

Übersicht zur geometrischen Interpretation des KQ-Schätzers

Falls $\boldsymbol{X}$ vollen Spaltenrang hat (und somit $(\boldsymbol{X}^T\boldsymbol{X})^{-1}$ existiert), kann man den KQ-Schätzer schreiben als:

$$\hat{\beta} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{y}$$

Damit ergeben sich die prognostizierten Werte und die Residuen:

$$\begin{aligned}\hat{\boldsymbol{y}} &= \boldsymbol{X}\hat{\beta} = \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{y}\\ \hat{\boldsymbol{u}} &= \boldsymbol{y} - \hat{\boldsymbol{y}} = \boldsymbol{y} - \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{y} = (\boldsymbol{I}_n - \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T)\boldsymbol{y} = (\boldsymbol{I}_n - \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T)u\end{aligned}$$

Man sieht, dass die prognostizierten Werte und Residuen linear von $\boldsymbol{y}$ abhängen und führt deshalb zwei Matrizen ein:

$$\begin{aligned}\boldsymbol{P}_X &:= \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T \quad \text{(damit gilt: } \hat{\boldsymbol{y}} = \boldsymbol{P}_X\boldsymbol{y})\\ \boldsymbol{M}_X &:= (\boldsymbol{I}_n - \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T) \quad \text{(damit gilt: } \hat{\boldsymbol{u}} = \boldsymbol{M}_X\boldsymbol{y})\end{aligned}$$

Geometrische Eigenschaften einer KQ-Schätzung mit Hilfe der Matrizen $\boldsymbol{P}_X$ und $\boldsymbol{M}_X$:

Eigenschaft / Formel	Interpretation
Symmetrie: $\boldsymbol{P}_X$ und $\boldsymbol{M}_X$ sind symmetrisch (Eine Matrix $\boldsymbol{A}$ heißt symmetrisch, falls gilt: $\boldsymbol{A}^T = \boldsymbol{A}$)	Aus der Symmetrie einer Matrix folgt, dass deren Zeilenraum gleich dem Spaltenraum ist.
Idempotenz: $\boldsymbol{P}_X$ und $\boldsymbol{M}_X$ sind idempotent (Eine Matrix $\boldsymbol{A}$ heißt idempotent, falls gilt: $\boldsymbol{A}\cdot\boldsymbol{A} = \boldsymbol{A}$)	Projektionen werden durch idempotente Matrizen charakterisiert, da eine doppelte, dreifache, ... Projektion auf den gleichen Raum das Gleiche wie eine einfache Projektion liefert. $\hat{\boldsymbol{y}}$ und $\hat{\boldsymbol{u}}$ werden also auf Räume projiziert!
Orthogonalität und Komplementarität: $\boldsymbol{P}_X, \boldsymbol{M}_X \in \mathbb{R}^{n\times n}$ mit $\boldsymbol{P}_X\cdot\boldsymbol{M}_X = \boldsymbol{M}_X\cdot\boldsymbol{P}_X = \boldsymbol{O}$ (wobei: $\mathrm{rk}(\boldsymbol{P}_X) = \mathrm{tr}(\boldsymbol{P}_X) = k$ und $\mathrm{rk}(\boldsymbol{M}_X) = \mathrm{tr}(\boldsymbol{M}_X) = n-k$) Beziehung von $\hat{\boldsymbol{y}}$ zu den Residuen: $\hat{\boldsymbol{y}}^T\hat{\boldsymbol{u}} = (\boldsymbol{P}_X\boldsymbol{y})^T(\boldsymbol{I}-\boldsymbol{P}_X)\boldsymbol{y} = 0$ Insbesondere gilt: $\boldsymbol{P}_X\boldsymbol{y} + \boldsymbol{M}_X\boldsymbol{y} = \boldsymbol{y}$ (bzw. $\hat{\boldsymbol{u}} + \hat{\boldsymbol{y}} = \boldsymbol{y}$)	Die Spalten von $\boldsymbol{P}_X$ stehen senkrecht auf den Spalten $\boldsymbol{M}_X$, d. h. sie bilden Vektoren in Räume ab, die wiederum senkrecht zueinander stehen. Da die Summe der Ränge wieder den ganzen Raum ergibt, sind die beiden Räume komplementär. Da der Vektor der prognostizierten Werte $\hat{\boldsymbol{y}}$ senkrecht auf dem Residuenvektor steht, lässt sich der Regressand als Summe der Residuen und der prognostizierten Werte schreiben.
Normalgleichungen: Beziehung von $\boldsymbol{X}$ zu den Residuen $\boldsymbol{x}_i^T\hat{\boldsymbol{u}} = <\boldsymbol{x}_i, \hat{\boldsymbol{u}}> = 0 \quad \forall i = 1, ...k \iff \hat{\boldsymbol{u}} \in \delta^{\perp}(\boldsymbol{X})$	Der Residuenvektor steht senkrecht auf den Spalten der Matrix $\boldsymbol{X}$. (Der Raum, der durch $\boldsymbol{X}$ aufgespannt wird, steht senkrecht zum Vektor der Residuen $\hat{\boldsymbol{u}}$)
Eigenschaften eines Modells mit Konstante (Intercept) 1. $\sum_{i=1}^n \hat{u}_i = 0$ (insb. der Mittelwert $\frac{1}{n}(\sum_{i=1}^n \hat{u}_i) = 0$) 2. $\bar{\hat{y}} = \frac{1}{n}(\sum_{i=1}^n \hat{y}_i) = \bar{y}$ 3. $\bar{y} = \hat{\beta}_1 + \hat{\beta}_2\bar{x}_2 + ... + \hat{\beta}_k\bar{x}_k$ 4. Streuungszerlegung: $SST = SSR + SSE$	Interpretation des Modells mit Kontante (Intercept): 1. Der Mittelwert der Residuen ist null, sie annulieren sich also im Mittel. 2. Der Mittelwert der Prognosen $\bar{\hat{y}}$ ist gleich dem Mittelwert der beobachteten Werte $\bar{y}$ 3. Die Regressionshyperebene geht durch den Schwerpunkt der Daten 4. $SST = \sum_{i=1}^n(y_i-\bar{y})^2, SSE = \sum_{i=1}^n(\hat{y}_i-\bar{y})^2, SSR = \sum_{i=1}^n \hat{u}_i^2$

Im Folgenden eine illustrierende Graphik im Fall eines Modells mit zwei Regressoren ($k = 2$) und drei Beobachtungen ($n = 3$) und dessen KQ-Schätzung:

$$\mathbf{y} = \begin{pmatrix}\mathbf{x}_1 & \mathbf{x}_2\end{pmatrix}\beta + \mathbf{u} \text{ wobei } \mathbf{y} = \begin{pmatrix}y_1\\ y_2\\ y_3\end{pmatrix} \text{ und } \begin{pmatrix}\mathbf{x}_1 & \mathbf{x}_2\end{pmatrix} = \begin{pmatrix}x_{11} & x_{12}\\ x_{21} & x_{22}\\ x_{31} & x_{32}\end{pmatrix} \text{ mit Schätzung } \hat{\mathbf{y}} = \begin{pmatrix}\mathbf{x}_1 & \mathbf{x}_2\end{pmatrix}\hat{\beta} + \hat{\mathbf{u}}$$

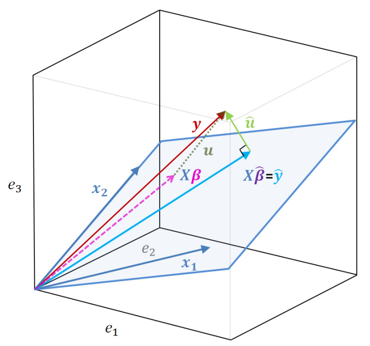

Abbildung 1: Raum der Beobachtungen

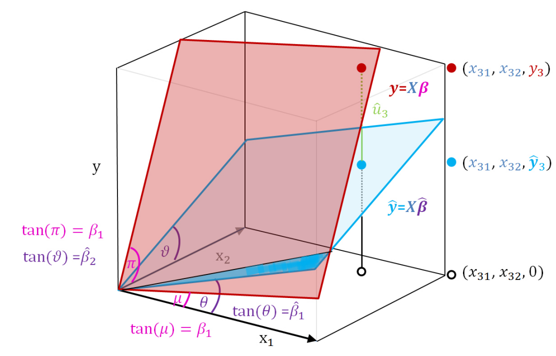

Abbildung 2: $3d$-Scatterplot mit wahrer und prognostizierter Regressionsebene

Interpretation:

Die Hyperebene (blau eingefärbte Fläche), in der $\hat{\boldsymbol{y}}$ liegt, wird von den Spalten der Matrix $\boldsymbol{X}$ aufgespannt. Diese ist $k = 2$-dimensional, während der Raum des Residuenvektors $\hat{\boldsymbol{u}}$ $n-k = 3-2 = 1$-dimensional ist. Beide Räume zusammen spannen den kompletten Raum, in dem sich Residuen, Regressoren und -anden befinden, auf (Dimension $n = 3$). Der Residuenvektor $\hat{\boldsymbol{u}}$ steht senkrecht auf dieser Fläche und bildet eine Basis des Komplementärraums; $\boldsymbol{y}$, $\hat{\boldsymbol{y}}$ und $\hat{\boldsymbol{u}}$ bilden ein rechtwinkliges Dreieck (nur rechtwinklig bei der KQ-Schätzung!). Zudem ist das wahre $\boldsymbol{\beta}$ mit dem wahren Fehler $\boldsymbol{u}$ eingezeichnet, der i. A. nicht senkrecht auf $\boldsymbol{X}$ steht.

In diesem Scatterplot, in dem nur eine Beobachtung (x_{31}, x_{32}, y_3) und dessen gefitteter Wert $(x_{31}, x_{32}, \hat{y}_3)$ zu sehen ist, sind zwei Regressionshyperebenen eingezeichnet: einmal die prognostizierte mit $\hat{\boldsymbol{y}} = \boldsymbol{X}\hat{\beta}$, wobei man deren geschätzten Parameter $\hat{\beta}_1, \hat{\beta}_2$ als Steigungen der Ebene bzw. als Tangens des Winkels zwischen Ebene und Regressor sehen kann und einmal die wahre $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta}$, wobei für die wahren Parameter β_1, β_2 das gleiche gilt.