Beweis zu Seite 1393 (4. Auflage)

Partielle und multiple Korrelation sowie Regressionskoeffizienten

In der Vertiefung auf Seite 1393 der 4. Auflage sind in der 1. Spalte unten ohne Beweis zwei nützliche Formeln angegeben, mit denen man aus der Inversen der Korrelationsmatrix die partiellen und multiplen Korrelationen aller beteiligten Variablen ableiten kann.

Wir wollen hier die Beweise nachliefern. Diese benutzen Ergebnisse, die wir zum Teil später im Kapitel 41 über die Regression kennenlernen werden. Dazu stellen wir hier das notwendige Werkzeug geschlossen zusammen, auch wenn dadurch die Darstellung etwas umfangreicher wird. Dabei betrachten wir einen von den Vektoren $\{\mathbf{x}_k, k = 1, \cdots, m\}$ erzeugten Unterraum $\mathbf{M} \subset \mathbb{R}^n$ und gegebenenfalls weitere Unterräume $\mathbf{N} \subset \mathbb{R}^n$ und $\mathbf{Z} \subset \mathbb{R}^n$.

Da wir Korrelationseigenschaften untersuchen wollen, setzen wir ohne Einschränkung der Allgemeinheit voraus, dass alle Vektoren zentriert und normiert sind: $\mathbf{x}^T\mathbf{1} = 0$ und $\|\mathbf{x}\| = 1$. Die hier nur für den $\mathbb{R}^n$ gemachten Aussagen gelten weitaus allgemeiner, zum Beispiel im Raum der zufälligen Variablen mit existierenen Varianz. Die Beweise lassen sich wörtlich verallgemeinern.

1 Projektionen

Ist $\mathbf{M} \subset \mathbb{R}^n$ und $\mathbf{P_M}$ der Projektionsoperator in diesen Unterraum, dann gilt:

$$\begin{aligned} \mathbf{P_M} &= \mathbf{P}_\mathbf{M}^T, \\ \mathbf{P_M} &= \mathbf{P_M}\mathbf{P_M}, \\ \mathbf{P_M}\left(\mathbf{I} - \mathbf{P_M}\right) &= \mathbf{0}. \end{aligned}$$

Dabei ist $\mathbf{I}$ die Einheitsmatrix bzw. die Identität. Wir verwenden der einfacheren Lesbarkeit willen für Operatoren und die sie abbildenden Matrizen die gleichen Buchstaben. Wird $\mathbf{M}$ von den Vektoren $\{\mathbf{x}_k, k = 1, \cdots, m\}$ erzeugt und ist $\mathbf{X} = (\mathbf{x}_1, \cdots, \mathbf{x}_m)$ eine Matrix, deren Spalten die Vektoren $\mathbf{x}_k$ bilden, dann schreiben wir

$$\mathbf{P_M} = \mathbf{P_X}.$$

Für jeden Vektor $\mathbf{x} \in \mathbf{M}$ und jeden Vektor $\mathbf{y} \in \mathbb{R}^n$ ist

$$\begin{aligned} \mathbf{P_M}\mathbf{x} &= \mathbf{x}, \\ \mathbf{P_M}\left(\mathbf{I} - \mathbf{x}\right) &= \mathbf{0}, \\ \mathbf{y}^T\mathbf{P_M}\mathbf{y} &= \left\|\mathbf{P_M}\mathbf{y}\right\|^2 \\ \left\|\mathbf{y} - \mathbf{P_M}\mathbf{y}\right\|^2 &= \left\|\mathbf{y}\right\|^2 - \left\|\mathbf{P_M}\mathbf{y}\right\|^2. \end{aligned}$$

Die Projektion auf einen Vektor $\mathbf{x}$, genauer gesagt, auf die von $\mathbf{x}$ aufgespannte Gerade, ist:

$$\mathbf{P_x}\mathbf{y} = \frac{\mathbf{x}^T\mathbf{y}}{\left\|\mathbf{x}\right\|^2}\mathbf{x}.$$

Sind die Spalten der Matrix $\mathbf{X}$ linear unabhängig, so ist

$$\mathbf{P_X} = \mathbf{X}\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T.$$

Sind $\mathbf{M}$ und $\mathbf{N}$ zwei orthogonale Unterräume und $\{\mathbf{M}, \mathbf{N}\}$ der von ihnen gemeinsam erzeugte Raum, so ist

$$\begin{aligned} \mathbf{P_M}\mathbf{P_N} &= \mathbf{0}, \\ \mathbf{P_M} + \mathbf{P_N} &= \mathbf{P}_{\{\mathbf{M},\mathbf{N}\}}. \end{aligned}$$

2 Orthogonale Zerlegung, Punktschreibweise

Jedes $\mathbf{y} \in \mathbb{R}^n$ wird durch

$$\begin{aligned} \mathbf{y} &= \mathbf{P_M}\mathbf{y} + \left(\mathbf{I} - \mathbf{P_M}\right)\mathbf{y} \\ &= \mathbf{y_M} + \mathbf{y}_{\bullet\mathbf{M}}. \end{aligned}$$

in zwei zueinander orthogonale Komponenten zerlegt. Dabei ist die $\mathbf{M}$-Komponente $\mathbf{y_M} = \mathbf{P_M}\mathbf{y}$ ein Element des Unterraums $\mathbf{M}$. $\mathbf{P_M}\mathbf{y}$ ist die beste lineare Approximation $\mathbf{y}$ durch $\mathbf{M}$:

$$\left\|\mathbf{y} - \mathbf{P_M}\mathbf{y}\right\|^2 \leq \left\|\mathbf{y} - \mathbf{m}\right\|^2 \quad \forall \mathbf{m} \in \mathbf{M}.$$

$\mathbf{y}_{\bullet\mathbf{M}} = (\mathbf{I} - \mathbf{P}_\mathbf{M})\,\mathbf{y}$ ist die vom linearen **M**-Anteil bereinigte Komponente oder das **M**-Residuum von **y**. Es steht orthogonal zu **M**. Man sagt auch, $\mathbf{y}_{\bullet\mathbf{M}}$ sei der durch **M** nicht linear erklärbare Rest. Mitunter lässt man das Wort "*linear*" weg und sagt vereinfachend, man habe den Einfluss von **M** *eliminiert*. Dies kann missverstanden werden und ist wörtlich genommen falsch, denn $\mathbf{y}_{\bullet\mathbf{M}}$ kann durchaus noch von **M** abhängen, nur eben nicht linear. Wird $\mathbf{M} = \{\mathbf{x}\}$ nur von einem Vektor **x** erzeugt, schreiben wir

$$\mathbf{y} = \mathbf{P}_\mathbf{x}\mathbf{y} + (\mathbf{I} - \mathbf{P}_\mathbf{x})\,\mathbf{y}$$

$$= \mathbf{y}_\mathbf{x} + \mathbf{y}_{\bullet\mathbf{x}}.$$

Sind **M** und **N** zwei Unterräume, dann wird jeder Vektor $\mathbf{y} \in \mathbf{N}$ durch

$$\mathbf{y} = \mathbf{y}_\mathbf{M} + \mathbf{y}_{\bullet\mathbf{M}}$$

in eine **M**-Komponente und ein **M**-Residuum $\mathbf{y}_{\bullet\mathbf{M}}$ zerlegt. Der von diesen **M**-Residuen erzeugte Raum ist

$$\mathbf{N}_{\bullet\mathbf{M}} = \{\mathbf{y}_{\bullet\mathbf{M}} : \mathbf{y} \in \mathbf{N}\}\,.$$

Die Räume **M** und $\mathbf{N}_{\bullet\mathbf{M}}$ sind zueinander orthogonal. Sie erzeugen den gleichen Raum wie die Räume **M** und **N** zusammen:

$$\{\mathbf{M}, \mathbf{N}\} = \{\mathbf{M}, \mathbf{N}_{\bullet\mathbf{M}}\} = \{\mathbf{M}_{\bullet\mathbf{N}}, \mathbf{N}\}\,.$$

Daher ist

$$\mathbf{P}_{\{\mathbf{M}\,,\mathbf{N}\}} = \mathbf{P}_\mathbf{M} + \mathbf{P}_{\mathbf{N}_{\bullet\mathbf{M}}} = \mathbf{P}_\mathbf{N} + \mathbf{P}_{\mathbf{M}_{\bullet\mathbf{N}}}.$$

3 Partielle Korrelationen und Kovarianzen

Sind alle Vektoren $\mathbf{y}, \mathbf{x}_1, \cdots, \mathbf{x}_m$ zentriert, $\mathbf{x}^T\mathbf{1} = 0$, dann ist

$$\operatorname{cov}(\mathbf{x}, \mathbf{y}) = \frac{1}{n}\mathbf{x}^T\mathbf{y},$$
$$r(\mathbf{x}, \mathbf{y}) = \frac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{x}\|\,\|\mathbf{y}\|}.$$

Sind die Vektoren darüber hinaus normiert, $\|\mathbf{x}\| = 1$, dann ist

$$r(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T\mathbf{y}.$$

Die Kovarianzmatrix ist

$$\operatorname{cov}(\mathbf{X}) = \frac{1}{n}\mathbf{X}^T\mathbf{X}.$$

Das Quadrat der multiplen Korrelation zwischen einem Vektor **x** und einem Unterraum **M**, bzw. den Vektoren, die diesen Unterraum erzeugen, ist

$$r^2(\mathbf{x}, \mathbf{M}) = \frac{\|\mathbf{P}_\mathbf{M}\mathbf{x}\|^2}{\|\mathbf{x}\|^2}.$$

$r^2(\mathbf{x},\mathbf{M})$ heißt auch das Bestimmtheitsmaß.

Nun betrachten wir zwei Vektoren $\mathbf{x}$ und $\mathbf{y}$, die von einem dritten Vektor $\mathbf{z}$ oder einem Komplex $\mathbf{Z}$ von Vektoren beeinflußt werden. Dabei ist $\mathbf{Z} \subset \mathbb{R}^n$ der von diesen Vektoren erzeugte Raum. Auch hier setzen wir die Zentrierung der Vektoren von $\mathbf{Z}$ oder zumindest $\mathbf{1} \in \mathbf{Z}$ voraus. Wegen $\mathbf{1} \in \mathbf{Z}$ sind auch die $\mathbf{Z}$-Residuen zentriert:

$$\mathbf{x}_{\bullet\mathbf{Z}}^T\mathbf{1} = \left[(\mathbf{I}-\mathbf{P_Z})\,\mathbf{x}\right]^T\mathbf{1} = \mathbf{x}^T(\mathbf{I}-\mathbf{P_Z})\,\mathbf{1} = \mathbf{x^T}(\mathbf{1}-\mathbf{1}) = \mathbf{0}.$$

Die partielle Kovarianz von $\mathbf{x}$ und $\mathbf{y}$ nach Elimination des linearen Einflußes von $\mathbf{Z}$ ist

$$\begin{aligned}\operatorname{cov}(\mathbf{x}_{\bullet\mathbf{Z}},\mathbf{y}_{\bullet\mathbf{Z}}) &= \frac{1}{n}\mathbf{x}_{\bullet\mathbf{Z}}^T\mathbf{y}_{\bullet\mathbf{Z}}\\ &= \frac{1}{n}\mathbf{x}^T(\mathbf{I}-\mathbf{P_Z})\,\mathbf{y}.\end{aligned}$$

Daher ist

$$\operatorname{cov}(\mathbf{x}_{\bullet\mathbf{Z}},\mathbf{y}) = \operatorname{cov}(\mathbf{x},\mathbf{y}_{\bullet\mathbf{Z}}) = \operatorname{cov}(\mathbf{x}_{\bullet\mathbf{Z}},\mathbf{y}_{\bullet\mathbf{Z}}).$$

Wir können daher auch kurz $\operatorname{cov}(\mathbf{x},\mathbf{y})_{\bullet\mathbf{Z}}$ schreiben. Analog schreiben wir für die partielle Korrelation von $\mathbf{x}$ und $\mathbf{y}$ nach Elimination des linearen Einflußes von $\mathbf{Z}$:

$$r(\mathbf{x},\mathbf{y})_{\bullet\mathbf{Z}} = r(\mathbf{x}_{\bullet\mathbf{Z}},\mathbf{y}_{\bullet\mathbf{Z}}) = \frac{\mathbf{x}^T(\mathbf{I}-\mathbf{P_Z})\,\mathbf{y}}{\|\mathbf{x}_{\bullet\mathbf{Z}}\|\,\|\mathbf{y}_{\bullet\mathbf{Z}}\|}.$$

4 Berechnung der partiellen Korrelation

Theorem 1 *Für drei Vektoren* $\mathbf{x}$, $\mathbf{y}$ *und* $\mathbf{z}$ *gilt:*

$$r(\mathbf{x},\mathbf{y})_{\bullet\mathbf{z}} = \frac{r(\mathbf{x},\mathbf{y}) - r(\mathbf{x},\mathbf{z})r(\mathbf{z},\mathbf{y})}{\sqrt{(1-r^2(\mathbf{x},\mathbf{z}))(1-r^2(\mathbf{y},\mathbf{z}))}}. \tag{1}$$

Beweis: Es ist

$$r(\mathbf{x},\mathbf{y})_{\bullet\mathbf{z}} = \frac{\mathbf{y}^T(\mathbf{I}-\mathbf{P_z})\,\mathbf{x}}{\|\mathbf{x}_{\bullet\mathbf{z}}\|\,\|\mathbf{y}_{\bullet\mathbf{z}}\|} = \frac{\mathbf{y}^T\mathbf{x} - \mathbf{y}^T\mathbf{P_z}\mathbf{x}}{\|\mathbf{x}_{\bullet\mathbf{z}}\|\,\|\mathbf{y}_{\bullet\mathbf{z}}\|}.$$

Ohne Einschränkung der Allgemeinheit können wir voraussetzen, dass alle Vektoren normiert sind, $\|\mathbf{z}\| = 1$, dann ist

$$\mathbf{P_z x} = \frac{\mathbf{z}^T\mathbf{x}}{\|\mathbf{z}\|^2}\mathbf{z} = r(\mathbf{z},\mathbf{x})\mathbf{z}.$$

Daher

$$\begin{aligned}\mathbf{y}^T(\mathbf{I}-\mathbf{P_z})\,\mathbf{x} = \mathbf{y}^T\mathbf{x} - \mathbf{y}^T r(\mathbf{z},\mathbf{x})\mathbf{z}\\ = r(\mathbf{x},\mathbf{y}) - r(\mathbf{y},\mathbf{z})r(\mathbf{z},\mathbf{x}).\end{aligned}$$

Ersetzt man $\mathbf{y}$ durch $\mathbf{x}$, folgt:

$$\|x_{\bullet \mathbf{z}}\|^2 = 1 - r^2(\mathbf{z}, \mathbf{x}).$$

■

Hat man es allgemeiner mit einem Matrix $\mathbf{Z}$ zu tun, ist die Formel ganz analog, nur ist im Beweis $\mathbf{P_Z} = \mathbf{Z}\left(\mathbf{Z}^T\mathbf{Z}\right)^{-1}\mathbf{Z}$ zu beachten.

5 Schrittweise Elimination

In der Praxis werden Einflüsse oft erst nach und nach erkannt. Zuerst entdeckt man den Komplex $\mathbf{M}$ und *schaltet ihn aus:*

$$\mathbf{y} \to \mathbf{y}_{\bullet \mathbf{M}}.$$

Dann entdeckt man einen weiteren Einflußkomplex $\mathbf{N}$ und *eliminiert* ihn ebenfalls:

$$\mathbf{y}_{\bullet \mathbf{M}} \to (\mathbf{y}_{\bullet \mathbf{M}})_{\bullet \mathbf{N}}.$$

Die umgekehrte Reihenfolge liefert

$$\mathbf{y} \to \mathbf{y}_{\bullet \mathbf{N}} \to (\mathbf{y}_{\bullet \mathbf{N}})_{\bullet \mathbf{M}}.$$

Schließlich wäre $\mathbf{y}_{\bullet(\mathbf{M},\mathbf{N})}$ das Ergebnis, wenn beide Variablenkomplexe gemeinsam auf einmal *eliminiert* worden wären. In der Regel gilt leider

$$(\mathbf{y}_{\bullet \mathbf{M}})_{\bullet \mathbf{N}} \neq (\mathbf{y}_{\bullet \mathbf{N}})_{\bullet \mathbf{M}} \neq \mathbf{y}_{\bullet\{\mathbf{M},\mathbf{N}\}},$$

es sei denn, die Räume $\mathbf{M}$ und $\mathbf{N}$ sind orthogonal.

Theorem 2 *Sind* $\mathbf{M}$ *und* $\mathbf{N}$ *zwei orthogonale Räume, so ist die iterative Bestimmung der Residuen unabhängig von der Reihenfolge:*

$$(\mathbf{y}_{\bullet \mathbf{N}})_{\bullet \mathbf{M}} = (\mathbf{y}_{\bullet \mathbf{M}})_{\bullet \mathbf{N}} = \mathbf{y}_{\bullet\{\mathbf{M},\mathbf{N}\}}.$$

Daher gilt für die partiellen Korrelationen

$$r(\mathbf{x}_{\bullet \mathbf{N}}, \mathbf{y}_{\bullet \mathbf{N}})_{\bullet \mathbf{M}} = r(\mathbf{x}_{\bullet \mathbf{M}}, \mathbf{y}_{\bullet \mathbf{M}})_{\bullet \mathbf{N}} = r(\mathbf{x}, \mathbf{y})_{\bullet\{\mathbf{M},\mathbf{N}\}}.$$

Beweis:

$$\begin{aligned}(\mathbf{y}_{\bullet \mathbf{M}})_{\bullet \mathbf{N}} &= (\mathbf{I} - \mathbf{P_N})(\mathbf{I} - \mathbf{P_M})\mathbf{y} = (\mathbf{I} - \mathbf{P_M} - \mathbf{P_N} + \mathbf{P_M}\mathbf{P_N})\mathbf{y} \\ &= (\mathbf{I} - \mathbf{P_M} - \mathbf{P_N})\mathbf{y} = (\mathbf{y}_{\bullet \mathbf{N}})_{\bullet \mathbf{M}}. \\ &= (\mathbf{I} - (\mathbf{P_M} + \mathbf{P_N}))\mathbf{y} = \mathbf{y}_{\bullet\{\mathbf{M},\mathbf{N}\}}.\end{aligned}$$

■

Sind nun $\mathbf{M}$ und $\mathbf{N}$ nicht orthogonal, orthogonalisiert man $\mathbf{N}$ bezüglich $\mathbf{M}$ und erhält $\mathbf{N}_{\bullet \mathbf{M}}$. Dann sind die beiden Räume $\mathbf{M}$ und $\mathbf{N}_{\bullet \mathbf{M}}$ orthogonal. Die Bereinigung kann nach Satz 2 in beliebiger Reihenfolge vorgenommen worden.

Theorem 3 *Sind* $\mathbf{x}$ *und* $\mathbf{y}$ *zwei Vektoren und* $\mathbf{M}$ *und* $\mathbf{N}$ *zwei beliebige Räume, so ist:*

$$\begin{array}{lclcl}(\mathbf{x}_{\bullet \mathbf{M}})_{\bullet \mathbf{N}_{\bullet \mathbf{M}}} & = & \mathbf{x}_{\bullet(\mathbf{M},\mathbf{N})} & = & (\mathbf{x}_{\bullet \mathbf{N}})_{\bullet \mathbf{M}_{\bullet \mathbf{N}}}, \\ r(\mathbf{x}_{\bullet \mathbf{M}}, \mathbf{y}_{\bullet \mathbf{M}})_{\bullet \mathbf{N}_{\bullet \mathbf{M}}} & = & r(\mathbf{x}, \mathbf{y})_{\bullet(\mathbf{M},\mathbf{N})} & = & r(\mathbf{x}_{\bullet \mathbf{N}}, \mathbf{y}_{\bullet \mathbf{N}})_{\bullet \mathbf{M}_{\bullet \mathbf{N}}}.\end{array}$$

6 Reziproke Partialisierung

Bereinigt man $\mathbf{x}$ vom linearen Einfluss $\mathbf{y}$, so sind die resultierenden Vektoren $\mathbf{x}_{\bullet\mathbf{y}}$ und $\mathbf{y}$ orthogonal. Ebenso sind $\mathbf{x}$ und $\mathbf{y}_{\bullet\mathbf{x}}$ orthogonal. Was geschieht, wenn man gleichzeitig $\mathbf{x}$ und $\mathbf{y}$ jeweils vom linearen Einfluss des anderen Partners bereinigt? Die Antwort ist überraschend: Die Korrelation ist bis auf das Vorzeichen dieselbe geblieben. Es gilt:

Theorem 4 *Für zwei Vektoren* $\mathbf{x}$ *und* $\mathbf{y}$ *ist:*

$$r(\mathbf{x}_{\bullet\mathbf{y}}, \mathbf{y}_{\bullet\mathbf{x}}) =- r(\mathbf{x}, \mathbf{y}).$$

Beweis: Es seien $\mathbf{x}$ und $\mathbf{y}$ zentriert und normiert. Dann ist $\mathbf{x}^T\mathbf{x} = 1$ und $\mathbf{x}^T\mathbf{y} =$

$r(\mathbf{x}, \mathbf{y}) = r$. Daher ist $\mathbf{P_y x} = \frac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{y}\|^2}\mathbf{y} = r\mathbf{y}$ und $\mathbf{x}_{\bullet\mathbf{y}} = \mathbf{x} - r\mathbf{y}$ sowie $\|\mathbf{x}_{\bullet\mathbf{y}}\|^2 = \|\mathbf{x} - \mathbf{P_y x}\|^2 = \|\mathbf{x}\|^2 - \|\mathbf{P_y x}\|^2 = 1 - r^2$. Und analog für $\mathbf{y}$. Daraus folgt

$$\mathbf{x}_{\bullet\mathbf{y}}^T\mathbf{y}_{\bullet\mathbf{x}} = (\mathbf{x} - r\mathbf{y})^T(\mathbf{y} - r\mathbf{x}) = -r\left(1 - r^2\right).$$

Nun ist

$$r(\mathbf{x}_{\bullet\mathbf{y}}, \mathbf{y}_{\bullet\mathbf{x}}) = \frac{\mathbf{x}_{\bullet\mathbf{y}}^T\mathbf{y}_{\bullet\mathbf{x}}}{\|\mathbf{x}_{\bullet\mathbf{y}}\|\,\|\mathbf{y}_{\bullet\mathbf{x}}\|} = \frac{-r\left(1 - r^2\right)}{1 - r^2} = -r.$$

■

Den gleichen Gedanken können wir nun auf eine ganze Gruppe von $k + 2$ Variablen $\mathbf{x}$, $\mathbf{y}$ und $\mathbf{Z} = (\mathbf{z}_1, \ldots, \mathbf{z}_k)$ erweitern. Betrachten wir $\mathbf{x}$ und $\mathbf{y}$ zusammen, so sind die $\mathbf{z}_i$ die **anderen** Variablen. Vom Standpunkt von $\mathbf{x}$ aus sind $(\mathbf{y}, \mathbf{Z})$ **alle anderen**. Analog sind vom Standpunkt $\mathbf{y}$ aus die Variablen $(\mathbf{x}, \mathbf{Z})$ **alle anderen**.

Nun gilt: Die Korrelation der von den **Anderen** bereinigten Variablen $\mathbf{x}$ und $\mathbf{y}$ ist bis auf das Vorzeichen dasselbe wie die Korrelation der von **alle anderen** bereinigten Variablen $\mathbf{x}$ und $\mathbf{y}$. Formal lautet der Satz:

Theorem 5

$$r(\mathbf{x}_{\bullet(\mathbf{Z},\mathbf{y})}, \mathbf{y}_{\bullet(\mathbf{Z},\mathbf{x})}) =- r(\mathbf{x}, \mathbf{y})_{\bullet\mathbf{Z}}.$$

Beweis: Wir zerlegen $\{\mathbf{Z}, \mathbf{y}\} = \{\mathbf{Z}, \mathbf{y}_{\bullet\mathbf{Z}}\}$. Da $\mathbf{Z}$ und $\mathbf{y}_{\bullet\mathbf{Z}}$ orthogonal sind, ist $\mathbf{x}_{\bullet(\mathbf{Z},\mathbf{y})} = (\mathbf{x}_{\bullet}\mathbf{z})_{\bullet\mathbf{y}_{\bullet}\mathbf{z}}$ und $\mathbf{y}_{\bullet(\mathbf{Z},\mathbf{x})} = (\mathbf{y}_{\bullet}\mathbf{z})_{\bullet\mathbf{x}_{\bullet}\mathbf{z}}$. Damit folgt nach Satz 4

$$r(\mathbf{x}_{\bullet(\mathbf{Z},\mathbf{y})}, \mathbf{y}_{\bullet(\mathbf{Z},\mathbf{x})}) = r((\mathbf{x}_{\bullet}\mathbf{z})_{\bullet\mathbf{y}_{\bullet}\mathbf{z}}, (\mathbf{y}_{\bullet}\mathbf{z})_{\bullet\mathbf{x}_{\bullet}\mathbf{z}}) =- r(\mathbf{x}_{\bullet\mathbf{Z}}, \mathbf{y}_{\bullet\mathbf{Z}}) =- r(\mathbf{x}, \mathbf{y})_{\bullet\mathbf{Z}}.$$

■

7 Die Regressionskoeffizienten

Es sei $\mathbf{M}$ ein Unterraum eines linearen Vektorraums. $\mathbf{M}$ sei von den linear unabhängigen Basisvektoren $\mathbf{x}_1, \cdots, \mathbf{x}_m$ erzeugt, die wir als Spalten der Matrix

$\mathbf{X} = (\mathbf{x}_1, \cdots, \mathbf{x}_m)$ zusammenfassen. Jeder Vektor $\mathbf{y}$ des Oberraums besitzt eine eindeutige Projektion $\mathbf{P_M y}$ nach $\mathbf{M}$.

$$\mathbf{P_M y} = \mathbf{X}\beta = \sum_{k=1}^{m} \mathbf{x}_k \beta_k.$$

Wir werden nun die $\mathbf{x}_k$ als Regressoren und die β_k, die wir zum Vektor β zusammenfassen, als Regressionskoeffizienten bezeichnen. β ist eindeutig bestimmt durch

$$\beta = \left(\mathbf{X}^T\mathbf{X}\right)^{-1} \mathbf{X}^T\mathbf{y}. \tag{2}$$

Dies folgt aus der eindeutigen Zerlegung

$$\mathbf{y} = \mathbf{y_M} + \mathbf{y_{\bullet M}} = \mathbf{X}\beta + \mathbf{y_{\bullet M}}$$

Da $\mathbf{y_{\bullet M}}$ orthogonal zu $\mathbf{M} = \{\mathbf{X}\}$ steht, ist $\mathbf{X}^T\mathbf{y_{\bullet M}} = \mathbf{0}$.Multiplikation mit $\mathbf{X}^T$ liefert daher

$$\mathbf{X}^T\mathbf{y} = \mathbf{X}^T\mathbf{X}\beta.$$

Da die Spalten von $\mathbf{X}$ nach Voraussetzung linear unabhängig sind, hat $\mathbf{X}^T\mathbf{X}$ vollen Rang und ist invertierbar. Damit folgt $\beta = \left(\mathbf{X}^T\mathbf{X}\right)^{-1} \mathbf{X}^T\mathbf{y}$.

Wollen wir aber die Rolle kennenlernen, die ein einzelner Regressor $\mathbf{x}_i$ und sein Koeffizienten β_i bei der Darstellung von $\mathbf{P_M y}$ spielen, so ist diese Darstellung nicht hilfreich. Besser ist hier ein anderer Weg. Dazu müssen wir zuerst den Einfluß der anderen Regressoren ausschalten. Wir verwenden dazu den Unterraum

$$\mathbf{M}_{\setminus i} = \{\mathbf{x}_k, k \neq i\} \subset \mathbf{M},$$

der von allen Regressoren mit Ausnahme von $\mathbf{x}_i$ aufgespannt wird. Wir projizieren $\mathbf{y} = \mathbf{y_M} + \mathbf{y_{\bullet M}}$ in diesen Teilraum

$$\begin{aligned}
\mathbf{P}_{\mathbf{M}_{\setminus i}}\mathbf{y} &= \mathbf{P}_{\mathbf{M}_{\setminus i}}\mathbf{y_M} + \mathbf{P}_{\mathbf{M}_{\setminus i}}\mathbf{y_{\bullet M}} \\
&= \mathbf{P}_{\mathbf{M}_{\setminus i}} \sum_{k \neq i}^{m} \mathbf{x}_k \beta_k \\
&= \sum_{k \neq i}^{m} \mathbf{x}_k \beta_k + \mathbf{P}_{\mathbf{M}_{\setminus i}}\mathbf{x}_i \beta_i.
\end{aligned}$$

Die zweite Gleichung gilt, da $\mathbf{y_{\bullet M}}$ orthogonal zu $\mathbf{M}$ und daher erst recht zum Unterraum $\mathbf{M}_{\setminus i}$ steht, die letzte folgt aus $\mathbf{x}_k, \in \mathbf{M}_{\setminus i}$ für alle $k \neq i$. Subtrahieren wir die Gleichungen

$$\begin{aligned}
\mathbf{y} &= \sum_{k=1}^{m} \mathbf{x}_k \beta_k + \mathbf{y_{\bullet M}}, \\
\mathbf{P}_{\mathbf{M}_{\setminus i}}\mathbf{y} &= \sum_{k \neq i}^{m} \mathbf{x}_k \beta_k + \mathbf{P}_{\mathbf{M}_{\setminus i}}\mathbf{x}_i \beta_i,
\end{aligned}$$

erhalten wir

$$\left(\mathbf{I}-\mathbf{P}_{\mathbf{M}\setminus i}\right)\mathbf{y}=\left(\mathbf{I}-\mathbf{P}_{\mathbf{M}\setminus i}\right)\mathbf{x}_i\beta_i+\mathbf{y}_{\bullet\mathbf{M}}$$

oder kurz

$$\mathbf{y}_{\bullet\mathbf{M}\setminus i}=\mathbf{x}_{i\bullet\mathbf{M}\setminus i}\beta_i+\mathbf{y}_{\bullet\mathbf{M}}.$$

Dies ist die Gleichung einer linearen Einfachregression mit dem Regressant $\mathbf{y}_{\bullet\mathbf{M}\setminus i}$ und dem einzigen Regressor $\mathbf{x}_{i\bullet\mathbf{M}\setminus i}$. Sowohl bei $\mathbf{y}$ wie bei $\mathbf{x}_i$ ist nun der lineare Einfluss der anderen Regressoren eliminiert worden. Dagegen ist die unerklärte Restgröße $\mathbf{y}_{\bullet\mathbf{M}}$ unverändert geblieben. Daher ist

$$\beta_i=\frac{\mathbf{x}_i^T\mathbf{y}()_{\bullet\mathbf{M}\setminus i}}{\left\|\mathbf{x}_{i\bullet\mathbf{M}\setminus i}\right\|^2}. \tag{3}$$

8 Die inverse Korrelationsmatrix als Quelle aller partieller und multipler Korrelationskoeffizienten

Wir haben mit den Gleichungen 2 und 3 zwei Darstellungen für den Regressionskoeffizienten β_i gefunden, direkt oder als die i.Zeile des Vektors β :

$$\beta_i=\frac{\left(\mathbf{x}_i^T\mathbf{y}\right)_{\bullet\mathbf{M}\setminus i}}{\left\|\mathbf{x}_{i\bullet\mathbf{M}\setminus i}\right\|^2}=\mathbf{1}_i^T\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{y}$$

Dabei ist $\mathbf{1}_i^T=(0,\cdots,0,1,0,\cdots,0)$ Vektor, der nur an der i-ten Stelle ein 1 besitzt und sonst 0 ist. Da diese Gleichung für alle $\mathbf{y}$ gilt , folgt

$$\frac{\mathbf{x}_i^T{}_{\bullet\mathbf{M}\setminus i}}{\left\|\mathbf{x}_{i\bullet\mathbf{M}\setminus i}\right\|^2}=\mathbf{1}_i^T\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T. \tag{4}$$

Wir bilden die quadrierte Norm dieses Vektors (wir multiplizieren die Gleichung mit ihrer Transponierten) und erhalten

$$\frac{1}{\left\|\mathbf{x}_{i\bullet\mathbf{M}\setminus i}\right\|^2}=\mathbf{1}_i^T\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{X}\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{1}_i \tag{5}$$

$$=\mathbf{1}_i^T\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{1}_i. \tag{6}$$

Damit erhalten wir bemerkenswerte und nützliche Ergebnisse. Im Korrelations- wie vor allem im Regressionsmodell müssen wir stets die Matrizen

$$\mathbf{X},\quad \mathbf{X}^T\mathbf{X},\quad \left(\mathbf{X}^T\mathbf{X}\right)^{-1},\quad \mathbf{X}^T\left(\mathbf{X}^T\mathbf{X}\right)^{-1},\quad \mathbf{X}^T\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}$$

berechnen. Sind die Vektoren $\mathbf{x}_1,\cdots,\mathbf{x}_m$ zentriert und normiert, dann ist $\mathbf{X}^T\mathbf{X}$ die Korrelationsmatrix und $\left(\mathbf{X}^T\mathbf{X}\right)^{-1}=\mathbf{V}$. Der Term auf der rechten Seite von Gleichung 5 ist V_{ii}, das i-te Diagonalelement. Der Term auf der linken Seite von

Gleichung 5 ist

$$\begin{aligned}V_{ii}^{-1} &= \left\|\mathbf{x}_{i\bullet\mathbf{M}_{\setminus i}}\right\|^2 \\ &= \left\|\left(\mathbf{I}-\mathbf{P}_{\mathbf{M}_{\setminus i}}\right)\mathbf{x}\right\|^2 \\ &= \|\mathbf{x}\|^2 - \left\|\mathbf{P}_{\mathbf{M}_{\setminus i}}\mathbf{x}\right\|^2 \\ &= 1 - r^2\left(\mathbf{x}_i, \mathbf{M}_{\setminus i}\right).\end{aligned} \tag{7}$$

$r\left(\mathbf{x}_i, \mathbf{M}_{\setminus i}\right)$ ist die multiple Korrelation zwischen dem Vektor $\mathbf{x}_i$ und allen anderen Vektoren. Je größer diese Korrelation ist, um so stärker ist $\mathbf{x}_i$ bereits durch die anderen Vektoren bestimmt und um so weniger bringt $\mathbf{x}_i$ neue Information ins Modell. Mit Gleichung 7 schreiben wir Gleichung 4 um:

$$\mathbf{x}_{i\bullet\mathbf{M}_{\setminus i}} = \frac{1}{V_{ii}}\mathbf{X}\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{1}_i.$$

Die i-te Spalte der Matrix $\mathbf{X}\left(\mathbf{X}^T\mathbf{X}\right)^{-1}$ liefert – bis auf den Faktor V_{ii}^{-1} – den vom linearen Einfluss aller anderen Vektoren bereinigten Vektor $\mathbf{x}_i$.

Nun bestimmen wir die Korrelation von $\mathbf{x}_{i\bullet\mathbf{M}_{\setminus i}}$ und $\mathbf{x}_{j\bullet\mathbf{M}_{\setminus j}}$.

$$\begin{aligned}r(\mathbf{x}_{i\bullet\mathbf{M}_{\setminus i}}, \mathbf{x}_{j\bullet\mathbf{M}_{\setminus j}}) &= \frac{\mathbf{x}_{i\bullet\mathbf{M}_{\setminus i}}^T}{\left\|\mathbf{x}_{i\bullet\mathbf{M}_{\setminus i}}\right\|}\frac{\mathbf{x}_{j\bullet\mathbf{M}_{\setminus j}}}{\left\|\mathbf{x}_{j\bullet\mathbf{M}_{\setminus j}}\right\|} \\ &= \frac{\mathbf{1}_i^T\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{X}\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{1}_j}{\sqrt{V_{ii}}\sqrt{V_{jj}}} \\ &= \frac{V_{ij}}{\sqrt{V_{ii}V_{jj}}}.\end{aligned}$$

Andererseits können wir diese Korrelation mit Satz 5 berechnen. Dazu bezeichnen wir den von allen Vektoren außer $\mathbf{x}_i$ und $\mathbf{x}_j$ erzeugte Raum mit $\mathbf{Z}$. Dann ist $\mathbf{M}_{\setminus i} = \{\mathbf{Z}, \mathbf{x}_j\}$: sowie $\mathbf{M}_{\setminus j} = \{\mathbf{Z}, \mathbf{x}_i\}$ Nach Satz 5 gilt nun

$$\begin{aligned}r(\mathbf{x}_{i\bullet\mathbf{M}_{\setminus i}}, \mathbf{x}_{j\bullet\mathbf{M}_{\setminus j}}) &= r(\mathbf{x}_{i\bullet\{\mathbf{Z},\mathbf{x}_j\}}, \mathbf{x}_{j\bullet\{\mathbf{Z},\mathbf{x}_i\}}) \\ &= -r(\mathbf{x}_i, \mathbf{x}_j)_{\bullet\mathbf{Z}}.\end{aligned}$$

Mit den Bezeichnungen aus der Vertiefung auf Seite 1393 der 4.Auflage haben wir nun den Satz bewiesen

Theorem 6 *Ist* $\mathbf{V}$ *die Inverse der Korrelationsmatrix, dann ist die partielle Korrelation zwischen* $\mathbf{x}_i$ *und* $\mathbf{x}_j$ *nach Elimination des linearen Einflusses der anderen Vektoren*

$$r(\mathbf{x}_i, \mathbf{x}_j)_{\bullet\{\mathbf{x}_k, k\neq i,j\}} = -\frac{V_{ij}}{\sqrt{V_{ii}V_{jj}}}.$$

Die multiple Korrelation zwischen $\mathbf{x}_i$ *und allen anderen Vektoren ist*

$$r(\mathbf{x}_i, \{\mathbf{x}_k, k \neq i,\}) = 1 - \frac{1}{V_{ii}}.$$

http://www.springer.com/978-3-662-56740-1

Mathematik
Arens, T.; Hettlich, F.; Karpfinger, C.; Kockelkorn, U.;
Lichtenegger, K.; Stachel, H.
2018, XXVIII, 1660 S. 836 Abb., 808 Abb. in Farbe.,
Hardcover
ISBN: 978-3-662-56740-1