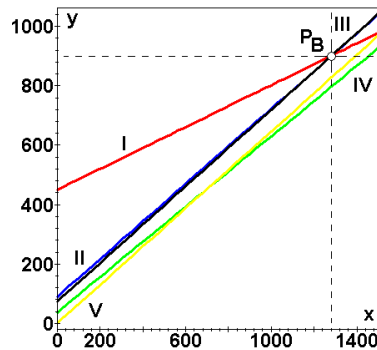


Reduzierte Lineare Regression



Reduced linear regression

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

www.Zenithpoint.de

Erstellt: 10. Juni 2017 - Letzte Revision: 5. Februar 2026

Inhaltsverzeichnis

1	Lineare Regression ohne Anstieg	3
1.1	MKQ	4
1.2	HKA	5
2	Lineare Regression ohne Inhomogenität	7
2.1	MKQ	8
2.2	HKA	9
3	Lineare Regression definierter Punkt	11
3.1	Inhomogenität aus Anstieg	12
3.2	Anstieg aus Inhomogenität	13
4	Beispiel	15

Literatur

- [Dipa] Dipl.-Ing. Björnstjerne Zindler, M.Sc. Elliptische Regression von Datenpunkten - HKA.
www.Zenithpoint.de.
- [Dipb] Dipl.-Ing. Björnstjerne Zindler, M.Sc. (Polynom)Regression von Datenpunkten.
www.Zenithpoint.de.
-

1 Lineare Regression ohne Anstieg

Ziel ist die Herleitung der Durchführung einer Linearen Regression **ohne Anstieg** a . So gilt allgemein für eine lineare Funktion:

Ohne
Anstieg

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

⇒

$$y = b$$

MKQ

1.1 Nach der Methode der kleinsten Quadrate - MKQ

Ziel ist es, die Fehlerfunktion F zu minimieren. Die Herleitung von F ist unter [Dipa] nachzulesen.

$$F(x, a, b) = \{y^2\} - 2a \cdot \{xy\} - 2b \cdot \{y\} + 2ab \cdot \{x\} + a^2 \{x^2\} + n \cdot b^2$$

Diese vereinfacht sich für ein $a = 0$ beträchtlich.

$$F(x, b) = \{y^2\} - 2b \cdot \{y\} + n \cdot b^2$$

Ein Extrema wird gesucht um später ein Minimum nachweisen zu können.

$$\frac{d}{db} F(x, b) = -2 \cdot \{y\} + 2 \cdot n \cdot b = 0$$

⇒

$$b = \frac{\{y\}}{n}$$

Der Nachweis, dass ein Minimum gefunden wurde.

$$\frac{d^2}{db^2} F(x, b) = 2 \cdot n > 0$$

Das Minimum wurde gefunden und die Regression ist beendet.

$$y = \frac{1}{n} \cdot \sum_{i=1}^n y_i$$

1.2 Nach der Methode der Hauptkomponentenanalyse - HKA

HKA

Aus der HKA ist eine Darstellung der linearen Funktion der Hauptachse bekannt.

$$y = \underbrace{\frac{V_{YY} - \lambda}{C}}_a \cdot x + \bar{y} - \underbrace{\frac{V_{YY} - \lambda}{C}}_b \cdot \bar{x}$$

Die Herleitung und Bedeutung ist nachzulesen unter [Dipb].

Es soll $a = 0$ gesetzt werden.

$$\frac{V_{YY} - \lambda}{C} = 0$$

\Rightarrow

$$y = \bar{y} = \frac{\{y\}}{n}$$

Damit ist b bekannt und die Regression beendet.

$$y = \frac{1}{n} \cdot \sum_{i=1}^n y_i$$

2 Lineare Regression ohne Inhomogenität

Ziel ist die Herleitung der Durchführung einer Linearen Regression **ohne Inhomogenität** b . So gilt allgemein für eine lineare Funktion:

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

⇒

$$y = a \cdot x$$

Ohne
Inhomogenität

MKQ

2.1 Nach der Methode der kleinsten Quadrate - MKQ

Ziel ist es, die Fehlerfunktion F zu minimieren. Die Herleitung von F ist unter [Dipa] nachzulesen.

$$F(x, a, b) = \{y^2\} - 2a \cdot \{xy\} - 2b \cdot \{y\} + 2ab \cdot \{x\} + a^2 \{x^2\} + n \cdot b^2$$

Diese vereinfacht sich für ein $b = 0$ beträchtlich.

$$F(x, a) = \{y^2\} - 2a \cdot \{xy\} + a^2 \{x^2\}$$

Ein Extrema wird gesucht um später ein Minimum nachweisen zu können.

$$\frac{d}{da} F(x, a) = -2 \cdot \{xy\} + 2a \cdot \{x^2\} = 0$$

⇒

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

Der Nachweis, dass ein Minimum gefunden wurde.

$$\frac{d^2}{da^2} F(x, a) = 2 \cdot \{x^2\} > 0$$

Das Minimum wurde gefunden und die Regression ist beendet.

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

2.2 Nach der Methode der Hauptkomponentenanalyse - HKA

HKA

Aus der HKA ist eine Darstellung der linearen Funktion der Hauptachse bekannt.

$$y = \underbrace{\frac{V_{YY} - \lambda}{C}}_a \cdot x + \bar{y} - \underbrace{\frac{V_{YY} - \lambda}{C}}_b \cdot \bar{x}$$

Die Herleitung und Bedeutung ist nachzulesen unter [Dipb].

Es soll $b = 0$ gesetzt werden.

$$\bar{y} - \frac{V_{YY} - \lambda}{C} \cdot \bar{x} = 0$$

\Rightarrow

$$\frac{V_{YY} - \lambda}{C} = \frac{\bar{y}}{\bar{x}}$$

Damit ist a bekannt und die Regression beendet.

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

3 Lineare Regression definierter Punkt

Es soll eine Lineare Regression durchgeführt werden, wobei ein Punkt $P_B(x_B, y_B)$ unbedingt auf der Regressionsgeraden liegen soll. Grundlage für die Funktion sei die Punkt- Richtungs- Form.

Definierter
Punkt

$$\frac{y - y_B}{x - x_B} = a'$$

\Rightarrow

$$y = \underbrace{a'}_a \cdot x + \underbrace{y_B - a' \cdot x_B}_b$$

Damit sind zwei Definitionen der gesuchten Regressionsgeraden möglich.

A → I

3.1 Ermittlung der Inhomogenität aus Anstieg

Der Anstieg a wird berechnet aus den Daten der Urliste mittels einer Linearen Regression ohne Inhomogenität b . Diese wird berechnet durch:

$$b = y_B - a \cdot x_B$$

⇒

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

3.2 Ermittlung des Anstiegs aus Inhomogenität

I → A

Die Inhomogenität b wird berechnet aus den Daten der Urliste mittels einer Linearen Regression ohne Anstieg a . Diese wird berechnet durch:

$$a = \frac{y_B - b}{x_B}$$

⇒

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

4 Beispiel

Beispiel

Gegeben sind Datenpunkte aus einer Messung (aus [Dipb] und [Dipa]).

i	X_i	Y_i	$X_i \cdot X_i$	$X_i \cdot Y_i$
1	128	100	16 384	12 800
2	256	250	65 536	64 000
3	440	210	193 600	224 400
4	640	160	409 600	102 400
5	768	400	589 824	307 200
6	896	520	802 816	465 920
7	1 152	750	1 327 104	864 000
8	1 280	900	1 638 400	1 152 000
Σ	5 560	3 590	5 043 264	3 192 720

Laut Festlegung soll der im 1. Quadranten äußerste Punkt unbedingt auf der Regressionsgeraden liegen.

$$P_B = P_8(1280; 900)$$

• **Für den Anstieg $a' = 0$**

Es wird zuerst $a' = 0$ gesetzt. Dadurch ist die Inhomogenität b bekannt und a berechenbar.

$$b = \frac{1}{n} \cdot \sum_{i=1}^n y_i = \frac{3590}{8} = 448,75 \quad \rightarrow \quad a = \frac{y_B - b}{x_B} = \frac{900 - 448,75}{1280} = 0,3525$$

\Rightarrow

$$y_{a'=0} = y_I = 0,3525 \cdot x + 448,75$$

• **Für die Inhomogenität $b' = 0$ und MKQ**

Es wird zuerst $b' = 0$ gesetzt. Dadurch ist der Anstieg a bekannt und b berechenbar.

$$a = \frac{\sum_{i=1}^n (x_i \cdot y_i)}{\sum_{i=1}^n x_i^2} = \frac{3192720}{5043264} = 0,6331 \rightarrow b = y_B - a \cdot x_B = 900 - 0,6331 \cdot 1280 = 89,632$$

\Rightarrow

$$y_{MKQ, b'=0} = y_{II} = 0,6331 \cdot x + 89,632$$

• **Für die Inhomogenität $b' = 0$ und HKA**

Es wird zuerst $b' = 0$ gesetzt. Dadurch ist der Anstieg a bekannt und b berechenbar.

$$a = \frac{\sum_{i=1}^n y_i}{\sum_{i=1}^n x_i} = \frac{3590}{5560} = 0,6457 \rightarrow b = y_B - a \cdot x_B = 900 - 0,6457 \cdot 1280 = 73,504$$

\Rightarrow

$$y_{HKA, b'=0} = y_{III} = 0,6457 \cdot x + 73,504$$

• Die Regressionsergebnisse der vollständigen Methoden

Aus [Dipb] und [Dipa] sind die Regressionsfunktionen aus den vollständigen Methoden bekannt.

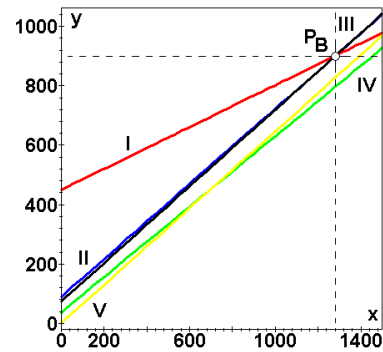
$$y_{MKQ} = y_{IV} = 0,5928 \cdot x + 37,5079$$

Sowie:

$$y_{HKA} = y_V = 0,6457 \cdot x$$

• Grafische Darstellung

Die gefundenen Funktionen grafisch dargestellt.



Dabei stellen folgende Farben die Funktionen
y_I, *y_{II}*, *y_{III}*, *y_{IV}*, *y_V* dar.