{y\} \\ \{x\} & \{x^2\} & \{x^3\} & \cdots & \{x\cdot y\} \\ \{x^2\} & \{x^3\} & \{x^4\} & \cdots & \{x^2\cdot y\} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \{x^n\} & \{x^{n+1}\} & \{x^{n+2}\} & \cdots & \{x^n\cdot y\} \end{pmatrix}$$

Wenn nun die Determinanten von M und $a_i \leftrightarrow$ errechnet werden, dann sind die Koeffizienten a_i von $P(\lambda)$ ermittelt.

$$a_0 = \frac{\det(a_0 \leftrightarrow)}{\det(M)} = \frac{|a_0 \leftrightarrow|}{|M|} \qquad a_1 = \frac{\det(a_1 \leftrightarrow)}{\det(M)} = \frac{|a_1 \leftrightarrow|}{|M|}$$

$$a_2 = \frac{\det\left(a_2 \leftrightarrow\right)}{\det\left(M\right)} = \frac{|a_2 \leftrightarrow|}{|M|} \qquad a_n = \frac{\det\left(a_n \leftrightarrow\right)}{\det\left(M\right)} = \frac{|a_n \leftrightarrow|}{|M|}$$

Dabei gilt aus den gegebenen Datenpunkten:

$$\{x\} = \sum_{i=1}^{n} \lambda_i \qquad \{y\} = \sum_{i=1}^{n} y_i$$

$$\left\{x^2\right\} = \sum_{i=1}^{n} \left(\lambda_i \cdot \lambda_i\right) \qquad \{x \cdot y\} = \sum_{i=1}^{n} \left(\lambda_i \cdot y_i\right)$$

$$\left\{x^3\right\} = \sum_{i=1}^{n} \left(\lambda_i \cdot \lambda_i \cdot \lambda_i\right) \qquad \left\{x^2 \cdot y\right\} = \sum_{i=1}^{n} \left(\lambda_i \cdot \lambda_i \cdot y_i\right)$$

$$\left\{x^4\right\} = \sum_{i=1}^{n} \left(\lambda_i \cdot \lambda_i \cdot \lambda_i \cdot \lambda_i\right) \qquad \cdots\cdots\cdots$$

$$\cdots\cdots\cdots \qquad \cdots\cdots\cdots$$

2.2 Regression n- ter Ordnung über den Gaußalgorithmus

[001]ff. Um das Polynom $P(\lambda)$ ermitteln zu können, wird folgendes Gleichungssystem aufgestellt:

$$P(\lambda) = a_n \cdot \lambda^n + a_{n-1} \cdot \lambda^{n-1} + \ldots + a_2 \cdot \lambda^2 + a_1 \cdot \lambda + a_0$$

$\Rightarrow$

$$\begin{pmatrix} n & \{x\} & \{x^2\} & \cdots & \{x^n\} \\ \{x\} & \{x^2\} & \{x^3\} & \cdots & \{x^{n+1}\} \\ \{x^2\} & \{x^3\} & \{x^4\} & \cdots & \{x^{n+2}\} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \{x^n\} & \{x^{n+1}\} & \{x^{n+2}\} & \cdots & \{x^{2\cdot n}\} \end{pmatrix} \cdot \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} \{y\} \\ \{x \cdot y\} \\ \{x^2 \cdot y\} \\ \vdots \\ \{x^n \cdot y\} \end{pmatrix}$$

$\Rightarrow$

$$\begin{aligned} a_0 \cdot n + a_1 \cdot \{x\} + a_2 \cdot \{x^2\} + \ldots + a_n \cdot \{x^n\} &= \{y\} \\ a_0 \cdot \{x\} + a_1 \cdot \{x^2\} + a_2 \cdot \{x^3\} + \ldots + a_n \cdot \{x^{n+1}\} &= \{x \cdot y\} \\ a_0 \cdot \{x^2\} + a_1 \cdot \{x^3\} + a_2 \cdot \{x^4\} + \ldots + a_n \cdot \{x^{n+2}\} &= \{x^2 \cdot y\} \\ \vdots \qquad\qquad \vdots \qquad\qquad & \vdots \\ a_0 \cdot \{x^n\} + a_1 \cdot \{x^{n+1}\} + a_2 \cdot \{x^{n+2}\} + \ldots + a_n \cdot \{x^{2n}\} &= \{x^n \cdot y\} \end{aligned}$$

Durch Lösen des Gleichungssystems ist das Polynom bestimmt. Dabei gilt aus den gegebenen Datenpunkten:

$$\{x\} = \sum_{i=1}^{n} \lambda_i \qquad \{y\} = \sum_{i=1}^{n} y_i$$

$$\{x^2\} = \sum_{i=1}^{n} (\lambda_i \cdot \lambda_i) \qquad \{x \cdot y\} = \sum_{i=1}^{n} (\lambda_i \cdot y_i)$$

$$\{x^3\} = \sum_{i=1}^{n} (\lambda_i \cdot \lambda_i \cdot \lambda_i) \qquad \{x^2 \cdot y\} = \sum_{i=1}^{n} (\lambda_i \cdot \lambda_i \cdot y_i)$$

$$\{x^4\} = \sum_{i=1}^{n} (\lambda_i \cdot \lambda_i \cdot \lambda_i \cdot \lambda_i) \qquad \cdots\cdots\cdots$$

$$\cdots\cdots\cdots \qquad\qquad \cdots\cdots\cdots$$

2.3 Probleme mit dem Ausdruck $\{x \cdot x \cdot x \cdot x \cdot \ldots\}$

Bei großen Datenmengen, sehr kleinen x- Werten aber auch sehr großen x- Werten kann es Probleme mit dem $\{x \cdot x \cdot x \cdot x \cdot \ldots\}$- Ausdruck geben. Von Rundungsfehlern, Divisionsproblemen bis Zahlenbereichsüberflüsse in verschiedenen Programmiersprachen ergeben sich unter Umständen Probleme bei der Berechnung des Polynoms. [001]ff.

Eine Möglichkeit diese Probleme zu umgehen, ist das Skalieren der Datenpunkte mit ε hin zu günstigen Werten.

$$P_n\left(\lambda_n; X_n\right) \quad \rightarrow \quad P_n\left(\varepsilon \cdot \lambda_n; \varepsilon \cdot X_n\right) \quad \rightarrow \quad P_n\left(\tilde{\lambda}_n; \tilde{X}_n\right)$$

Das Polynom lautet dann:

$$X = a_n \cdot \tilde{\lambda}^n + a_{n-1} \cdot \tilde{\lambda}^{n-1} + a_{n-2} \cdot \tilde{\lambda}^{n-2} + \ldots + a_2 \cdot \tilde{\lambda}^2 + a_1 \cdot \tilde{\lambda} + a_0$$

Eine Rückskalierung ergibt die exakte Regression.

$$X = \varepsilon^{n-1} \cdot a_n \cdot \lambda^n + \varepsilon^{n-2} \cdot a_{n-1} \cdot \lambda^{n-1} + \varepsilon^{n-3} \cdot a_{n-2} \cdot \lambda^{n-2} + \ldots + \varepsilon \cdot a_2 \cdot \lambda^2 + a_1 \cdot \lambda + \varepsilon^{-1} \cdot a_0$$

3 Sextische (Triquadratische) Regression

3.1 Sextische Regression über den Gaußalgorithmus

Es soll eine Sextische Regression mit Datenpunkten $P_n(\lambda_n; X_n)$ durchgeführt werden. Wobei λ
[001]ff. eine Wellenlänge darstellt und X eine Ortskoordinate.

Um das Polynom $P(\lambda)$ ermitteln zu können, wird folgendes Gleichungssystem aufgestellt:

$$P(\lambda) = a \cdot \lambda^6 + b \cdot \lambda^5 + c \cdot \lambda^4 + d \cdot \lambda^3 + e \cdot \lambda^2 + f \cdot \lambda + g$$

$\Rightarrow$

$$\begin{pmatrix} n & \{x\} & \{x^2\} & \{x^3\} & \{x^4\} & \{x^5\} & \{x^6\} \\ \{x\} & \{x^2\} & \{x^3\} & \{x^4\} & \{x^5\} & \{x^6\} & \{x^7\} \\ \{x^2\} & \{x^3\} & \{x^4\} & \{x^5\} & \{x^6\} & \{x^7\} & \{x^8\} \\ \{x^3\} & \{x^4\} & \{x^5\} & \{x^6\} & \{x^7\} & \{x^8\} & \{x^9\} \\ \{x^4\} & \{x^5\} & \{x^6\} & \{x^7\} & \{x^8\} & \{x^9\} & \{x^{10}\} \\ \{x^5\} & \{x^6\} & \{x^7\} & \{x^8\} & \{x^9\} & \{x^{10}\} & \{x^{11}\} \\ \{x^6\} & \{x^7\} & \{x^8\} & \{x^9\} & \{x^{10}\} & \{x^{11}\} & \{x^{12}\} \end{pmatrix} \cdot \begin{pmatrix} g \\ f \\ e \\ d \\ c \\ b \\ a \end{pmatrix} = \begin{pmatrix} \{x\} \\ \{x \cdot y\} \\ \{x^2 \cdot y\} \\ \{x^3 \cdot y\} \\ \{x^4 \cdot y\} \\ \{x^5 \cdot y\} \\ \{x^6 \cdot y\} \end{pmatrix}$$

Letztendlich ist es „nur“ nötig eine inverse Matrix von M zu finden, um die Lösung zu generieren. Ist M^{-1} gefunden ist die Regression vollendet.

$$M^{-1} = \begin{pmatrix} n & \{x\} & \{x^2\} & \{x^3\} & \{x^4\} & \{x^5\} & \{x^6\} \\ \{x\} & \{x^2\} & \{x^3\} & \{x^4\} & \{x^5\} & \{x^6\} & \{x^7\} \\ \{x^2\} & \{x^3\} & \{x^4\} & \{x^5\} & \{x^6\} & \{x^7\} & \{x^8\} \\ \{x^3\} & \{x^4\} & \{x^5\} & \{x^6\} & \{x^7\} & \{x^8\} & \{x^9\} \\ \{x^4\} & \{x^5\} & \{x^6\} & \{x^7\} & \{x^8\} & \{x^9\} & \{x^{10}\} \\ \{x^5\} & \{x^6\} & \{x^7\} & \{x^8\} & \{x^9\} & \{x^{10}\} & \{x^{11}\} \\ \{x^6\} & \{x^7\} & \{x^8\} & \{x^9\} & \{x^{10}\} & \{x^{11}\} & \{x^{12}\} \end{pmatrix}^{-1}$$

$\Rightarrow$

$$\begin{pmatrix} g \\ f \\ e \\ d \\ c \\ b \\ a \end{pmatrix} = M^{-1}V = M^{-1} \begin{pmatrix} \{x\} \\ \{x \cdot y\} \\ \{x^2 \cdot y\} \\ \{x^3 \cdot y\} \\ \{x^4 \cdot y\} \\ \{x^5 \cdot y\} \\ \{x^6 \cdot y\} \end{pmatrix}$$

Dabei gilt aus den gegebenen Datenpunkten:

$$\{x\} = \sum_{i=1}^{n} \lambda_i \quad \{x\} = \sum_{i=1}^{n} \lambda_i$$

$$\{x^2\} = \sum_{i=1}^{n} (\lambda_i \cdot \lambda_i) \quad \{x \cdot y\} = \sum_{i=1}^{n} (\lambda_i \cdot X_i)$$

$$\{x^3\} = \sum_{i=1}^{n} (\lambda_i \cdot \lambda_i \cdot \lambda_i) \quad \{x^2 \cdot y\} = \sum_{i=1}^{n} (\lambda_i \cdot \lambda_i \cdot X_i)$$

$$\{x^4\} = \sum_{i=1}^{n} (\lambda_i \cdot \lambda_i \cdot \lambda_i \cdot \lambda_i) \qquad \cdots\cdots\cdots$$

$$\cdots\cdots\cdots \qquad \cdots\cdots\cdots$$

3.2 Probleme mit dem Ausdruck $\{x \cdot x \cdot x \cdot x \cdot \ldots\}$

Bei großen Datenmengen, sehr kleinen x- Werten aber auch sehr großen x- Werten kann es Probleme mit dem $\{x \cdot x \cdot x \cdot x \cdot \ldots\}$- Ausdruck geben. Von Rundungsfehlern, Divisionsproblemen bis Zahlenbereichsüberflüsse in verschiedenen Programmiersprachen ergeben sich unter Umständen Probleme bei der Berechnung des Polynoms. [001]ff.

Eine Möglichkeit diese Probleme zu umgehen, ist das Skalieren der Datenpunkte mit ε hin zu günstigen Werten.

$$P_n\left(\lambda_n; X_n\right) \quad \rightarrow \quad P_n\left(\varepsilon \cdot \lambda_n; \varepsilon \cdot X_n\right) \quad \rightarrow \quad P_n\left(\tilde{\lambda}_n; \tilde{X}_n\right)$$

Das Polynom lautet dann:

$$\tilde{X} = a \cdot \tilde{\lambda}^6 + b \cdot \tilde{\lambda}^5 + c \cdot \tilde{\lambda}^4 + d \cdot \tilde{\lambda}^3 + e \cdot \tilde{\lambda}^2 + f \cdot \tilde{\lambda} + g$$

Eine Rückskalierung ergibt die exakte Regression.

$$X = \varepsilon^5 \cdot a \cdot \lambda^6 + \varepsilon^4 \cdot b \cdot \lambda^5 + \varepsilon^3 \cdot c \cdot \lambda^4 + \varepsilon^2 \cdot d \cdot \lambda^3 + \varepsilon \cdot e \cdot \lambda^2 + f \cdot \lambda + \frac{g}{\varepsilon}$$

4 Quintische Regression

4.1 Quintische Regression über den Gaußalgorithmus

Es soll eine Quintische Regression mit Datenpunkten $P_n\left(\lambda_n; X_n\right)$ durchgeführt werden. Wobei λ eine Wellenlänge darstellt und X eine Ortskoordinate. [001]ff.

Um das Polynom $P\left(\lambda\right)$ ermitteln zu können, wird folgendes Gleichungssystem aufgestellt:

$$P\left(\lambda\right) = a \cdot \lambda^5 + b \cdot \lambda^4 + c \cdot \lambda^3 + d \cdot \lambda^2 + e \cdot \lambda + f$$

$\Rightarrow$

$$\begin{pmatrix} n & \{x\} & \{x^2\} & \{x^3\} & \{x^4\} & \{x^5\} \\ \{x\} & \{x^2\} & \{x^3\} & \{x^4\} & \{x^5\} & \{x^6\} \\ \{x^2\} & \{x^3\} & \{x^4\} & \{x^5\} & \{x^6\} & \{x^7\} \\ \{x^3\} & \{x^4\} & \{x^5\} & \{x^6\} & \{x^7\} & \{x^8\} \\ \{x^4\} & \{x^5\} & \{x^6\} & \{x^7\} & \{x^8\} & \{x^9\} \\ \{x^5\} & \{x^6\} & \{x^7\} & \{x^8\} & \{x^9\} & \{x^{10}\} \end{pmatrix} \cdot \begin{pmatrix} f \\ e \\ d \\ c \\ b \\ a \end{pmatrix} = \begin{pmatrix} \{x\} \\ \{x \cdot y\} \\ \{x^2 \cdot y\} \\ \{x^3 \cdot y\} \\ \{x^4 \cdot y\} \\ \{x^5 \cdot y\} \end{pmatrix}$$

Letztendlich ist es über den Gauß- Jordan- Algorithmus „nur“ nötig die Matrix M in eine Einheitsmatrix E um zu wandeln, um eine Lösung zu generieren. Ist E entwickelt ist die Regression vollendet. Mit Hilfe der Determinanten lassen sich dann die Koeffizienten darstellen.

$$|M| \cdot \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} f \\ e \\ d \\ c \\ b \\ a \end{pmatrix} = \begin{pmatrix} |f \leftrightarrow| \\ |e \leftrightarrow| \\ |d \leftrightarrow| \\ |c \leftrightarrow| \\ |b \leftrightarrow| \\ |a \leftrightarrow| \end{pmatrix}$$

Dabei gilt aus den gegebenen Datenpunkten:

$$\{x\} = \sum_{i=1}^{n} \lambda_i \quad \{x\} = \sum_{i=1}^{n} \lambda_i$$

$$\{x^2\} = \sum_{i=1}^{n} \left(\lambda_i \cdot \lambda_i\right) \quad \{x \cdot y\} = \sum_{i=1}^{n} \left(\lambda_i \cdot X_i\right)$$

$$\{x^3\} = \sum_{i=1}^{n} \left(\lambda_i \cdot \lambda_i \cdot \lambda_i\right) \quad \{x^2 \cdot y\} = \sum_{i=1}^{n} \left(\lambda_i \cdot \lambda_i \cdot X_i\right)$$

$$\{x^4\} = \sum_{i=1}^{n} \left(\lambda_i \cdot \lambda_i \cdot \lambda_i \cdot \lambda_i\right) \quad \cdots\cdots\cdots$$

$$\cdots\cdots\cdots \qquad \cdots\cdots\cdots$$

4.2 Probleme mit dem Ausdruck $\{x \cdot x \cdot x \cdot x \cdot \ldots\}$

Bei großen Datenmengen, sehr kleinen x- Werten aber auch sehr großen x- Werten kann es Probleme mit dem $\{x \cdot x \cdot x \cdot x \cdot \ldots\}$- Ausdruck geben. Von Rundungsfehlern, Divisionsproblemen bis Zahlenbereichsüberflüsse in verschiedenen Programmiersprachen ergeben sich unter Umständen Probleme bei der Berechnung des Polynoms. [001]ff.

Eine Möglichkeit diese Probleme zu umgehen, ist das Skalieren der Datenpunkte mit ε hin zu günstigen Werten.

$$P_n\left(\lambda_n; X_n\right) \quad \rightarrow \quad P_n\left(\varepsilon \cdot \lambda_n; \varepsilon \cdot X_n\right) \quad \rightarrow \quad P_n\left(\tilde{\lambda}_n; \tilde{X}_n\right)$$

Das Polynom lautet dann:

$$\tilde{X} = a \cdot \tilde{\lambda}^5 + b \cdot \tilde{\lambda}^4 + c \cdot \tilde{\lambda}^3 + d \cdot \tilde{\lambda}^2 + e \cdot \tilde{\lambda} + f$$

Eine Rückskalierung ergibt die exakte Regression.

$$X = \varepsilon^4 \cdot a \cdot \lambda^5 + \varepsilon^3 \cdot b \cdot \lambda^4 + \varepsilon^2 \cdot c \cdot \lambda^3 + \varepsilon \cdot d \cdot \lambda^2 + e \cdot \lambda + \frac{f}{\varepsilon}$$

5 Biquadratische Regression

5.1 Biquadratische Regression über den Gaußalgorithmus

Es soll eine Biquadratische Regression mit den bekannten Datenpunkten $P_n(\lambda_n; X_n)$ durchgeführt
[001]ff. werden. Wobei λ eine Wellenlänge darstellt und X eine Ortskoordinate.

Um das Polynom $P(\lambda)$ ermitteln zu können, wird folgendes Gleichungssystem aufgestellt:

$$P(\lambda) = a \cdot \lambda^4 + b \cdot \lambda^3 + c \cdot \lambda^2 + d \cdot \lambda + e$$

$\Rightarrow$

$$\begin{pmatrix} n & \{x\} & \{x^2\} & \{x^3\} & \{x^4\} \\ \{x\} & \{x^2\} & \{x^3\} & \{x^4\} & \{x^5\} \\ \{x^2\} & \{x^3\} & \{x^4\} & \{x^5\} & \{x^6\} \\ \{x^3\} & \{x^4\} & \{x^5\} & \{x^6\} & \{x^7\} \\ \{x^4\} & \{x^5\} & \{x^6\} & \{x^7\} & \{x^8\} \end{pmatrix} \cdot \begin{pmatrix} e \\ d \\ c \\ b \\ a \end{pmatrix} = \begin{pmatrix} \{y\} \\ \{x \cdot y\} \\ \{x^2 \cdot y\} \\ \{x^3 \cdot y\} \\ \{x^4 \cdot y\} \end{pmatrix}$$

$\Rightarrow$

$$\begin{aligned} n \cdot e + \{x\} \cdot d + \{x^2\} \cdot c + \{x^3\} \cdot b + \{x^4\} \cdot a &= \{y\} \\ \{x\} \cdot e + \{x^2\} \cdot d + \{x^3\} \cdot c + \{x^4\} \cdot b + \{x^5\} \cdot a &= \{x \cdot y\} \\ \{x^2\} \cdot e + \{x^3\} \cdot d + \{x^4\} \cdot c + \{x^5\} \cdot b + \{x^6\} \cdot a &= \{x^2 \cdot y\} \\ \{x^3\} \cdot e + \{x^4\} \cdot d + \{x^5\} \cdot c + \{x^6\} \cdot b + \{x^7\} \cdot a &= \{x^3 \cdot y\} \\ \{x^4\} \cdot e + \{x^5\} \cdot d + \{x^6\} \cdot c + \{x^7\} \cdot b + \{x^8\} \cdot a &= \{x^4 \cdot y\} \end{aligned}$$

Durch Lösen des Gleichungssystems wird das Polynom bestimmt.

$$\begin{aligned} A \cdot d + B \cdot c + C \cdot b + D \cdot a &= E \\ F \cdot d + G \cdot c + H \cdot b + I \cdot a &= J \\ K \cdot d + L \cdot c + M \cdot b + N \cdot a &= O \\ P \cdot d + Q \cdot c + R \cdot b + S \cdot a &= T \end{aligned}$$

$\Rightarrow$

$$\begin{aligned} \tilde{A} \cdot c + \tilde{B} \cdot b + \tilde{C} \cdot a &= \tilde{D} \\ \tilde{E} \cdot c + \tilde{F} \cdot b + \tilde{G} \cdot a &= \tilde{H} \\ \tilde{I} \cdot c + \tilde{J} \cdot b + \tilde{K} \cdot a &= \tilde{L} \end{aligned}$$

$\Rightarrow$

$$\begin{aligned} \hat{A} \cdot b + \hat{B} \cdot a &= \hat{C} \\ \hat{D} \cdot b + \hat{E} \cdot a &= \hat{F} \end{aligned}$$

$\Rightarrow$

$$\bar{A} \cdot a = \bar{B}$$

Damit sind die Koeffizienten $a; b; c; d$ und e betimmt. (Rot hervorgehobene Terme sind bestehend aus nativen Hilfsvariablen)

$$a = \frac{\bar{B}}{\bar{A}}$$

$\Rightarrow$

$$b = \frac{\hat{C}}{\hat{A}} - \frac{\hat{B}}{\hat{A}} \cdot a$$

$\Rightarrow$

$$c = \frac{\tilde{D}}{\tilde{A}} - \frac{\tilde{C}}{\tilde{A}} \cdot a - \frac{\tilde{B}}{\tilde{A}} \cdot b$$

$\Rightarrow$

$$d = \frac{E}{A} - \frac{D}{A} \cdot a - \frac{C}{A} \cdot b - \frac{B}{A} \cdot c$$

$\Rightarrow$

$$e = Y - X \cdot a - W \cdot b - V \cdot c - U \cdot d$$

Mit den nativen Hilfsvariablen.

$$A = U - \frac{\{x^2\}}{\{x\}} \quad B = V - \frac{\{x^3\}}{\{x\}} \quad C = W - \frac{\{x^4\}}{\{x\}} \quad D = X - \frac{\{x^5\}}{\{x\}}$$

$$E = Y - \frac{\{x \cdot y\}}{\{x\}} \quad F = U - \frac{\{x^3\}}{\{x^2\}} \quad G = V - \frac{\{x^4\}}{\{x^2\}} \quad H = W - \frac{\{x^5\}}{\{x^2\}}$$

$$I = X - \frac{\{x^6\}}{\{x^2\}} \quad J = Y - \frac{\{x^2 \cdot y\}}{\{x^2\}} \quad K = U - \frac{\{x^4\}}{\{x^3\}} \quad L = V - \frac{\{x^5\}}{\{x^3\}}$$

$$M = W - \frac{\{x^6\}}{\{x^3\}} \quad N = X - \frac{\{x^7\}}{\{x^3\}} \quad O = Y - \frac{\{x^3 \cdot y\}}{\{x^3\}} \quad P = U - \frac{\{x^5\}}{\{x^4\}}$$

$$Q = V - \frac{\{x^6\}}{\{x^4\}} \quad R = W - \frac{\{x^7\}}{\{x^4\}} \quad S = X - \frac{\{x^8\}}{\{x^4\}} \quad T = Y - \frac{\{x^4 \cdot y\}}{\{x^4\}}$$

Und:

$$U = \frac{\{x\}}{n} \quad V = \frac{\{x^2\}}{n} \quad W = \frac{\{x^3\}}{n} \quad X = \frac{\{x^4\}}{n} \quad Y = \frac{\{y\}}{n}$$

Sowie den nichtnativen Hilfsvariablen.

$$\tilde{A} = \frac{B}{A} - \frac{G}{F} \quad \tilde{B} = \frac{C}{A} - \frac{H}{F} \quad \tilde{C} = \frac{D}{A} - \frac{I}{F} \quad \tilde{D} = \frac{E}{A} - \frac{J}{F}$$

$$\tilde{E} = \frac{B}{A} - \frac{L}{K} \quad \tilde{F} = \frac{C}{A} - \frac{M}{K} \quad \tilde{G} = \frac{D}{A} - \frac{N}{K} \quad \tilde{H} = \frac{E}{A} - \frac{O}{K}$$

$$\tilde{I} = \frac{B}{A} - \frac{Q}{P} \quad \tilde{J} = \frac{C}{A} - \frac{R}{P} \quad \tilde{K} = \frac{D}{A} - \frac{S}{P} \quad \tilde{L} = \frac{E}{A} - \frac{T}{P}$$

Und:

$$\hat{A} = \frac{\tilde{B}}{\tilde{A}} - \frac{\tilde{F}}{\tilde{E}} \quad \hat{B} = \frac{\tilde{C}}{\tilde{A}} - \frac{\tilde{G}}{\tilde{E}} \quad \hat{C} = \frac{\tilde{D}}{\tilde{A}} - \frac{\tilde{H}}{\tilde{E}}$$

$$\hat{D} = \frac{\tilde{B}}{\tilde{A}} - \frac{\tilde{J}}{\tilde{I}} \quad \hat{E} = \frac{\tilde{C}}{\tilde{A}} - \frac{\tilde{K}}{\tilde{I}} \quad \hat{F} = \frac{\tilde{D}}{\tilde{A}} - \frac{\tilde{L}}{\tilde{I}}$$

Und:

$$\bar{A} = \frac{\hat{B}}{\hat{A}} - \frac{\hat{E}}{\hat{D}} \quad \bar{B} = \frac{\hat{C}}{\hat{A}} - \frac{\hat{F}}{\hat{D}}$$

Dabei gilt aus den gegebenen Datenpunkten:

$$\{x\} = \sum_{i=1}^{n} \lambda_i \quad \{x\} = \sum_{i=1}^{n} \lambda_i$$

$$\{x^2\} = \sum_{i=1}^{n} (\lambda_i \cdot \lambda_i) \quad \{x \cdot y\} = \sum_{i=1}^{n} (\lambda_i \cdot X_i)$$

$$\{x^3\} = \sum_{i=1}^{n} (\lambda_i \cdot \lambda_i \cdot \lambda_i) \quad \{x^2 \cdot y\} = \sum_{i=1}^{n} (\lambda_i \cdot \lambda_i \cdot X_i)$$

$$\{x^4\} = \sum_{i=1}^{n} (\lambda_i \cdot \lambda_i \cdot \lambda_i \cdot \lambda_i) \qquad \cdots\cdots\cdots$$

$$\cdots\cdots\cdots \qquad \cdots\cdots\cdots$$

5.2 Probleme mit dem Ausdruck $\{x \cdot x \cdot x \cdot x \cdot \ldots\}$

[001]ff. Bei großen Datenmengen, sehr kleinen x- Werten aber auch sehr großen x- Werten kann es Probleme mit dem $\{x \cdot x \cdot x \cdot x \cdot \ldots\}$- Ausdruck geben. Von Rundungsfehlern, Divisionsproblemen bis Zahlenbereichsüberflüsse in verschiedenen Programmiersprachen ergeben sich unter Umständen Probleme bei der Berechnung des Polynoms.

Eine Möglichkeit diese Probleme zu umgehen, ist das Skalieren der Datenpunkte mit ε hin zu günstigen Werten.

$$P_n\left(\lambda_n; X_n\right) \quad \rightarrow \quad P_n\left(\varepsilon \cdot \lambda_n; \varepsilon \cdot X_n\right) \quad \rightarrow \quad P_n\left(\tilde{\lambda}_n; \tilde{X}_n\right)$$

Das Polynom lautet dann:

$$\tilde{X} = a \cdot \tilde{\lambda}^4 + b \cdot \tilde{\lambda}^3 + c \cdot \tilde{\lambda}^2 + d \cdot \tilde{\lambda} + e$$

Eine Rückskalierung ergibt die exakte Regression.

$$X = \varepsilon^3 \cdot a \cdot \lambda^4 + \varepsilon^2 \cdot b \cdot \lambda^3 + \varepsilon \cdot c \cdot \lambda^2 + d \cdot \lambda + \frac{e}{\varepsilon}$$

6 Kubikregression

6.1 Kubikregression über den Gaußalgorithmus

Es soll eine Kubikregression mit den Datenpunkten $P_n\left(\lambda_n; X_n\right)$ durchgeführt werden. Wobei λ eine Wellenlänge darstellt und X eine Ortskoordinate. [001]ff.

Um das Polynom $P\left(\lambda\right)$ ermitteln zu können, wird folgendes Gleichungssystem aufgestellt:

$$P\left(\lambda\right) = a \cdot \lambda^3 + b \cdot \lambda^2 + c \cdot \lambda + d$$

$\Rightarrow$

$$\begin{pmatrix} n & \{x\} & \left\{x^2\right\} & \left\{x^3\right\} \\ \{x\} & \left\{x^2\right\} & \left\{x^3\right\} & \left\{x^4\right\} \\ \left\{x^2\right\} & \left\{x^3\right\} & \left\{x^4\right\} & \left\{x^5\right\} \\ \left\{x^3\right\} & \left\{x^4\right\} & \left\{x^5\right\} & \left\{x^6\right\} \end{pmatrix} \cdot \begin{pmatrix} d \\ c \\ b \\ a \end{pmatrix} = \begin{pmatrix} \{y\} \\ \{x \cdot y\} \\ \left\{x^2 \cdot y\right\} \\ \left\{x^3 \cdot y\right\} \end{pmatrix}$$

$\Rightarrow$

$$\begin{aligned} d \cdot n + c \cdot \{x\} + b \cdot \left\{x^2\right\} + a \cdot \left\{x^3\right\} &= \{y\} \\ d \cdot \{x\} + c \cdot \left\{x^2\right\} + b \cdot \left\{x^3\right\} + a \cdot \left\{x^4\right\} &= \{x \cdot y\} \\ d \cdot \left\{x^2\right\} + c \cdot \left\{x^3\right\} + b \cdot \left\{x^4\right\} + a \cdot \left\{x^5\right\} &= \left\{x^2 \cdot y\right\} \\ d \cdot \left\{x^3\right\} + c \cdot \left\{x^4\right\} + b \cdot \left\{x^5\right\} + a \cdot \left\{x^6\right\} &= \left\{x^3 \cdot y\right\} \end{aligned}$$

Durch Lösen des Gleichungssystems ist das Polynom bestimmt. Dabei gilt aus den gegebenen Datenpunkten:

$$\{x\} = \sum_{i=1}^{n} \lambda_i \quad \{x\} = \sum_{i=1}^{n} \lambda_i$$

$$\left\{x^2\right\} = \sum_{i=1}^{n} \left(\lambda_i \cdot \lambda_i\right) \quad \{x \cdot y\} = \sum_{i=1}^{n} \left(\lambda_i \cdot X_i\right)$$

$$\left\{x^3\right\} = \sum_{i=1}^{n} \left(\lambda_i \cdot \lambda_i \cdot \lambda_i\right) \quad \left\{x^2 \cdot y\right\} = \sum_{i=1}^{n} \left(\lambda_i \cdot \lambda_i \cdot X_i\right)$$

$$\left\{x^4\right\} = \sum_{i=1}^{n} \left(\lambda_i \cdot \lambda_i \cdot \lambda_i \cdot \lambda_i\right) \qquad \cdots\cdots\cdots$$

$$\cdots\cdots\cdots \qquad \cdots\cdots\cdots$$

Zur Lösung, der Gaußalgorithmus:

$$\begin{aligned} d + \frac{\{x\}}{n} \cdot c + \frac{\left\{x^2\right\}}{n} \cdot b + \frac{\left\{x^3\right\}}{n} \cdot a &= \frac{\{y\}}{n} \\ d + \frac{\left\{x^2\right\}}{\{x\}} \cdot c + \frac{\left\{x^3\right\}}{\{x\}} \cdot b + \frac{\left\{x^4\right\}}{\{x\}} \cdot a &= \frac{\{x \cdot y\}}{\{x\}} \\ d + \frac{\left\{x^3\right\}}{\left\{x^2\right\}} \cdot c + \frac{\left\{x^4\right\}}{\left\{x^2\right\}} \cdot b + \frac{\left\{x^5\right\}}{\left\{x^2\right\}} \cdot a &= \frac{\left\{x^2 \cdot y\right\}}{\left\{x^2\right\}} \\ d + \frac{\left\{x^4\right\}}{\left\{x^3\right\}} \cdot c + \frac{\left\{x^5\right\}}{\left\{x^3\right\}} \cdot b + \frac{\left\{x^6\right\}}{\left\{x^3\right\}} \cdot a &= \frac{\left\{x^3 \cdot y\right\}}{\left\{x^3\right\}} \end{aligned}$$

$\Rightarrow$

$$\begin{aligned} A \cdot c + B \cdot b + C \cdot a &= D \\ E \cdot c + F \cdot b + G \cdot a &= H \\ I \cdot c + J \cdot b + K \cdot a &= L \end{aligned}$$

$\Rightarrow$

$$\begin{aligned} M \cdot b + N \cdot a &= O \\ P \cdot b + Q \cdot a &= R \end{aligned}$$

$\Rightarrow$

$$S \cdot a = T$$

Damit sind alle Koeffizienten $a; b; c$ und d ermittelt. (Rot hervorgehobene Terme sind bestehend aus nativen Hilfsvariablen)

$$a = \frac{T}{S}$$

$\Rightarrow$

$$b = \frac{O}{M} - \frac{N}{M} \cdot a$$

$\Rightarrow$

$$c = \frac{D}{A} - \frac{B}{A} \cdot b - \frac{C}{A} \cdot a$$

$\Rightarrow$

$$d = U - V \cdot c - W \cdot b - X \cdot a$$

Mit den nativen Hilfsvariablen:

$$U = \frac{\{y\}}{n} \qquad V = \frac{\{x\}}{n} \qquad W = \frac{\{x^2\}}{n} \qquad X = \frac{\{x^3\}}{n}$$

Und:

$$I = V - \frac{\{x^4\}}{\{x^3\}} \quad J = W - \frac{\{x^5\}}{\{x^3\}} \quad K = X - \frac{\{x^6\}}{\{x^3\}} \quad L = U - \frac{\{x^3 \cdot y\}}{\{x^3\}}$$

$$E = V - \frac{\{x^3\}}{\{x^2\}} \quad F = W - \frac{\{x^4\}}{\{x^2\}} \quad G = X - \frac{\{x^5\}}{\{x^2\}} \quad H = U - \frac{\{x^2 \cdot y\}}{\{x^2\}}$$

$$A = V - \frac{\{x^2\}}{\{x\}} \quad B = W - \frac{\{x^3\}}{\{x\}} \quad C = X - \frac{\{x^4\}}{\{x\}} \quad D = U - \frac{\{x \cdot y\}}{\{x\}}$$

Sowie den nichtnativen Hilfsvariablen.

$$S = \frac{N}{M} - \frac{Q}{P} \qquad T = \frac{O}{M} - \frac{R}{P}$$

Und:

$$P = \frac{B}{A} - \frac{J}{I} \qquad Q = \frac{C}{A} - \frac{K}{I} \qquad R = \frac{D}{A} - \frac{L}{I}$$

$$M = \frac{B}{A} - \frac{F}{E} \qquad N = \frac{C}{A} - \frac{G}{E} \qquad O = \frac{D}{A} - \frac{H}{E}$$

Die eben aufgezeigte Umformung obiger Matrizen mittels des Gauß- (Jordan-) Verfahren ergibt ohne Auflösung der Arrays:

$$\begin{pmatrix} n & \{x\} & \{x^2\} & \{x^3\} \\ \{x\} & \{x^2\} & \{x^3\} & \{x^4\} \\ \{x^2\} & \{x^3\} & \{x^4\} & \{x^5\} \\ \{x^3\} & \{x^4\} & \{x^5\} & \{x^6\} \end{pmatrix} \cdot \begin{pmatrix} d \\ c \\ b \\ a \end{pmatrix} = \begin{pmatrix} \{x\} \\ \{x \cdot y\} \\ \{x^2 \cdot y\} \\ \{x^3 \cdot y\} \end{pmatrix}$$

$\Rightarrow$

$$|M| \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} d \\ c \\ b \\ a \end{pmatrix} = \begin{pmatrix} |d \leftrightarrow| \\ |c \leftrightarrow| \\ |b \leftrightarrow| \\ |a \leftrightarrow| \end{pmatrix}$$

Wobei letztendlich das der Lösungsweg über die Determinanten darstellt.

6.2 Probleme mit dem Ausdruck $\{x \cdot x \cdot x \cdot x \cdot \ldots\}$

Bei großen Datenmengen, sehr kleinen x- Werten aber auch sehr großen x- Werten kann es Probleme mit dem $\{x \cdot x \cdot x \cdot x \cdot \ldots\}$- Ausdruck geben. Von Rundungsfehlern, Divisionsproblemen bis Zahlenbereichsüberflüsse in verschiedenen Programmiersprachen ergeben sich unter Umständen Probleme bei der Berechnung des Polynoms. [001]ff.

Eine Möglichkeit diese Probleme zu umgehen, ist das Skalieren der Datenpunkte mit ε hin zu günstigen Werten.

$$P_n\left(\lambda_n; X_n\right) \quad \rightarrow \quad P_n\left(\varepsilon \cdot \lambda_n; \varepsilon \cdot X_n\right) \quad \rightarrow \quad P_n\left(\tilde{\lambda}_n; \tilde{X}_n\right)$$

Das Polynom lautet dann:

$$\tilde{X} = a \cdot \tilde{\lambda}^3 + b \cdot \tilde{\lambda}^2 + c \cdot \tilde{\lambda} + d$$

Eine Rückskalierung ergibt die exakte Regression.

$$X = \varepsilon^2 \cdot a \cdot \lambda^3 + \varepsilon \cdot b \cdot \lambda^2 + c \cdot \lambda + \frac{d}{\varepsilon}$$

7 Quadratregression

7.1 Quadratregression über die Determinanten

Es soll eine Quadratregression mit Datenpunkten $P_n\left(\lambda_n; X_n\right)$ durchgeführt werden. Wobei λ eine
[001]ff. Wellenlänge darstellt und X eine Ortskoordinate.

Definiert werden Summen folgender Form:

$$\{x\} = \sum_{i=1}^{n} \lambda_i \quad \{y\} = \sum_{i=1}^{n} X_i$$

$$\{x^2\} = \sum_{i=1}^{n} (\lambda_i \cdot \lambda_i) \quad \{x \cdot y\} = \sum_{i=1}^{n} (\lambda_i \cdot X_i)$$

$$\{x^3\} = \sum_{i=1}^{n} (\lambda_i \cdot \lambda_i \cdot \lambda_i) \quad \{x^2 \cdot y\} = \sum_{i=1}^{n} (\lambda_i \cdot \lambda_i \cdot X_i)$$

$$\{x^4\} = \sum_{i=1}^{n} (\lambda_i \cdot \lambda_i \cdot \lambda_i \cdot \lambda_i)$$

Damit sind die Determinanten definiert.

$$D = \begin{vmatrix} n & \{x\} & \{x^2\} \\ \{x\} & \{x^2\} & \{x^3\} \\ \{x^2\} & \{x^3\} & \{x^4\} \end{vmatrix}$$

$$C = \begin{vmatrix} \{y\} & \{x\} & \{x^2\} \\ \{x \cdot y\} & \{x^2\} & \{x^3\} \\ \{x^2 \cdot y\} & \{x^3\} & \{x^4\} \end{vmatrix}$$

$$B = \begin{vmatrix} n & \{y\} & \{x^2\} \\ \{x\} & \{x \cdot y\} & \{x^3\} \\ \{x^2\} & \{x^2 \cdot y\} & \{x^4\} \end{vmatrix}$$

$$A = \begin{vmatrix} n & \{x\} & \{y\} \\ \{x\} & \{x^2\} & \{x \cdot y\} \\ \{x^2\} & \{x^3\} & \{x^2 \cdot y\} \end{vmatrix}$$

Jetzt sind die Koeffizienten a, b und c berechenbar.

$$a = \frac{A}{D} \quad b = \frac{B}{D} \quad c = \frac{C}{D}$$

Mit:

$$D = \begin{cases} +n \cdot \{x^2\} \cdot \{x^4\} - n \cdot \{x^3\} \cdot \{x^3\} \\ -\{x\} \cdot \{x\} \cdot \{x^4\} + \{x\} \cdot \{x^2\} \cdot \{x^3\} \\ +\{x\} \cdot \{x^2\} \cdot \{x^3\} - \{x^2\} \cdot \{x^2\} \cdot \{x^2\} \end{cases}$$

$$C = \begin{cases} +\{y\} \cdot \{x^2\} \{x^4\} - \{y\} \cdot \{x^3\} \cdot \{x^3\} \\ -\{x \cdot y\} \cdot \{x\} \cdot \{x^4\} + \{x \cdot y\} \cdot \{x^2\} \cdot \{x^3\} \\ +\{x^2 \cdot y\} \cdot \{x\} \cdot \{x^3\} - \{x^2 \cdot y\} \cdot \{x^2\} \cdot \{x^2\} \end{cases}$$

$$B = \begin{cases} +n \cdot \{x \cdot y\} \cdot \{x^4\} - n \cdot \{x^3\} \cdot \{x^2 \cdot y\} \\ -\{x\} \cdot \{y\} \cdot \{x^4\} + \{x\} \cdot \{x^2\} \cdot \{x^2 \cdot y\} \\ +\{x^2\} \cdot \{y\} \cdot \{x^3\} - \{x^2\} \cdot \{x^2\} \cdot \{x \cdot y\} \end{cases}$$

$$A = \left\{ \begin{array}{l} +n \cdot \left\{x^2\right\} \cdot \left\{x^2 \cdot y\right\} - n \cdot \{x \cdot y\} \cdot \left\{x^2 \cdot y\right\} \\ - \{x\} \cdot \{x\} \cdot \left\{x^2 \cdot y\right\} + \{x\} \cdot \{y\} \cdot \left\{x^3\right\} \\ + \left\{x^2\right\} \cdot \{x\} \cdot \{x \cdot y\} - \{y\} \cdot \left\{x^2\right\} \cdot \left\{x^2\right\} \end{array} \right.$$

Das Regressionspolynom y ist bestimmt.

$$y = a \cdot x^2 + b \cdot x + c$$

$\Rightarrow$

$$X = a \cdot \lambda^2 + b \cdot \lambda + c$$

7.2 Quadratregression über den Gaußalgorithmus

[001]ff. Es soll eine Quadratregression mit Datenpunkten $P_n(\lambda_n; X_n)$ ausgeführt werden. Wobei λ eine Wellenlänge darstellt und X eine Ortskoordinate.

Definiert werden Summen folgender Form:

$$\{x\} = \sum_{i=1}^{n} \lambda_i \quad \{y\} = \sum_{i=1}^{n} X_i$$

$$\{x^2\} = \sum_{i=1}^{n} (\lambda_i \cdot \lambda_i) \quad \{x \cdot y\} = \sum_{i=1}^{n} (\lambda_i \cdot X_i)$$

$$\{x^3\} = \sum_{i=1}^{n} (\lambda_i \cdot \lambda_i \cdot \lambda_i) \quad \{x^2 \cdot y\} = \sum_{i=1}^{n} (\lambda_i \cdot \lambda_i \cdot X_i)$$

$$\{x^4\} = \sum_{i=1}^{n} (\lambda_i \cdot \lambda_i \cdot \lambda_i \cdot \lambda_i)$$

Damit ist das Gleichungssystem für die Koeffizienten a, b und c berechenbar.

$$\begin{aligned} n \cdot c + \{x\} \cdot b + \{x^2\} \cdot a &= \{y\} \\ \{x\} \cdot c + \{x^2\} \cdot b + \{x^3\} \cdot a &= \{x \cdot y\} \\ \{x^2\} \cdot c + \{x^3\} \cdot b + \{x^4\} \cdot a &= \{x^2 \cdot y\} \end{aligned}$$

Mittels Gaußalgorithmus wird aufgelöst.

$$\begin{aligned} c + \frac{\{x\}}{n} \cdot b + \frac{\{x^2\}}{n} \cdot a &= \frac{\{y\}}{n} \\ c + \frac{\{x^2\}}{\{x\}} \cdot b + \frac{\{x^3\}}{\{x\}} \cdot a &= \frac{\{x \cdot y\}}{\{x\}} \\ c + \frac{\{x^3\}}{\{x^2\}} \cdot b + \frac{\{x^4\}}{\{x^2\}} \cdot a &= \frac{\{x^2 \cdot y\}}{\{x^2\}} \end{aligned}$$

$\Rightarrow$

$$\begin{aligned} A \cdot b + B \cdot a &= C \\ D \cdot b + E \cdot a &= F \end{aligned}$$

Der nächste Schritt.

$$\begin{aligned} b + \frac{B}{A} \cdot a &= \frac{C}{A} \\ b + \frac{E}{D} \cdot a &= \frac{F}{D} \end{aligned}$$

$\Rightarrow$

$$G \cdot a = H$$

Damit ist a bestimmt (Rot hervorgehobene Terme sind bestehend aus nativen Hilfsvariablen).

$$a = \frac{H}{G} = \frac{A \cdot F - C \cdot D}{A \cdot E - B \cdot D}$$

Rückwärts eingesetzt ergibt b.

$$b + \frac{B}{A} \cdot a = \frac{C}{A}$$

$\Rightarrow$

$$b = \frac{C}{A} - \frac{B}{A} \cdot a$$

Für c gilt:

$$c + I \cdot b + J \cdot a = K$$

$\Rightarrow$

$$c = K - I \cdot b - J \cdot a$$

Das Regressionspolynom y ist bestimmt.

$$y = a \cdot x^2 + b \cdot x + c$$

$\Rightarrow$

$$X = a \cdot \lambda^2 + b \cdot \lambda + c$$

Mit den nativen Hilfsvariablen:

$$I = \frac{\{x\}}{n} \quad J = \frac{\{x^2\}}{n} \quad K = \frac{\{y\}}{n}$$

Und:

$$A = I - \frac{\{x^2\}}{\{x\}} \quad D = I - \frac{\{x^3\}}{\{x^2\}}$$
$$B = J - \frac{\{x^3\}}{\{x\}} \quad E = J - \frac{\{x^4\}}{\{x^2\}}$$
$$C = K - \frac{\{x \cdot y\}}{\{x\}} \quad F = K - \frac{\{x^2 \cdot y\}}{\{x^2\}}$$

Sowie den nichtnativen Hilfsvariablen:

$$G = \frac{B}{A} - \frac{E}{D} \qquad H = \frac{C}{A} - \frac{F}{D}$$

Die eben aufgezeigte Umformung obiger Matrizen mittels des Gauß- (Jordan-) Verfahren ergibt ohne Auflösung der Arrays:

$$\begin{pmatrix} n & \{x\} & \{x^2\} \\ \{x\} & \{x^2\} & \{x^3\} \\ \{x^2\} & \{x^3\} & \{x^4\} \end{pmatrix} \cdot \begin{pmatrix} c \\ b \\ a \end{pmatrix} = \begin{pmatrix} \{y\} \\ \{x \cdot y\} \\ \{x^2 \cdot y\} \end{pmatrix}$$

$\Rightarrow$

$$|M| \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} c \\ b \\ a \end{pmatrix} = \begin{pmatrix} |c \leftrightarrow| \\ |b \leftrightarrow| \\ |a \leftrightarrow| \end{pmatrix}$$

Wobei letztendlich das der Lösungsweg über die Determinanten darstellt.

$$|M| \cdot c = |c \leftrightarrow| = \begin{cases} +\{y\} \cdot \{x^2\} \cdot \{x^4\} - \{y\} \cdot \{x^3\}^2 - \{x \cdot y\} \cdot \{x\} \cdot \{x^4\} + \{x \cdot y\} \cdot \{x^2\} \cdot \{x^3\} \\ +\{x^2 \cdot y\} \cdot \{x\} \cdot \{x^3\} - \{x^2 \cdot y\} \cdot \{x^2\}^2 \end{cases}$$

$$|M| \cdot b = |b \leftrightarrow| = \begin{cases} +\{y\} \cdot \{x^2\} \cdot \{x^3\} - \{y\} \cdot \{x\} \cdot \{x^4\} + \{x \cdot y\} \cdot n \cdot \{x^4\} - \{x \cdot y\} \cdot \{x^2\}^2 \\ -\{x^2 \cdot y\} \cdot n \cdot \{x^3\} + \{x^2 \cdot y\} \cdot \{x\} \cdot \{x^2\} \end{cases}$$

$$|M| \cdot a = |a \leftrightarrow| = \begin{cases} +\{y\} \cdot \{x\} \cdot \{x^3\} - \{y\} \cdot \{x^2\}^2 - \{x \cdot y\} \cdot n \cdot \{x^3\} + \{x \cdot y\} \cdot \{x\} \cdot \{x^2\} \\ +\{x^2 \cdot y\} \cdot n \cdot \{x^2\} - \{x^2 \cdot y\} \cdot \{x\}^2 \end{cases}$$

Womit die Koeffizienten $a; b$ und c direkt ermittelt werden können.

$$|M| = n \cdot \{x^2\} \cdot \{x^4\} - n \cdot \{x^3\}^2 - \{x\}^2 \cdot \{x^4\} + 2 \cdot \{x\} \cdot \{x^2\} \cdot \{x^3\} - \{x^2\}^3$$

$\Rightarrow$

$$|M| = \left(n \cdot \{x^2\} - \{x\}^2\right) \cdot \{x^4\} + 2 \cdot \{x\} \cdot \{x^2\} \cdot \{x^3\} - \{x^2\}^3 - n \cdot \{x^3\}^2$$

$\Rightarrow$

$$|M|^{(2)} = |M|^{(1)} \cdot \{x^4\} + 2 \cdot \{x\} \cdot \{x^2\} \cdot \{x^3\} - \{x^2\}^3 - n \cdot \{x^3\}^2$$

Wobei $|M|^{(2)}$ die Determinante der Matrix M aus der Quadratregression ist und $|M|^{(1)}$ die der Linearregression.

7.3 Probleme mit dem Ausdruck $\{x \cdot x \cdot x \cdot x\}$

[001]ff. Bei großen Datenmengen, sehr kleinen x- Werten aber auch sehr großen x- Werten kann es Probleme mit dem $\{x \cdot x \cdot x \cdot x\}$- Ausdruck geben. Von Rundungsfehlern, Divisionsproblemen bis Zahlenbereichsüberflüsse in verschiedenen Programmiersprachen ergeben sich unter Umständen Probleme bei der Berechnung des Polynoms.

Eine Möglichkeit diese Probleme zu umgehen, ist das Skalieren der Datenpunkte mit ε hin zu günstigen Werten.

$$P_n\left(\lambda_n; X_n\right) \quad \rightarrow \quad P_n\left(\varepsilon \cdot \lambda_n; \varepsilon \cdot X_n\right) \quad \rightarrow \quad P_n\left(\tilde{\lambda}_n; \tilde{X}_n\right)$$

Das Polynom lautet dann:

$$\tilde{X} = a \cdot \tilde{\lambda}^2 + b \cdot \tilde{\lambda} + c$$

Eine Rückskalierung ergibt die exakte Regression.

$$X = \varepsilon \cdot a \cdot \lambda^2 + b \cdot \lambda + \frac{c}{\varepsilon}$$

8 Linearregression

Es wird eine Linearregression mit Datenpunkten in der Form $P_n\left(\lambda_n; X_n\right)$ durchgeführt. Wobei λ eine Wellenlänge darstellt und X eine Ortskoordinate. [001]ff.

Definiert werden Summen folgender Form:

$$\{x\} = \sum_{i=1}^{n} \lambda_i \qquad \{y\} = \sum_{i=1}^{n} X_i$$

$$\left\{x^2\right\} = \sum_{i=1}^{n} \left(\lambda_i \cdot \lambda_i\right) \qquad \{x \cdot y\} = \sum_{i=1}^{n} \left(\lambda_i \cdot X_i\right)$$

8.1 Linearregression über die Determinanten

Aufstellen der Matrizen.

$$M = \begin{pmatrix} n & \{x\} \\ \{x\} & \{x^2\} \end{pmatrix} \qquad V = \begin{pmatrix} \{y\} \\ \{x \cdot y\} \end{pmatrix}$$

Somit:

$$b \leftrightarrow \begin{pmatrix} \{y\} & \{x\} \\ \{x \cdot y\} & \{x^2\} \end{pmatrix} \qquad a \leftrightarrow \begin{pmatrix} n & \{y\} \\ \{x\} & \{x \cdot y\} \end{pmatrix}$$

Die dazu gehörigen Determinanten:

$$\begin{aligned} \det(M) &= |M| = n \cdot \left\{x^2\right\} - \{x\} \cdot \{x\} \\ \det(a \leftrightarrow) &= |a \leftrightarrow| = n \cdot \{x \cdot y\} - \{x\} \cdot \{y\} \\ \det(b \leftrightarrow) &= |b \leftrightarrow| = \{y\} \cdot \left\{x^2\right\} - \{x \cdot y\} \cdot \{x\} \end{aligned}$$

Damit sind die Koeffizienten a und b berechenbar.

$$a = \frac{|a \leftrightarrow|}{|M|} \qquad b = \frac{|b \leftrightarrow|}{|M|}$$

$\Rightarrow$

$$a = \frac{\{x \cdot y\} \cdot n - \{x\} \cdot \{y\}}{\{x^2\} \cdot n - \{x\}^2} \qquad b = \frac{\left\{x^2\right\} \cdot \{y\} - \{x \cdot y\} \cdot \{x\}}{\{x^2\} \cdot n - \{x\}^2}$$

Das Linearpolynom y ist bestimmt.

$$y = a \cdot x + b$$

$\Rightarrow$

$$X = a \cdot \lambda + b$$

8.2 Linearregression über den Gaußalgorithmus

Die zur Linearregression dazu gehörige Matrize: [001]

$$\begin{pmatrix} n & \{x\} \\ \{x\} & \{x^2\} \end{pmatrix} \cdot \begin{pmatrix} b \\ a \end{pmatrix} = \begin{pmatrix} \{y\} \\ \{x \cdot y\} \end{pmatrix}$$

$\Rightarrow$

$$n \cdot b + \{x\} \cdot a = \{y\}$$
$$\{x\} \cdot b + \{x^2\} \cdot a = \{x \cdot y\}$$

$\Rightarrow$

$$b + \frac{\{x\}}{n} \cdot a = \frac{\{y\}}{n}$$
$$b + \frac{\{x^2\}}{\{x\}} \cdot a = \frac{\{x \cdot y\}}{\{x\}}$$

$\Rightarrow$

$$\left(\frac{\{x\}}{n} - \frac{\{x^2\}}{\{x\}} \right) \cdot a = \frac{\{y\}}{n} - \frac{\{x \cdot y\}}{\{x\}}$$

$\Rightarrow$

$$a = \frac{\{x \cdot y\} \cdot n - \{x\} \cdot \{y\}}{\{x^2\} \cdot n - \{x\}^2}$$

Für b gilt:

$$b = \frac{\{y\}}{n} - \frac{\{x\}}{n} \cdot \frac{\{x\} \cdot \{y\} - \{x \cdot y\} \cdot n}{\{x\}^2 - \{x^2\} \cdot n}$$

$\Rightarrow$

$$b = \frac{\{x^2\} \cdot \{y\} - \{x \cdot y\} \cdot \{x\}}{\{x^2\} \cdot n - \{x\}^2}$$

Das Linearpolynom y ist bestimmt.

$$y = a \cdot x + b$$

$\Rightarrow$

$$X = a \cdot \lambda + b$$

Die eben aufgezeigte Umformung obiger Matrizen mittels des Gauß-(Jordan-)Verfahren ergibt ohne Auflösung der Arrays:

$$\begin{pmatrix} n & \{x\} \\ \{x\} & \{x^2\} \end{pmatrix} \cdot \begin{pmatrix} b \\ a \end{pmatrix} = \begin{pmatrix} \{y\} \\ \{x \cdot y\} \end{pmatrix}$$

$\Rightarrow$

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} b \\ a \end{pmatrix} = \begin{pmatrix} \frac{\{x^2\} \cdot \{y\} - \{x \cdot y\} \cdot \{x\}}{\{x^2\} \cdot n - \{x\}^2} \\ \frac{\{x \cdot y\} \cdot n - \{x\} \cdot \{y\}}{\{x^2\} \cdot n - \{x\}^2} \end{pmatrix}$$

$\Rightarrow$

$$\left(\{x^2\} \cdot n - \{x\}^2\right) \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} b \\ a \end{pmatrix} = \begin{pmatrix} \{x^2\} \cdot \{y\} - \{x \cdot y\} \cdot \{x\} \\ \{x \cdot y\} \cdot n - \{x\} \cdot \{y\} \end{pmatrix}$$

$\Rightarrow$

$$\det \begin{pmatrix} n & \{x\} \\ \{x\} & \{x^2\} \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} b \\ a \end{pmatrix} = \begin{pmatrix} \det \begin{pmatrix} \{y\} & \{x\} \\ \{x \cdot y\} & \{x^2\} \end{pmatrix} \\ \det \begin{pmatrix} n & \{y\} \\ \{x\} & \{x \cdot y\} \end{pmatrix} \end{pmatrix}$$

$\Rightarrow$

$$|M| \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} b \\ a \end{pmatrix} = \begin{pmatrix} |b \leftrightarrow| \\ |a \leftrightarrow| \end{pmatrix}$$

Wobei letztendlich das der Lösungsweg über die Determinanten darstellt.

8.3 Nichtlineare Regressionen als Ableitungen der Linearregression

8.3.1 Potenzfunktionelle Regression

[001]ff. Definiert ist eine Potenzfunktion:

$$y = b \cdot x^a$$

Die Rückführung in eine lineare Funktion ist denkbar einfach.

$$\ln y = \ln (b \cdot x^a)$$

$\Rightarrow$

$$\ln y = \ln b + a \cdot \ln x \quad \leftrightarrow \quad Y = B + A \cdot X$$

Damit ist folgendes Vorangehen zwecks Regression gegeben.

• Die gegebenen zu regressierenden Datenpunkte $P_i\,(x_i; y_i)$ werden transformiert.

$$X_i = \ln x_i \qquad Y_i = \ln y_i$$

Die Datenpunkte liegen nun in der Form $P_i\,(X_i; Y_i)$ vor.

• Durchführen einer linearen Regression. Als Ergebnis sind die Koeffizienten A und B bekannt.

• (Rück)Transformierung der Linearkoeffizienten A und B in die potenzfunktionelle Koeffizienten a und b.

$$a = A$$

Sowie:

$$B = \ln b$$

$\Rightarrow$

$$b = e^B$$

Schon bei der Transformation der Datenpunkte ist ersichtlich, dass diese Regression nur angewandt werden kann, wenn y_i und x_i nichtnegative von Null verschiedene Werte besitzen.Im Besonderen ist bei jeder Transformation Vorsicht geboten, dass die gegeben Randbedingungen und Voraussetzungen eingehalten werden.

Selbstverständlich kann hier auch der dekadische Logarithmus $\log_{10} = \lg$ zur Transformation genutzt werden.

Für $b \in N$. Diese Regression ist eine konsequent „abgespeckte“ Polynomregression. Alle Koeffizienten außer a_b sind Null. So gilt für diese Regression in Matrixschreibweise:

$$|M| \cdot \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \cdot \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_b \end{pmatrix} = \begin{pmatrix} |a_0 \leftrightarrow| \\ |a_1 \leftrightarrow| \\ \vdots \\ |a_b \leftrightarrow| \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ |a_b \leftrightarrow| \end{pmatrix}$$

Schon aus der Komplexizität von $|a_i \leftrightarrow|$ ist mit der Forderung, dass diese Determinanten Null sein sollen ersichtlich, dass diese Regression selten den praktischen und realen Anforderungen genügt und mehr theoretischen bzw. Spezialfällen vorbehalten sein wird. Exemplarisch sei hier die einfachste Linearregression $y = x$ genannt. Mit $b = 1$ gilt:

$$|b \leftrightarrow| = 0 = \{x \cdot x\} \cdot \{y\} - \{x \cdot y\} \cdot \{x\}$$

Die Variable y wird durch x substituiert (oder umgekehrt).

$$0 = \{x \cdot x\} \cdot \{x\} - \{x \cdot x\} \cdot \{x\} = \{y \cdot y\} \cdot \{y\} - \{y \cdot y\} \cdot \{y\}$$

Was eine wahre Aussage liefert und mit $b = 2$ dann $y = x^2$ und höhere b genauso funktioniert.

8.3.2 Exponentielle Regression

Definiert ist eine Exponentialfunktion: [001]ff.

$$y = b \cdot e^{a \cdot x}$$

Die Rückführung in eine lineare Funktion ist denkbar einfach.

$$\ln y = \ln \left(b \cdot e^{a \cdot x}\right)$$

$\Rightarrow$

$$\ln y = \ln b + a \cdot x \quad \leftrightarrow \quad Y = B + A \cdot X$$

Damit ist folgendes Vorangehen zwecks Regression gegeben.

• Die gegebenen zu regressierenden Datenpunkte $P_i\left(x_i; y_i\right)$ werden transformiert.

$$X_i = x_i \qquad Y_i = \ln y_i$$

Die Datenpunkte liegen nun in der Form $P_i\left(X_i; Y_i\right)$ vor.

• Durchführen einer linearen Regression. Als Ergebnis sind die Koeffizienten A und B bekannt.

•(Rück)Transformierung der Linearkoeffizienten A und B in die exponentialfunktionelle Koeffizienten a und b.

$$a = A$$

Sowie:

$$B = \ln b$$

$\Rightarrow$

$$b = e^B$$

Schon bei der Transformation der Datenpunkte ist ersichtlich, dass diese Regression nur angewandt werden kann, wenn y_i nichtnegative von Null verschiedene Werte besitzt. Im Besonderen ist bei jeder Transformation Vorsicht geboten, dass die gegeben Randbedingungen und Voraussetzungen eingehalten werden.

8.3.3 Rationalfunktionelle Regression

[001]ff. Definiert ist eine rationale Funktion:

$$y = \frac{b \cdot x}{a + x}$$

Die Rückführung zu einer linearen Funktion ist denkbar einfach.

$$\frac{1}{y} = \frac{1}{b} + \frac{a}{b} \cdot \frac{1}{x} \quad \leftrightarrow \quad Y = B + A \cdot X$$

Damit ist folgendes Vorangehen zwecks Regression gegeben.

• Die gegebenen, sigmoidal zu regressierenden Datenpunkte $P_i\left(x_i; y_i\right)$ werden transformiert.

$$X_i = \frac{1}{x_i} \qquad Y_i = \frac{1}{y_i}$$

Die Datenpunkte liegen nun in der Form $P_i\left(X_i; Y_i\right)$ vor.

• Durchführen einer linearen Regression. Als Ergebnis sind die Koeffizienten A und B bekannt.

• (Rück)Transformierung der Linearkoeffizienten A und B in die rationalfunktionelle Koeffizienten a und b.

$$b = \frac{1}{B}$$

Sowie:

$$A = \frac{a}{b}$$

$\Rightarrow$

$$a = A \cdot b$$

Der rationalfunktionelle Koeffizient b ist bekannt, wird eingesetzt und umgeformt. Ergebnis ist eine Berechnungsgrundlage für a.

$$a = \frac{A}{B}$$

Die betrachtete rationale Funktion besitzt im Zähler kein Absolutglied. Folge ist, dass für ein $x = 0$ auch ein $y = 0$ gilt. Daher müssen die zu regressierenden Datenpunkte sich um $P\left(0; 0\right)$ versammeln, bzw. weit entfernt von diesem sein, um dem Regressionsgrafen es zu ermöglichen, sich vom Koordinatenursprung hinaus in die Datenwolke hinein zu regressieren. Im Besonderen ist bei jeder Transformation Vorsicht geboten, dass die gegeben Randbedingungen und Voraussetzungen eingehalten werden.

8.3.4 Sigmoide Regression nach Mitscherlich

Definiert ist eine sigmoidale Wachstumsfunktion, auch nach dessen Beschreiber Mitscherlich- Funktion [1] genannt: [001]ff.

$$y = \frac{a \cdot b}{a + (b-a) \cdot e^{-c \cdot x}}$$

Die Rückführung zu einer linearen Funktion ist denkbar einfach.

$$\frac{1}{y} = \frac{1}{b} + \frac{b-a}{a \cdot b} \cdot e^{-c \cdot x} \quad \leftrightarrow \quad Y = B + A \cdot X$$

Damit ist folgendes Vorangehen zwecks Regression gegeben.

• Die gegebenen, sigmoidal zu regressierenden Datenpunkte $P_i\,(x_i; y_i)$ werden transformiert.

$$X_i = e^{-c \cdot x_i} \qquad Y_i = \frac{1}{y_i}$$

Der Koeffizient c ist unkritisch und dient lediglich der Skalierung zu kleiner oder zu großer Beträge von x. Die Datenpunkte liegen nun in der Form $P_i\,(X_i; Y_i)$ vor.

• Durchführen einer linearen Regression. Als Ergebnis sind die Koeffizienten A und B bekannt.

• (Rück)Transformierung der Linearkoeffizienten A und B in die Mitscherlich-Koeffizienten a und b.

$$b = \frac{1}{B}$$

Sowie:

$$A = \frac{b-a}{a \cdot b}$$

$\Rightarrow$

$$a = \frac{b}{b \cdot A + 1}$$

Der Mitscherlich- Koeffizient b ist bekannt, wird eingesetzt und umgeformt. Ergebnis ist eine Berechnungsgrundlage für a.

$$a = \frac{1}{A + B}$$

Sind die Punkte $P_i\,(x_i; y_i)$ streng sigmoidal um den Regressionsgrafen der Mitscherlich-Funktion angeordnet, dann kann diese Funktion als Verteilungsfunktion F angesehen werden. Wird diese differenziert, erhält man eine Dichtefunktion f. Auf diesem Wege ist es leicht möglich einen Algorithmus zu implementieren um f zu erhalten.

$$F\,(x; a; b; c) = \frac{a \cdot b}{a + (b-a) \cdot e^{-c \cdot x}}$$

$\Rightarrow$

$$f\,(x; a; b; c) = F\,(x; a; b; c) \cdot \frac{(b-a) \cdot c \cdot e^{-c \cdot x}}{a + (b-a) \cdot e^{-c \cdot x}}$$

[1] Eilhard Alfred Mitscherlich, 1874-1956, Pflanzenbauwissenschaftler und Bodenkundler, beschrieb das mathematisch formulierte Wirkungsgesetz der Wachstumsfaktoren

9 Programmierbeispiele für Regressionen

9.1 Für die Biquadratische Regression

```
' Berechnungsgrundlage mit „0.01“ skaliert infolge numerischen Überlaufs großer Datensätze
' Vordefinition der Datensätze
'
xxxxy = 0.0;
xxxy = 0.0;
xxy = 0.0;
xy = 0.0;
x = 0.0;
y = 0.0;
xx = 0.0;
xxx = 0.0;
xxxx = 0.0;
xxxxx = 0.0;
xxxxxx = 0.0;
xxxxxxx = 0.0;
xxxxxxxx = 0.0;
'
for ( z = 0 ; z < n ; z = z + 1 )
{
xxxxy = xxxxy + ( Math.pow(0.01 * o[z] , 4 ) * 0.01 * w[z]);
xxxy = xxxy + ( Math.pow(0.01 * o[z] , 3 ) * 0.01 * w[z]);
xxy = xxy + ( Math.pow(0.01 * o[z] , 2 ) * 0.01 * w[z]);
xy = xy + (0.01 * o[z] * 0.01 * w[z]);
x = x + (0.01 * o[z]);
y = y + (0.01 * w[z]);
xx = xx + Math.pow(0.01 * o[z] , 2);
xxx = xxx + Math.pow(0.01 * o[z] , 3);
xxxx = xxxx + Math.pow(0.01 * o[z] , 4 );
xxxxx = xxxxx + Math.pow(0.01 * o[z] , 5 );
xxxxxx = xxxxxx + Math.pow(0.01 * o[z] , 6 );
xxxxxxx = xxxxxxx + Math.pow(0.01 * o[z] , 7 );
xxxxxxxx = xxxxxxxx + Math.pow(0.01 * o[z] , 8 );
}
'
' Regressionskern
' Diskriminantenberechnung - bewirkt kleinere Rundungsfehler durch Skalierungsmöglichkeiten
' Kolonne der 1. Diskriminantenschicht
'
Y = y / n;
X = xxxx / n;
W = xxx / n;
V = xx / n;
U = x / n;
'
' Kolonne der 2. Diskriminantenschicht
'
A = U - ( xx / x );
B = V - ( xxx / x );
C = W - ( xxxx / x );
D = X - ( xxxxx / x );
E = Y - ( xy / x );
F = U - ( xxx / xx );
G = V - ( xxxx / xx );
H = W - ( xxxxx / xx );
```

```
I = X - ( xxxxxx / xx );
J = Y - ( xxy / xx );
K = U - ( xxxx / xxx );
L = V - ( xxxxx / xxx );
M = W - ( xxxxxx / xxx );
N = X - ( xxxxxxx / xxx );
O = Y - ( xxxy / xxx );
P = U - ( xxxxx / xxxx );
Q = V - ( xxxxxx / xxxx );
R = W - ( xxxxxxx / xxxx );
S = X - ( xxxxxxxx / xxxx );
T = Y - ( xxxxy / xxxx );
'
' Kolonne der 3. Diskriminantenschicht
'
AA = ( B / A ) - ( G / F );
BB = ( C / A ) - ( H / F );
CC = ( D / A ) - ( I / F );
DD = ( E / A ) - ( J / F );
EE = ( B / A ) - ( L / K );
FF = ( C / A ) - ( M / K );
GG = ( D / A ) - ( N / K );
HH = ( E / A ) - ( O / K );
II = ( B / A ) - ( Q / P );
JJ = ( C / A ) - ( R / P );
KK = ( D / A ) - ( S / P );
LL = ( E / A ) - ( T / P );
'
' Kolonne der 4. Diskriminantenschicht
'
AAA = ( BB / AA ) - ( FF / EE );
BBB = ( CC / AA ) - ( GG / EE );
CCC = ( DD / AA ) - ( HH / EE );
DDD = ( BB / AA ) - ( JJ / II );
EEE = ( CC / AA ) - ( KK / II );
FFF = ( DD / AA ) - ( LL / II );
'
' Kolonne der 5. Diskriminantenschicht
'
AAAA = ( BBB / AAA ) - ( EEE / DDD );
BBBB = ( CCC / AAA ) - ( FFF / DDD );
'
' Koeffizientenberechnung
'
a4 = BBBB / AAAA;
b4 = ( CCC / AAA ) - ( ( BBB / AAA ) * a4 );
c4 = ( DD / AA ) - ( ( CC / AA ) * a4 ) - ( ( BB / AA ) * b4 );
d4 = ( E / A ) - ( ( D / A ) * a4 ) - ( ( C / A ) * b4 ) - ( ( B / A ) * c4 );
e4 = Y - ( X * a4 ) - ( W * b4 ) - ( V * c4 ) - ( U * d4 );
'
' Entskalierung der Koeffizienten
'
e4 = e4 / 0.01;
c4 = c4 * 0.01;
b4 = b4 * 0.01 * 0.01;
a4 = a4 * 0.01 * 0.01 * 0.01;
```

9.2 Für die Sigmoide Regression nach Mitscherlich

```
' Umbau der Rohdaten in Mitscherlich- kompatible Daten
'
for ( z = 0 ; z < n ; z = z + 1 )
{
o_inv[z] = Math.pow(e , ( (-o[z]) / alpha ) );
w_inv[z] = Math.pow( w[z] , -1 );
}
'
' Beginn Lineare Regression der Mitscherlich Daten
' Vordefinition der Datensätze
'
xy = 0.0;
x = 0.0;
y = 0.0;
xx = 0.0;
'
for ( z = 0 ; z < n ; z = z + 1 )
{
xy = xy + (o_inv[z] * w_inv[z]);
x = x + o_inv[z];
y = y + w_inv[z];
xx = xx + (o_inv[z] * o_inv[z]);
}
'
' Koeffizientenberechnung der Mitscherlich- Daten
'
a6 = ( ( xy * n ) - ( x * y) ) / ( ( xx * n ) - ( x * x )
b6 = ( ( xx * y ) - (xy * x) ) / ( ( xx * n ) - ( x * x ) );
'
' Koeffizientenüberführung in kartesisch- kompatible Koeffizienten
'
A6 = Math.pow( (a6 + b6) , -1);
B6 = Math.pow( b6 , -1);
```

$\LaTeX\,2_\varepsilon$