Die geometrisch echte Ausgleichsgerade

Peter Gallin, Universität Zürich

1 Einleitung

Gegeben sind in der xy-Ebene n Punkte $P_i(x_i/y_i)$ mit den Koordinaten $(x_1, x_2, \ldots, x_n)$ und $(y_1, y_2, \ldots, y_n)$. Im Stochastikunterricht des Gymnasiums[1] berechnet man normalerweise die Gleichung der sogenannten Ausgleichsgeraden, indem man sie durch $y - \overline{y} = m(x - \overline{x})$ ansetzt, wobei die Mittelwerte $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ und $\overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ die Koordinaten des Schwerpunkts $S(\overline{x}/\overline{y})$ aller n Punkte liefern. Durch den plausiblen Trick, die Ausgleichsgerade durch den Schwerpunkt laufen zu lassen, vermeidet man das in der Hochschule übliche Verfahren, bei dem die Ausgleichsgerade durch die Gleichung $y = mx + q$ angesetzt wird und damit zwei Unbekannte m und q zu bestimmen sind, was auf das Differenzieren einer Funktion mit zwei Variablen hinausläuft.

2 Die normale Ausgleichsgerade

Das entscheidende Kriterium für die Festlegung der Ausgleichsgeraden ist das Minimum einer Summe von Abstandsquadraten. Dabei berechnet man üblicherweise die **vertikalen**, also in y-Richtung gemessenen Abstände der n Punkte von der Geraden. Dies liefert die folgende zu minimierende Funktion $f(m)$:

$$f(m) = \sum_{i=1}^{n}\big(y_i - (\overline{y} + m(x_i - \overline{x}))\big)^2 = \sum_{i=1}^{n}\big(y_i - \overline{y} - mx_i + m\overline{x}\big)^2$$

Durch mechanisches Ausmultiplizieren des Quadrats mit vier Summanden erhält man wohl am schnellsten — sicher nicht am elegantesten — eine vereinfachte Form für $f(m)$:

$$\sum_{i=1}^{n} y_i^2 + n\overline{y}^2 + m^2\sum_{i=1}^{n} x_i^2 + nm^2\overline{x}^2 - 2\overline{y}\sum_{i=1}^{n} y_i - 2m\sum_{i=1}^{n} x_i y_i + 2m\overline{x}\sum_{i=1}^{n} y_i + 2m\overline{y}\sum_{i=1}^{n} x_i - 2nm\overline{x}\,\overline{y} - 2m^2\overline{x}\sum_{i=1}^{n} x_i$$

Ersetzen wir hier die Summen $\sum_{i=1}^{n} x_i = n\overline{x}$ und $\sum_{i=1}^{n} y_i = n\overline{y}$, so erhalten wir:

$$f(m) = \sum_{i=1}^{n} y_i^2 - n\overline{y}^2 + m^2\sum_{i=1}^{n} x_i^2 - nm^2\overline{x}^2 - 2m\sum_{i=1}^{n} x_i y_i + 2nm\overline{x}\,\overline{y}$$

Kürzen wir schliesslich noch mittels der Varianzen[2] $\sigma_x^2 = \frac{1}{n}(\sum_{i=1}^{n} x_i^2 - n\overline{x}^2)$ und $\sigma_y^2 = \frac{1}{n}(\sum_{i=1}^{n} y_i^2 - n\overline{y}^2)$, sowie der Kovarianz $c_{xy} = \frac{1}{n}(\sum_{i=1}^{n} x_i y_i - n\overline{x}\,\overline{y})$ ab, so ergibt sich:

$$f(m) = n(\sigma_y^2 + m^2\sigma_x^2 - 2mc_{xy})$$

Über das Nullsetzen der Ableitung ergibt sich derjenige Wert von m, der f minimal macht:

$$f'(m) = n(2m\sigma_x^2 - 2c_{xy}) \stackrel{!}{=} 0 \quad \Longrightarrow \quad m = \frac{c_{xy}}{\sigma_x^2}$$

Dass wir mit diesem Extremum tatsächlich ein Minimum erhalten haben, lässt sich aufgrund der nach oben offenen parabolischen Gestalt des Graphen von $f(m)$ sofort einsehen. So zeigt sich im Nachhinein der Vorteil von diesem nicht-eleganten Bestimmen der vereinfachten Form von $f(m)$, anstatt gleich mit der ersten Form von $f(m)$ ins Ableiten einzusteigen.

[1] Robert Ineichen, Hansjürg Stocker: Stochastik. Einführung in die elementare Statistik und Wahrscheinlichkeitsrechnung. Raeber Verlag Luzern 2007, 11. unveränd. Auflage, ISBN: 3723900429.

[2] Es ist hier unerheblich, ob man durch n oder $n-1$ teilt.

3 Die Ausgleichsgerade der schräg gemessenen Abstandsquadrate

Nun haben meine Schüler aufgrund der geometrischen Fragestellung zurecht beanstandet, dass es sehr willkürlich sei, die vertikalen Abstände zu betrachten. Auch die horizontalen Abstände könnte man ja analog behandeln, was aber an der ungerechten Behandlung der Variablen x und y nichts ändert. Einige schlugen vor, dass man die Summe der horizontalen und vertikalen Abstandsquadrate minimieren solle, was aber — wegen dem Auftreten der Steigung $\frac{1}{m}$ — auf ein Problem 4. Grades in m führte und nicht befriedigte. So versuchten wir die wirklichen, schräg gemessenen Abstände der Punkte von der Ausgleichsgeraden zu betrachten. (Dass dies bei stochastischen Fragestellungen nicht sinnvoll ist, wurde diskutiert.) Damit war eine willkommene Gelegenheit gekommen, die Hessesche Normalform der Geradengleichung zu repetieren oder ein zweites Mal herzuleiten. Es ist nämlich dank Ähnlichkeit von Steigungsdreiecken offensichtlich, dass man den vertikalen Abstand eines Punktes von einer Geraden mit Steigung m nur durch $\sqrt{1+m^2}$ dividieren muss, um den schräg gemessenen Abstand zu erhalten. Damit ergibt sich die neue zu minimierende Funktion $g(m)$ direkt aus der Funktion $f(m)$:

$$g(m) = \frac{1}{\left(\sqrt{1+m^2}\right)^2} \sum_{i=1}^{n} \left(y_i - \overline{y} - mx_i + m\overline{x}\right)^2 = \frac{n}{(1+m^2)}(\sigma_y^2 + m^2\sigma_x^2 - 2mc_{xy})$$

Erst beim Ableiten wird der Unterschied grösser:

$$g'(m) = \frac{n}{(1+m^2)^2}\Big((2m\sigma_x^2 - 2c_{xy})(1+m^2) - (\sigma_y^2 + m^2\sigma_x^2 - 2mc_{xy}) \cdot 2m\Big) \stackrel{!}{=} 0$$

$$\Longrightarrow \quad (m\sigma_x^2 - c_{xy})(1+m^2) = m\sigma_y^2 + m^3\sigma_x^2 - 2m^2c_{xy}$$

In dieser Gleichung fällt die dritte Potenz von m weg und wir erhalten eine quadratische Gleichung in m:

$$m^2c_{xy} + m(\sigma_x^2 - \sigma_y^2) - c_{xy} = 0$$

$$\boxed{m_{1/2} = \frac{\sigma_y^2 - \sigma_x^2 \overset{+}{\scriptstyle(-)} \sqrt{(\sigma_y^2 - \sigma_x^2)^2 + 4c_{xy}^2}}{2c_{xy}} \ .}$$

Da das Vorzeichen von m gleich dem Vorzeichen von c_{xy} sein muss, muss der Zähler positiv sein, und damit gilt nur das Pluszeichen vor der Wurzel. In der Tat rechnet man nach, dass $g''(m_1) > 0$ und $g''(m_2) < 0$. Ausserdem kann man nachweisen, dass beim Berechnen von $\frac{1}{m_1}$ (natürlich nach Erweitern mit dem konjugierten Wurzelterm) in der obigen Fomel genau x und y vertauscht werden, womit also vollständige Gerechtigkeit unter den Variablen x und y nachgewiesen ist. Abschliessend sei angemerkt, dass man unter dem Stichwort „orthogonal regression" einiges zu diesem Thema im Internet findet.

4 Beispiel

Will man durch die Punkte $P_1(1/1)$, $P_2(3/1)$, $P_3(4/3)$ und $P_4(5/4)$ und deren Schwerpunkt $S(\frac{13}{4}/\frac{9}{4})$ die **normale** Ausgleichsgerade g_{nv} mit minimaler vertikaler Abstandsquadratsumme legen, berechnet man

$$\sigma_x^2 = \frac{1}{4}(51 - 4 \cdot (\frac{13}{4})^2) = \frac{35}{16} \quad \text{und} \quad c_{xy} = \frac{1}{4}(36 - 4 \cdot \frac{13}{4} \cdot \frac{9}{4}) = \frac{27}{16} \ .$$

Daraus ergibt sich die Steigung $m_v = \frac{c_{xy}}{\sigma_x^2} = \frac{27}{35} \approx 0.7714$ und die Gleichung der Geraden g_{nv}

$$g_{nv}: \quad y = \frac{27}{35} \cdot (x - \frac{13}{4}) + \frac{9}{4} = \frac{27}{35}x - \frac{9}{35} \ .$$

Will man die horizontale Abstandsquadratsumme minimieren, berechnet man zusätzlich

$$\sigma_y^2 = \frac{1}{4}(27 - 4 \cdot (\frac{9}{4})^2) = \frac{27}{16} \ .$$

Dies liefert die Steigung $m_h = \frac{\sigma_y^2}{c_{xy}} = 1$ und die Gleichung der Geraden g_{nh}:

$$g_{nh}: \quad y = 1 \cdot (x - \frac{13}{4}) + \frac{9}{4} = x - 1$$

Berechnen wir aus $16\sigma_x^2 = 35$, $16\sigma_y^2 = 27$ und $16c_{xy} = 27$ die Steigung m_1, so erhalten wir

$$m_1 = \frac{-8 + \sqrt{64 + 4 \cdot 27^2}}{2 \cdot 27} = \frac{-4 + \sqrt{745}}{27} \approx 0.862766$$

und die Gleichung der „echten“ Ausgleichsgeraden g_e, welche „zwischen“ den anderen beiden liegt:

$$g_e: \quad y = \frac{-4 + \sqrt{745}}{27} \cdot (x - \frac{13}{4}) + \frac{9}{4} \approx 0.86x - 0.55\ .$$