Zentrieren und Rückkippen einer Ellipse, gewonnen aus der Regression nach der Hauptkomponentenanalyse

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

www.Zenithpoint.de

Erstellt: 6. Juli 2017 – Letzte Revision: 27. Januar 2021

Inhaltsverzeichnis

Literatur

[Dipa] Dipl.- Ing. Björnstjerne Zindler, M.Sc. Elliptische Regression von Datenpunkten. www.Zenithpoint.de.

[Dipb] Dipl.- Ing. Björnstjerne Zindler, M.Sc. Elliptische Regression von Datenpunkten über die Hauptkomponentenanalyse. www.Zenithpoint.de.

[Dipc] Dipl.- Ing. Björnstjerne Zindler, M.Sc. Ermittlung der abszissen und ordinatenparallelen Ellipse über die Singularwertzerlegung. www.Zenithpoint.de.

---

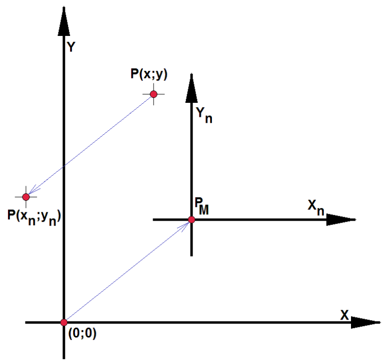
1 Einleitung zum Thema

Ziel dieses Arbeitsblattes ist es, die Elliptische Regression über die Hauptkomponentenanalyse (HKA) in der Ebene zu vervollständigen. In [Dipb] war am Ende eine unzentrierte und gekippte Ellipse als Ergebnis ermittelt. Diese soll zentriert und zurückgekippt werden. Die Eigenschaften der Varianzen als ein Indikator wird aufgezeigt.

2 Durchführung der Drehung und Zentrierung

2.1 Verschiebung des Systems in den Ellipsenmittelpunkt $P_M\,(x_M, y_M)$

Der erste Schritt ist es, die Ellipse zu zentrieren. Dazu wird das Koordinatensystem in den Mittelpunkt der Ellipse verschoben, Bekannt sind der Mittelpunkt der Ellipse mit $P_M\,(x_M, y_M)$



und

$$x_M = \bar{X} = \frac{\{X_i\}}{n} \qquad y_M = \bar{Y} = \frac{\{Y_i\}}{n}$$

aus der Verschiebung ergeben sich neue Datenpunktangaben $P_i\,(x_i; y_i)$ mit:

$$x_n = x - x_M \qquad y_n = y - y_M$$

Beispiel aus [Dipb]:

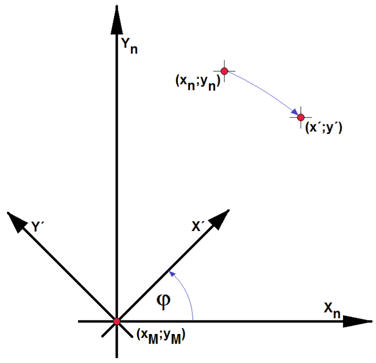
i	x_i	y_i	x_n	y_n
1	128	100	-567	-349
2	256	250	-439	-199
3	440	510	-255	+61
4	640	160	-55	-289
5	768	400	+73	-49
6	896	520	+201	+71
7	1152	750	+457	+301
8	1280	900	+585	+451
$\sum$	5560	3590	0	≈ 0

Mit:

$$x_M = \frac{5560}{8} = 695 \qquad y_M = \frac{3590}{8} = 449$$

2.2 Drehung des Systems auf die Abszisse $Y^{(b=0)}$

Die nun zentrierte jedoch noch gekippte Ellipse wird nun abszissenparallel gestellt. Dazu wird der bekannte Kippwinkel φ der Ellipse genutzt. Die Vorzeichen werden so gesetzt, dass eine Drehung gegen den Uhrzeigersinn ein positives φ erfordert.



$\Rightarrow$

$$x' = x_n \cdot \cos\varphi + y_n \cdot \sin\varphi \qquad y' = y_n \cdot \cos\varphi - x_n \cdot \sin\varphi$$

Auch für den Winkel φ gibt es aus der HKA eine Berechnungsgrundlage.

$$\varphi = \arctan\frac{y_M}{x_M}$$

$\Rightarrow$

$$x' = \frac{x_n \cdot x_M + y_n \cdot y_M}{\sqrt{x_M^2 + y_M^2}} \qquad y' = \frac{y_n \cdot x_M - x_n \cdot y_M}{\sqrt{x_M^2 + y_M^2}}$$

Damit ist die Ellipse zentriert und ungekippt. Das genutzte Beispiel wird weiter entwickelt.

i	x_n	y_n	x'	y'
1	-567	-349	-666	+13
2	-439	-199	-477	+70
3	-255	+61	-181	+189
4	-55	-289	-203	-213
5	+73	-49	+35	-81
6	+201	+71	+207	-49
7	+457	+301	+547	+6
8	+585	+451	+736	+63
$\sum$	0	≈ 0	≈ 0	≈ 0

2.3 Zusammenfassen von Verschiebung und Drehung

Die Berechnungsgrundlagen für die Verschiebung und Drehung können zusammen gefasst werden. Daraus ergeben sich die folgenden Gesamtgleichungen:

$$x_n = x - x_M \quad y_n = y - y_M \quad \text{mit} \quad x' = \frac{x_n \cdot x_M + y_n \cdot y_M}{\sqrt{x_M^2 + y_M^2}} \quad y' = \frac{y_n \cdot x_M - x_n \cdot y_M}{\sqrt{x_M^2 + y_M^2}}$$

$\Rightarrow$

$$x' = \frac{(x - x_M) \cdot x_M + (y - y_M) \cdot y_M}{\sqrt{x_M^2 + y_M^2}}$$
$$y' = \frac{(y - y_M) \cdot x_M - (x - x_M) \cdot y_M}{\sqrt{x_M^2 + y_M^2}}$$

2.4 Neuermittlung von Haupt- und Nebenachse $Y^{(\varphi=0°)}, Y^{(\varphi=90°)}$

Die allgemeine Berechnungsgrundlage der Haupt- und Nebenachse der Ellipse ist aus [Dipa] gegeben mit:

$$Y^{(\varphi,b=0)} = \frac{\bar{Y}}{\bar{X}} \cdot X \qquad Y^{(\varphi,b\neq 0)} = \left(\frac{V_{YY} - V_{XX}}{C} - \frac{\bar{Y}}{\bar{X}}\right) \cdot X + \frac{V_{XX} - V_{YY}}{C} \cdot \bar{X} + 2 \cdot \bar{Y}$$

Die Achse $Y^{(\varphi,b=0)}$ ist aus den Werten $\bar{X}$ und $\bar{Y}$ nicht definiert, da beide nach der Verschiebung den Wert Null ergeben. Durch die anschließende Drehung wurde jedoch festgelegt:

$$Y^{(\varphi=0,b=0)} = \frac{\bar{Y}}{\bar{X}} \overset{def}{=} 0$$

Damit kann dann auch $Y^{(\varphi=0,b\neq 0)}$ ermittelt werden.

$$Y^{(\varphi,b\neq 0)} = \frac{V_{YY} - V_{XX}}{C} \cdot X - \frac{\bar{Y}}{\bar{X}} \cdot X + \frac{V_{XX} - V_{YY}}{C} \cdot \bar{X} + 2 \cdot \bar{Y}$$

$\Rightarrow$

$$Y^{(\varphi=0,b\neq 0)} = \frac{V_{YY} - V_{XX}}{C} \cdot X$$

Ideal müsste die Achse $Y^{(\varphi=0,b\neq 0)}$ senkrecht auf $Y^{(\varphi=0,b=0)}$ stehen. Daraus folgt, dass $C \rightarrow 0$ geht. Das ist jedoch nur dann der Fall, wenn die Varianzen der Originaldaten vollständig aufgelöst werden können. Der Winkel zwischen den Achsen ist demnach ein Indikator der Regression und der Hauptkomponentenanalyse.

$$\tan \varphi = \frac{V_{YY} - V_{XX}}{C}$$

3 Eigenschaft und Rolle von $\tan\varphi$

Die allgemeine Berechnungsgrundlage der Haupt- und Nebenachse der Ellipse ist aus [Dipa] ist gegeben mit:

$$Y^{(\varphi,b=0)} = \frac{\bar{Y}}{\bar{X}} \cdot X \qquad Y^{(\varphi,b\neq 0)} = \left(\frac{V_{YY}-V_{XX}}{C} - \frac{\bar{Y}}{\bar{X}}\right) \cdot X + \frac{V_{XX}-V_{YY}}{C} \cdot \bar{X} + 2 \cdot \bar{Y}$$

Definiert man als Hauptachse die Achse, welche durch den Koordinatenursprung verläuft $Y^{(\varphi,b=0)}$ und als Nebenachse, die orthogonal dazu verlaufende $Y^{(\varphi,b\neq 0)}$, dann gilt für die Anstiege:

$$Y^{(\varphi,b=0)} = -\frac{1}{Y^{(\varphi,b\neq 0)}}$$

$\Rightarrow$

$$-\frac{\bar{X}}{\bar{Y}} = \frac{V_{YY}-V_{XX}}{C} - \frac{\bar{Y}}{\bar{X}}$$

$\Rightarrow$

$$\frac{\bar{Y}}{\bar{X}} - \frac{\bar{X}}{\bar{Y}} = \tan\varphi = \frac{V_{YY}-V_{XX}}{C}$$

Diese Aussage ist für die zentrierte Ellipse richtig, da $x_M = y_M = 0$ gilt. Damit gilt für den Normalfall und den vordefinierten Quotienten:

$$\frac{\bar{Y}}{\bar{X}} = 0 \qquad \frac{\bar{X}}{\bar{Y}} = \pm\infty$$

$\Rightarrow$

$$\tan\varphi = \pm\infty \qquad V_{YY} - V_{XX} \neq 0 \qquad C \to 0$$

Das Pendant dazu hat folgende Form:

$$V_{YY} - V_{XX} = 0 \qquad C \neq 0$$

$\Rightarrow$

$$\tan\varphi = 0$$

Damit sind zwei globale Fälle unterscheidbar.

- $\tan\varphi = \pm\infty$ Die Haupt- und die Nebenachse sind orthogonal zueinander. Die Varianzen sind auflösbar / unterscheidbar.
- $\tan\varphi = 0$ Die Haupt- und die Nebenachse sind identisch. Es liegt eine Kreisregression vor. Die Varianzen sind nicht auflösbar / unterscheidbar.

4 Ermitteln der neuen Ellipsenfunktion $Y_{1;2}^{(\varphi=x_M=y_M=0)}$

Aus den gewonnen Werten x' und y' kann die Ellipsenfunktion ermittelt werden. Die allgemeine Berechnungsgrundlage aus [Dipa] ist gegeben mit:

$$Y_{1;2}^{(\varphi)} = y_M + \frac{B}{A} \cdot (x - x_M) \pm \frac{f \cdot e}{A} \cdot \sqrt{A - (x - x_M)^2}$$

Der Kippwinkel φ der neugewonnenen Ellipse ist Null, da diese ungekippt nun vorliegt. Gleichzeitig fällt x_M und y_M weg, da die Zentrierung abgeschlossen ist.

$$Y_{1;2}^{(\varphi=x_M=y_M=0)} = \frac{B}{A} \cdot x \pm \frac{f \cdot e}{A} \cdot \sqrt{A - x^2}$$

Für ein $\varphi = 0$ ist der Anstieg einer Achse $a = 0$. Für die Koeffizienten A und B ist dann bekannt:

$$A = \frac{e^2 \cdot a^2 + f^2}{1 + a^2} \qquad B = a \cdot \frac{f^2 - e^2}{1 + a^2}$$

$\Rightarrow$

$$A^{(\varphi=0)} = f^2 \qquad B^{(\varphi=0)} = 0$$

$\Rightarrow$

$$Y_{1;2}^{(\varphi=x_M=y_M=0)} = \pm\frac{e}{f} \cdot \sqrt{f^2 - x^2}$$

Die Haupt- und die Nebenachse ist definiert.

$$e^2 = \frac{V_{YY}}{n} \qquad f^2 = \frac{V_{XX}}{n}$$

$\Rightarrow$

$$Y_{1;2}^{(\varphi=x_M=y_M=0)} = \pm\sqrt{\frac{V_{YY}}{V_{XX}}} \cdot \sqrt{\frac{V_{XX}}{n} - x^2}$$

Für vorliegendes Beispiel ergibt sich somit ein $Y_{1;2}^{(\varphi=x_M=y_M=0)}$ von:

i	x'	y'	x'^2	y'^2	$x' \cdot y'$
1	-666	+13	443.556	169	-8.658
2	-477	+70	227.529	4.900	-33.390
3	-181	+189	32.761	35.721	-34.209
4	-203	-213	41.209	45.369	+43.239
5	+35	-81	1.225	6.561	-2.835
6	+207	-49	42.849	2.401	-10.143
7	+547	+6	299.209	36	+3.282
8	+736	+63	541.696	3.969	+46.368
$\sum$	0	≈ 0	1.630.034	99.126	+3.654

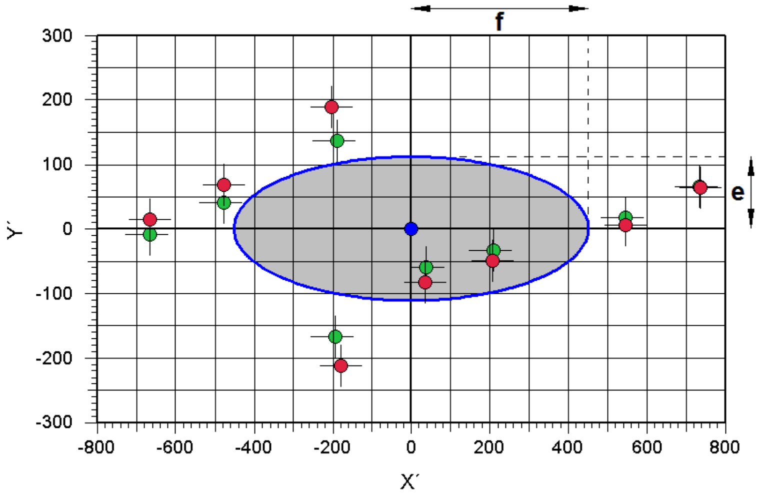
$$V_{XX} = 1630034 \qquad V_{YY} = 99126$$

$$C = C_{XY} = C_{YX} = 3654$$

$\Rightarrow$

$$Y_{1;2}^{(\varphi=x_M=y_M=0)} = \pm\sqrt{\frac{99126}{1630034}} \cdot \sqrt{\frac{1630034}{8} - x^2} = \pm\sqrt{12390,75 - 0,0608 \cdot x^2}$$

Die grafische Darstellung dazu folgend, wobei ROT ermittelte Datenpunkte aus der Hauptkomponentenanalyse (HKA) und GRÜN die nach der Methode der kleinsten Quadrate (MKQ) darstellt.



Mit:

$$e = \sqrt{\frac{V_{YY}}{n}} = 111,3 \qquad f = \sqrt{\frac{V_{XX}}{n}} = 451,4$$

Und:

$$\varphi = \arctan \frac{V_{YY} - V_{XX}}{C} = -89,9^\circ \equiv -0,499 \cdot \pi$$

Aus Interesse sollen die gedrehten (Ko)Varianzen ermittelt werden. Aus [Dipc] folgt:

$$V_{XX}^{(\varphi)} = \frac{V_{YY}}{n} \cdot \cos^2 \varphi + \frac{V_{XX}}{n} \cdot \sin^2 \varphi$$

$$V_{Y}^{(\varphi)} = \frac{V_{YY}}{n} \cdot \sin^2 \varphi + \frac{V_{XX}}{n} \cdot \cos^2 \varphi$$

$$C^{(\varphi)} = \sin \varphi \cdot \cos \varphi \cdot \left(\frac{V_{YY}}{n} - \frac{V_{XX}}{n} \right)$$

$\Rightarrow$

$$V_{XX}^{(\varphi=0)} = \frac{V_{YY}}{n} \qquad C^{(\varphi=0)} = 0 \qquad V_{YY}^{(\varphi=0)} = \frac{V_{XX}}{n}$$

Die gedrehten Varianzen sind gleich der ungedrehten. Die für die Ellipsendrehung verantwortliche gedrehte Kovarianz $C^{(\varphi=0)}$ ist gleich Null.

5 Zusammenfassung

Das Verschieben und Drehen eines Datensatzes im Zusammenhang mit der Elliptischen Regression nach HKA erfolgt durch folgende Berechnungsgrundlage.

$$x' = \frac{(x - x_M) \cdot x_M + (y - y_M) \cdot y_M}{\sqrt{x_M^2 + y_M^2}}$$
$$y' = \frac{(y - y_M) \cdot x_M - (x - x_M) \cdot y_M}{\sqrt{x_M^2 + y_M^2}}$$

Wobei $P_i\,(x_i; y_i)$ den alte Datenpunkt und $P_i\,(x_i'; y_i')$ den neuen darstellt. Es werden dazu lediglich die Mittelwerte x_M und y_M benötigt.

Die nun zentrierte und ungekippte Ellipse besitzt die Form:

$$Y_{1;2}^{(\varphi = x_M = y_M = 0)} = \pm\sqrt{\frac{V_{YY}}{V_{XX}}} \cdot \sqrt{\frac{V_{XX}}{n} - x^2}$$

Mit den Varianzen V_{XX} und V_{YY}. Die Haupt- und Nebenachse besitzen einen Winkel $var\phi$ zueinander. Dieser ist definiert durch:

$$\tan\varphi = \frac{V_{YY} - V_{XX}}{C}$$

Idealerweise ist dieser 90°. Bedeutet, dass die Kovarianz C des Datensatzes $P_i\,(x_i'; y_i')$ gegen Null geht. Ist das der Fall, konnten die Varianzen des Originaldatensatzes $P_i\,(x_i; y_i)$ vollständig aufgelöst werden.

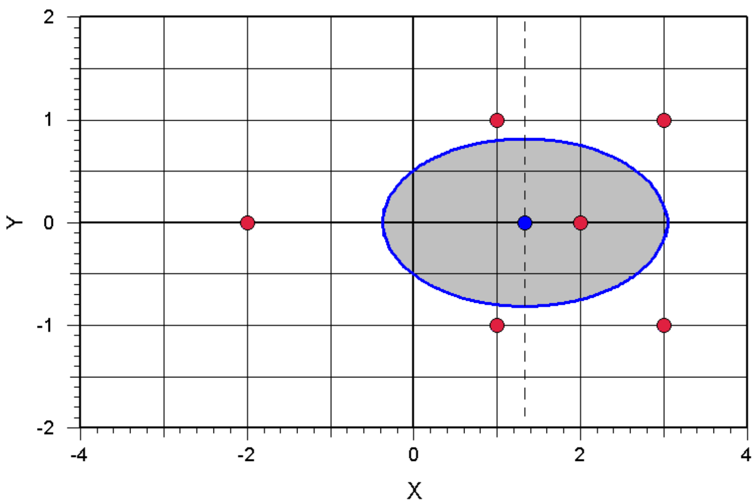
6 Beispiele

6.1 Beispiel I – vollständige Auflösung

Gegeben ist folgender Datensatz

i	x_i	y_i	$(x_i - x_M)^2$	$(y_i - y_M)^2$	$(x_i - x_M) \cdot (y_i - y_M)$
1	+3	-1	2,778	1	-1,667
2	-2	0	11,112	0	0,000
3	+1	+1	0,112	1	-0,334
4	+1	-1	0,112	1	+0,334
5	+2	0	0,445	0	0,000
6	+3	+1	2,778	1	+1,667
$\sum$	+8	0	17,337	4	0,000

$\Rightarrow$



$\Rightarrow$

$$V_{XX} = 17,337 \qquad C = 0 \qquad V_{YY} = 4$$

$\Rightarrow$

$$f = \sqrt{\frac{V_{XX}}{n}} = 1,700 \qquad e = \sqrt{\frac{V_{YY}}{n}} = 0,816$$

Mit:

$$y_H = 0 \qquad x_N = 1,334$$

Bei:

$$x_M = \frac{8}{6} = 1,334 \qquad y_M = \frac{0}{6} = 0$$

$\Rightarrow$

$$A = \frac{e^2 \cdot a^2 + f^2}{1 + a^2} = 2,890 \qquad B = a \cdot \frac{f^2 - e^2}{1 + a^2} = 0$$

$\Rightarrow$

$$Y_{1;2}^{(\varphi)} = y_M + \frac{B}{A} \cdot (x - x_M) \pm \frac{f \cdot e}{A} \cdot \sqrt{A - (x - x_M)^2}$$

$\Rightarrow$

$$Y_{1;2}^{(\varphi)} = \pm 0,480 \cdot \sqrt{2,890 - (x - 1,334)^2}$$

Sowie:

$$\tan\varphi = \frac{V_{YY} - V_{XX}}{C} = -\infty$$

$\Rightarrow$

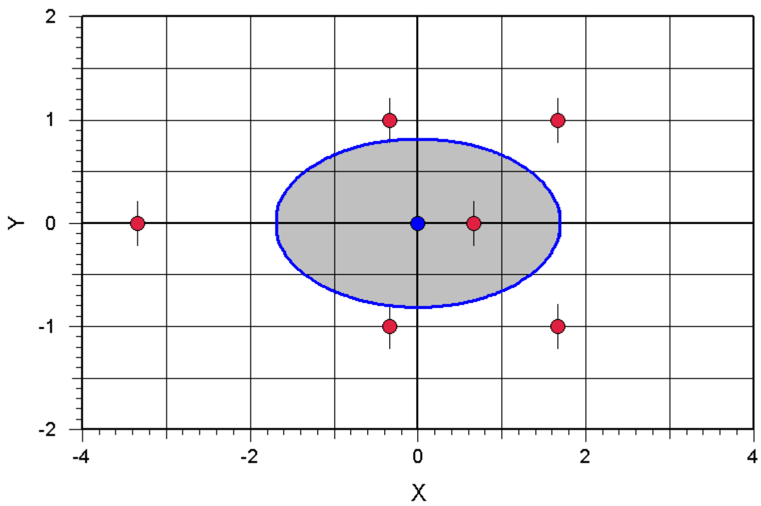
$$\varphi = 90^\circ$$

Damit können die Varianzen vollständig aufgelöst werden.

Die Ermittlung der Werte x' und y'.

i	x_i	y_i	x'_i	y'_i	$(x'_i - x'_M)^2$	$(y'_i - y'_M)^2$	$(x'_i - x'_M) \cdot (y'_i - y'_M)$
1	+3	-1	+1,667	-1	2,779	1	-1,667
2	-2	0	-3,334	0	11,116	0	0,000
3	+1	+1	-0,334	+1	0,112	1	-0,334
4	+1	-1	-0,334	-1	0,112	1	+0,334
5	+2	0	+0,667	0	0,445	0	0,000
6	+3	+1	+1,667	+1	2,779	1	+1,667
$\sum$	+8	0	0	0	17,334	4	0

$\Rightarrow$



$\Rightarrow$

$$V_{XX} = 17,334 \qquad C = 0 \qquad V_{YY} = 4$$

$\Rightarrow$

$$f = \sqrt{\frac{V_{XX}}{n}} = 1,700 \qquad e = \sqrt{\frac{V_{YY}}{n}} = 0,816$$

$\Rightarrow$

$$Y_{1;2}^{(\varphi = x_M = y_M = 0)} = \pm\sqrt{\frac{V_{YY}}{V_{XX}}} \cdot \sqrt{\frac{V_{XX}}{n} - x^2}$$

$\Rightarrow$

$$Y_{1;2}^{(\varphi = x_M = y_M = 0)} = \pm 0,480 \cdot \sqrt{2,890 - x^2}$$

Sowie:

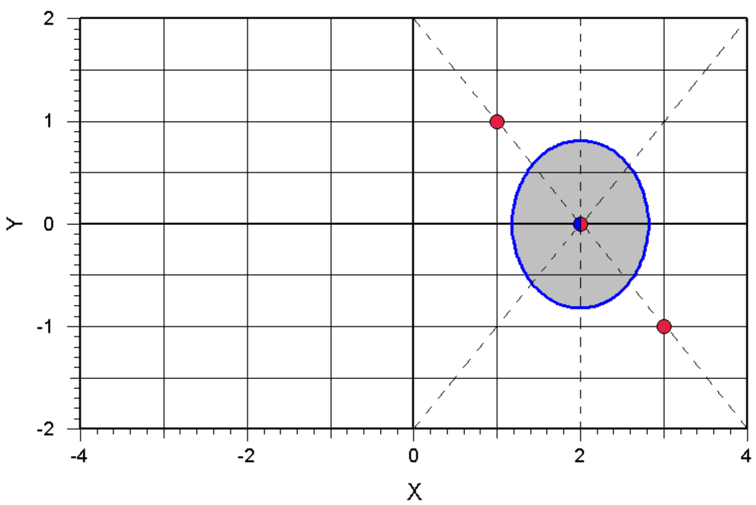
$$\tan\varphi = \frac{V_{YY} - V_{XX}}{C} = -\infty \qquad \rightarrow \qquad \varphi = 90^\circ$$

6.2 Beispiel II – unvollständige Auflösung

Gegeben ist folgender Datensatz

i	x_i	y_i	$(x_i - x_M)^2$	$(y_i - y_M)^2$	$(x_i - x_M) \cdot (y_i - y_M)$
1	3	-1	1	1	-1
2	2	0	0	0	0
3	1	+1	1	1	-1
4	1	+1	1	1	-1
5	2	0	0	0	0
6	3	-1	1	1	-1
$\sum$	12	0	4	4	-4

$\Rightarrow$



$\Rightarrow$

$$V_{XX} = 4 \qquad C = -4 \qquad V_{YY} = 4$$

Es liegt eine Kreisregression vor.

$$f = \sqrt{\frac{V_{XX}}{n}} = 0{,}816 \qquad e = \sqrt{\frac{V_{YY}}{n}} = 0{,}816$$

Mit Linearer Regression – MKQ:

$$y_{H;MKQ} = -x + 2 \qquad y_{N;MKQ} = x - 2$$

Mit Linearer Regression – HKA:

$$y_{H;HKA} = 0 \qquad x_{N;HKA} = 2$$

$\Rightarrow$

$$A = \frac{e^2 \cdot a^2 + f^2}{1 + a^2} = 0{,}667 \qquad B = a \cdot \frac{f^2 - e^2}{1 + a^2} = 0$$

$\Rightarrow$

$$Y_{1;2}^{(\varphi)} = y_M + \frac{B}{A} \cdot (x - x_M) \pm \frac{f \cdot e}{A} \cdot \sqrt{A - (x - x_M)^2}$$

⇒

$$Y_{1;2}^{(\varphi)} = \pm\sqrt{0,667 - (x-2)^2}$$

Bei:

$$x_M = \frac{12}{6} = 2 \qquad y_M = \frac{0}{6} = 0$$

Sowie:

$$\tan\varphi = \frac{V_{YY} - V_{XX}}{C} = 0$$

⇒

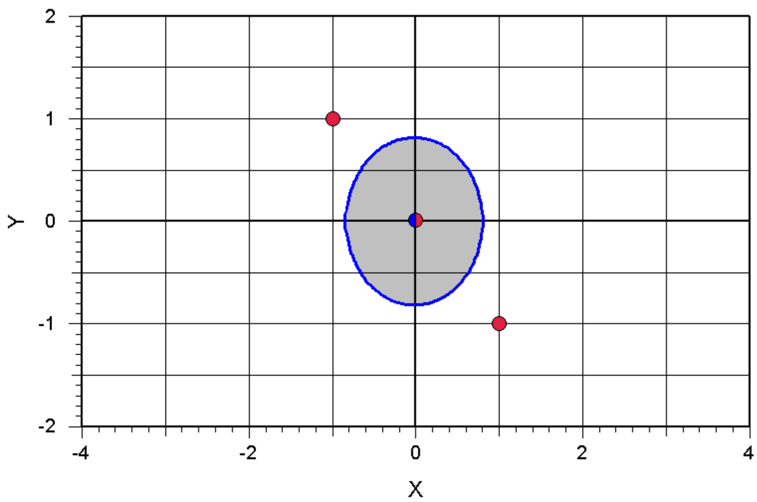
$$\varphi = 0°$$

Damit können die Varianzen nicht aufgelöst werden.

Die Ermittlung der Werte x' und y'.

i	x_i	y_i	x'_i	y'_i	$(x'_i - x'_M)^2$	$(y'_i - y'_M)^2$	$(x'_i - x'_M) \cdot (y'_i - y'_M)$
1	3	-1	+1	-1	1	1	-1
2	2	0	0	0	0	0	0
3	1	+1	-1	+1	1	1	-1
4	1	+1	-1	+1	1	1	-1
5	2	0	0	0	0	0	0
6	3	-1	+1	-1	1	1	-1
$\sum$	12	0	0	0	4	4	-4

⇒



⇒

$$V_{XX} = 4 \qquad C = -4 \qquad V_{YY} = 4$$

Es liegt eine Kreisregression vor.

$$f = \sqrt{\frac{V_{XX}}{n}} = 0,816 \qquad e = \sqrt{\frac{V_{YY}}{n}} = 0,816$$

⇒

$$Y_{1;2}^{(\varphi = x_M = y_M = 0)} = \pm\sqrt{\frac{V_{YY}}{V_{XX}}} \cdot \sqrt{\frac{V_{XX}}{n} - x^2}$$

$\Rightarrow$
$$Y_{1;2}^{(\varphi=x_M=y_M=0)} = \pm\sqrt{0,667 - x^2}$$

Sowie:
$$\tan\varphi = \frac{V_{YY} - V_{XX}}{C} = 0 \qquad \rightarrow \qquad \varphi = 0^\circ$$

LaTeX 2_ε