Verallgemeinerte Kleinste Quadrate (GLS-Schätzer)

OLS-Schätzer sind effizient, wenn die Gauss-Markov Annahmen erfüllt sind. Diese implizieren u.a.

$$\mathrm{E}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'|\boldsymbol{X}) = \sigma^2 \boldsymbol{I}$$

Störterme, die diese Annahme erfüllen, werden häufig auch *'spherical disturbances'* genannt. Konkret bedeutet dies, dass

- die bedingte Varianz der Störterme ist nicht abhängig von den erklärenden Variablen, d.h. $\mathrm{var}(\varepsilon_i|\boldsymbol{X}) = \sigma^2$ (keine Heteroskedastizität), und
- die Kovarianzen der Störterme für unterschiedliche Zeit- bzw. Beobachtungspunkte alle gleich Null sind ($\mathrm{cov}(\varepsilon_i \varepsilon_j|\boldsymbol{X}) = 0$ für $i \neq j$), d.h. keine Autokorrelation.

Wenn die Annahme $\mathrm{E}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'|\boldsymbol{X}) = \sigma^2 \boldsymbol{I}$ *nicht* erfüllt ist, dann sind OLS-Schätzer zwar immer noch erwartungstreu, aber nicht effizient.

Dies haben wir bereits mehrfach gezeigt, aber noch einmal zur Erinnerung: im Fall von *'nonspherical disturbances'* (d.h. wenn $\mathrm{E}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}') = \sigma^2 \boldsymbol{\Omega} := \boldsymbol{V}$, wobei $\boldsymbol{\Omega}$ bzw. $\boldsymbol{V}$ eine beliebige *positiv-definite* Matrix sein kann) gilt

$$\begin{aligned}
&\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad \mathrm{E}(\boldsymbol{\varepsilon}) = 0, \quad \mathrm{E}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}') = \sigma_\varepsilon^2 \boldsymbol{\Omega} := \boldsymbol{V} \\
&\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}
\end{aligned}$$

$$\mathrm{E}(\hat{\boldsymbol{\beta}}) = \mathrm{E}[(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'(\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon})] = \boldsymbol{\beta} \quad \text{wenn } \mathrm{E}(\boldsymbol{X}'\boldsymbol{\varepsilon}) = 0 \quad \text{(Erwartungstreue)}$$

$$\begin{aligned}
\mathrm{var}(\hat{\boldsymbol{\beta}}) &= \mathrm{E}\left[(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})'\right] \\
&= \mathrm{E}\left[(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\right] \\
&= \sigma_\varepsilon^2(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Omega}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1} \neq \sigma_\varepsilon^2(\boldsymbol{X}'\boldsymbol{X})^{-1} \quad \text{wenn } \boldsymbol{\Omega} \neq \sigma^2\boldsymbol{I}
\end{aligned}$$

das heißt, wenn $\boldsymbol{\Omega} \neq \sigma^2\boldsymbol{I}$ ist OLS zwar erwartungstreu, aber *nicht* effizient!

Generalized Least Squares (GLS)

Angenommen, wir multiplizieren das Modell $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ mit einer deterministischen nichtsingulären $n \times n$ Transformationsmatrix $\boldsymbol{P}$

$$\boldsymbol{P}\boldsymbol{y} = (\boldsymbol{P}\boldsymbol{X})\boldsymbol{\beta} + \boldsymbol{P}\boldsymbol{\varepsilon}$$

Die Varianz-Kovarianzmatrix dieser transformierten Störterme ist (mit $\mathrm{E}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}') = \sigma^2\boldsymbol{\Omega}$)

$$\mathrm{E}(\boldsymbol{P}\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\boldsymbol{P}') = \boldsymbol{P}\,\mathrm{E}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}')\boldsymbol{P}' = \sigma^2\boldsymbol{P}\boldsymbol{\Omega}\boldsymbol{P}'$$

Da $\boldsymbol{P}$ exogen ist gilt $\mathrm{E}(\boldsymbol{P}\boldsymbol{\varepsilon}) = \boldsymbol{P}\,\mathrm{E}(\boldsymbol{\varepsilon}) = 0$. Wenn also eine Transformationsmatrix $\boldsymbol{P}$ gefunden werden könnte für die gilt

$$\boldsymbol{P}\boldsymbol{\Omega}\boldsymbol{P}' = \boldsymbol{I}$$

könnten die Parameter des transformierten Modells $\boldsymbol{P}\boldsymbol{y} = (\boldsymbol{P}\boldsymbol{X})\boldsymbol{\beta} + \boldsymbol{P}\boldsymbol{\varepsilon}$ mit OLS effizient geschätzt werden, da dieses transformierte Modell die Gauss-Markov Annahme erfüllt. Tatsächlich kann man zeigen, dass eine solche Transformationsmatrix $\boldsymbol{P}$ existiert.

Aus der linearen Algebra ist bekannt, dass für jede symmetrische und positiv definite Matrix $\boldsymbol{\Omega}$ eine nichtsinguläre Matrix $\boldsymbol{T}$ existiert für die gilt

$$\boldsymbol{\Omega} = \boldsymbol{T}\boldsymbol{T}'$$

(siehe Appendix: Eigenwerte und Eigenvektoren)
Da $\boldsymbol{T}$ nichtsingulär ist kann dies umgeformt werden zu

$$\boldsymbol{T}^{-1}\boldsymbol{\Omega}\boldsymbol{T}'^{-1} = \boldsymbol{I}$$

Für $\boldsymbol{P} = \boldsymbol{T}^{-1}$ erfüllt dies genau die gesuchte Bedingung $\boldsymbol{P}\boldsymbol{\Omega}\boldsymbol{P}' = \boldsymbol{I}$, da $(\boldsymbol{A}^{-1})' = (\boldsymbol{A}')^{-1}$, also kann das solcherart transformierte Modell mit OLS effizient geschätzt werden.

Man beachte außerdem, dass für $\boldsymbol{\Omega} = \boldsymbol{T}\boldsymbol{T}'$ folgt

$$\begin{aligned}
\boldsymbol{\Omega}^{-1} &= (\boldsymbol{T}\boldsymbol{T}')^{-1} \\
&= (\boldsymbol{T}')^{-1}\boldsymbol{T}^{-1} \\
&= (\boldsymbol{T}^{-1})'\boldsymbol{T}^{-1} \\
&= \boldsymbol{P}'\boldsymbol{P}
\end{aligned}$$

Daraus kann nun einfach der GLS oder *Aitken*-Schätzer für eine allgemeine Varianz-Kovarianzmatrix $\mathrm{E}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}') = \sigma^2\boldsymbol{\Omega}$ ermittelt werden, indem der OLS-Schätzer $\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}$ auf das transformierte Modell $\boldsymbol{P}\boldsymbol{y} = (\boldsymbol{P}\boldsymbol{X})\boldsymbol{\beta} + \boldsymbol{P}\boldsymbol{\varepsilon}$ angewandt wird, da $\mathrm{E}(\boldsymbol{P}\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\boldsymbol{P}') = \sigma^2\boldsymbol{P}\boldsymbol{\Omega}\boldsymbol{P}' = \sigma^2\boldsymbol{I}$

$$\begin{aligned}
\hat{\boldsymbol{\beta}}_{\mathrm{GLS}} &= \left[(\boldsymbol{P}\boldsymbol{X})'(\boldsymbol{P}\boldsymbol{X})\right]^{-1}(\boldsymbol{P}\boldsymbol{X})'(\boldsymbol{P}\boldsymbol{y}) \\
&= (\boldsymbol{X}'\boldsymbol{P}'\boldsymbol{P}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{P}'\boldsymbol{P}\boldsymbol{y} \\
&= (\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{y}
\end{aligned}$$

Ähnlich kann die Varianz-Kovarianzmatrix der Koeffizienten $\hat{\boldsymbol{\beta}}_{\mathrm{GLS}}$ bestimmt werden. Der unter diesen Umständen verzerrte OLS-Schätzer für die Varianz- Kovarianzmatrix von $\hat{\boldsymbol{\beta}}$ ist $\mathrm{var}(\hat{\boldsymbol{\beta}}_{\mathrm{OLS}}) = \sigma^2\left[\boldsymbol{X})'\boldsymbol{X})\right]^{-1}$, durch die Transformation mit $\boldsymbol{P}$ erhalten wir einen effizienten Schätzer für die Varianz- Kovarianzmatrix von $\hat{\boldsymbol{\beta}}$

$$\begin{aligned}
\mathrm{var}(\hat{\boldsymbol{\beta}}_{\mathrm{GLS}}) &= \sigma^2\left[(\boldsymbol{P}\boldsymbol{X})'(\boldsymbol{P}\boldsymbol{X})\right]^{-1} \\
&= \sigma^2(\boldsymbol{X}'\boldsymbol{P}'\boldsymbol{P}\boldsymbol{X})^{-1} \\
&= \sigma^2(\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X})^{-1}
\end{aligned}$$

Zusammenfassend halten wir fest, dass im Modell $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ mit $\mathrm{E}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}') = \sigma^2\boldsymbol{\Omega} = \boldsymbol{V}$ der GLS-Schätzer $\hat{\boldsymbol{\beta}}_{\mathrm{GLS}}$ ein BLUE für $\boldsymbol{\beta}$ ist.

Bemerkung: Manchmal wird der GLS-Schätzer auch als Funktion von $\boldsymbol{V} := \sigma^2\boldsymbol{\Omega}$ ausgedrückt. Aus

$$\mathrm{E}(\boldsymbol{P}\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\boldsymbol{P}') = \boldsymbol{P}\boldsymbol{V}\boldsymbol{P}'$$

folgt für $\boldsymbol{P}\boldsymbol{V}\boldsymbol{P}' = \boldsymbol{I}$ wieder $\boldsymbol{V}^{-1} = \boldsymbol{P}'\boldsymbol{P}$ und

$$\begin{aligned}\hat{\boldsymbol{\beta}}_{\mathrm{GLS}} &= (\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{y} \\ \mathrm{var}(\hat{\boldsymbol{\beta}}_{\mathrm{GLS}}) &= (\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}\end{aligned}$$

¶

Einen Schätzer s^2 für das unbekannte σ^2 erhält man wieder aus der Anwendung von OLS auf das transformierte Modell

$$\begin{aligned} s^2 &= \left[\boldsymbol{P}\boldsymbol{y} - \boldsymbol{P}\boldsymbol{X}\hat{\boldsymbol{\beta}}_{\mathrm{GLS}}\right]' \left[\boldsymbol{P}\boldsymbol{y} - \boldsymbol{P}\boldsymbol{X}\hat{\boldsymbol{\beta}}_{\mathrm{GLS}}\right] /(n-k) \\ &= \left[\boldsymbol{P}(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}_{\mathrm{GLS}})\right]' \left[\boldsymbol{P}(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}_{\mathrm{GLS}})\right] /(n-k) \\ &= \left(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}_{\mathrm{GLS}}\right)' \boldsymbol{\Omega}^{-1} \left(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}_{\mathrm{GLS}}\right) /(n-k) \\ &= \frac{\hat{\boldsymbol{\varepsilon}}'_{\mathrm{GLS}}\hat{\boldsymbol{\varepsilon}}_{\mathrm{GLS}}}{n-k} \end{aligned}$$

Da die Anwendung von OLS auf das transformierte Modell alle Gauss-Markov Annahmen erfüllt kann ein kleiner Stichprobentest der linearen Nullhypothese

$$\mathrm{H}_0: \qquad \boldsymbol{R}\boldsymbol{\beta} = \boldsymbol{r}$$

anhand der folgenden F-Statistik mit q Zähler- und $n-k$ Nennerfreiheitsgraden erfolgen

$$F = \frac{(\boldsymbol{R}\hat{\boldsymbol{\beta}}_{\mathrm{GLS}} - \boldsymbol{r})' \left[\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X})^{-1}\boldsymbol{R}'\right]^{-1} (\boldsymbol{R}\hat{\boldsymbol{\beta}}_{\mathrm{GLS}} - \boldsymbol{r})/q}{s^2} \sim \mathrm{F}_{(q,n-k)}$$

wobei q die Anzahl der Restriktionen ist.

Dieser GLS-Schätzer hat viele wichtige Anwendungen in der Ökonometrie, jedoch haben wir bislang ein zentrales Problem noch gar nicht angesprochen, nämlich wie die unbekannte Matrix $\boldsymbol{\Omega}$ (bzw. $\boldsymbol{V}$) geschätzt werden kann.

Wann immer diese Matrix aus den Daten geschätzt werden muss spricht man von *'Feasible Generalized Least Squares'* (FGLS) Schätzer. Man beachte aber, dass aus den Daten nicht alle Elemente von $\boldsymbol{\Omega}$ geschätzt werden können, da diese Matrix von der Dimension $n \times n$ mehr Parameter enthält als Beobachtungen zur Verfügung stehen. Hier müssen vereinfachende Annahmen getroffen werden, wie wir sie z.B. schon für die Schätzung von Modellen mit heteroskedastischen oder autokorrelierten Störtermen getroffen haben. Tatsächlich sind die dort diskutierten Datentransformationen Spezialfälle des hier vorgestellten GLS-Schätzers.

In manchen Fällen hat man auch a priori Information über die wahrscheinliche Form der Heteroskedastizität, die man mit Hilfe dieses GLS Schätzers nützen kann. Beispiele dafür sind das SUR Mehrgleichungsmodell und das Random Effects Modell der Panelökonometrie.

Anhang A

Matrixalgebra

A.1 Quadratische Formen und positiv definite Matrizen

Quadratische Formen spielen in der Ökonometrie eine wichtige Rolle. Die quadratische Form einer symmetrischen $n \times n$ Matrix $\boldsymbol{V}$ ist eine reellwertige Funktion definiert für alle $n \times 1$ Spaltenvektoren $\boldsymbol{x}$

$$f(\boldsymbol{x}) = \boldsymbol{x}'\boldsymbol{V}\boldsymbol{x}$$

Beispiel:
Gegeben sei eine 2×2 symmetrische Matrix $\boldsymbol{V}$ und ein 2×1 Spaltenvektor $\boldsymbol{x}$. Dann ist die quadratische Form definiert als ein Skalar

$$\begin{aligned}
\boldsymbol{x}'\boldsymbol{V}\boldsymbol{x} &= \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\
&= \begin{pmatrix} x_1 v_{11} + x_2 v_{21} & x_1 v_{12} + x_2 v_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\
&= x_1^2 v_{11} + x_2 x_1 v_{21} + x_1 x_2 v_{12} + x_2^2 v_{22} \\
&= \begin{matrix} v_{11}x_1^2 & + & 2v_{12}x_1x_2 \\ & + & v_{22}x_2^2 \end{matrix} \\
&= \sum_{i=1}^{2} \sum_{j=1}^{2} v_{ij} x_i x_j
\end{aligned}$$

(aufgrund der angenommenen Symmetrie von $\boldsymbol{V}$ ist $v_{12} = v_{21}$).
Beispiel:

$$\begin{aligned}
\boldsymbol{x}'\boldsymbol{V}\boldsymbol{x} &= \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 2 & 3 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\
&= \begin{pmatrix} 2x_1 + 3x_2 & 3x_1 + 1x_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\
&= 2x_1^2 + 3x_2x_1 + 3x_1x_2 + 1x_2^2 \\
&= 2x_1^2 + 6x_1x_2 + x_2^2
\end{aligned}$$

Man beachte, dass die *Summe der Hochzahlen* von x für jeden Term 2 ergibt, daher der Name quadratische Form.

Die quadratische Form einer symmetrischen 3×3 Matrix $\boldsymbol{V}$ und eines 3×1 Spaltenvektor $\boldsymbol{x}$ ist

$$
\begin{aligned}
\boldsymbol{x}'\boldsymbol{V}\boldsymbol{x} &= \begin{array}{ccccc} v_{11}x_1^2 & + & 2v_{12}x_1x_2 & + & 2v_{13}x_1x_3 \\ & & v_{22}x_2^2 & + & 2v_{12}x_1x_2 \\ & & & + & v_{33}x_3^2 \end{array} \\
&= \sum_{i=1}^{3}\sum_{j=1}^{3} v_{ij}x_ix_j
\end{aligned}
$$

Schließlich ist die quadratische Form einer symmetrischen $n \times n$ Matrix $\boldsymbol{V}$ und eines $n \times 1$ Spaltenvektor $\boldsymbol{x}$

$$
\begin{aligned}
\boldsymbol{x}'\boldsymbol{V}\boldsymbol{x} &= \begin{array}{ccccccccc} v_{11}x_1^2 & + & 2v_{12}x_1x_2 & + & 2v_{13}x_1x_3 & + & \cdots & + & 2v_{1n}x_1x_n \\ & & v_{22}x_2^2 & + & 2v_{12}x_1x_2 & + & \cdots & + & 2v_{2n}x_2x_n \\ & & & + & v_{33}x_3^2 & + & \cdots & + & 2v_{3n}x_3x_n \\ & & & & & & \ddots & & \vdots \\ & & & & & & & & v_{nn}x_n^2 \end{array} \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} v_{ij}x_ix_j = \sum_{i=1}^{n} v_{ii}x_i^2 + 2\sum_{j>i} v_{ij}x_ix_j
\end{aligned}
$$

A.1.1 Positiv definite Matrizen

Eine symmetrische Matrix $\boldsymbol{V}$ heißt **positiv definit** wenn

$$\boldsymbol{x}'\boldsymbol{V}\boldsymbol{x} > 0 \qquad \text{für alle} \quad \boldsymbol{x} \neq 0$$

d.h. wenn die quadratische Form für jeden beliebigen $\boldsymbol{x}$-Vektor – mit Ausnahme des Nullvektors – positiv ist.

Wenn $\boldsymbol{x}'\boldsymbol{V}\boldsymbol{x} \geq 0$ für alle $\boldsymbol{x} \neq 0$ heißt die quadratische Form positiv semidefinit.

Umdrehen der Ungleichheitszeichen definiert negativ definite und negativ semidefinite quadratische Formen. Ist die quadratische Form schließlich positiv für manche $\boldsymbol{x} \neq 0$ Vektoren und negativ für andere heißt sie indefinit.

A.1.2 Eigenschaften positiv definiter Matrizen

1. Die Hauptdiagonalelemente einer positiv definiten Matrix sind strikt positiv; die Hauptdiagonalelemente einer positiv semidefiniten Matrix sind nicht negativ.

2. Wenn $\boldsymbol{V}$ positiv definit ist existiert $\boldsymbol{V}^{-1}$ und ist ebenfalls positiv definit.

3. Wenn $\boldsymbol{V}$ eine $n \times k$ Matrix dann sind $\boldsymbol{V}'\boldsymbol{V}$ und $\boldsymbol{V}\boldsymbol{V}'$ positiv semidefinit.

4. Wenn $\boldsymbol{V}$ eine $n \times k$ Matrix und der $\text{rk}(\boldsymbol{V}) = k$ dann ist $\boldsymbol{V}'\boldsymbol{V}$ positiv definit (und damit auch nichtsingulär) (siehe z.B. Johnston & DiNardo 1997, 484).

   Man beachte, dass aus diesem Grund die $\boldsymbol{X}'\boldsymbol{X}$ positiv definit ist; deshalb sind die Bedingungen 2. Ordnung bei der Herleitung des OLS Schätzers erfüllt!

A.2 Differentiation linearer und quadratischer Formen

Seien $\boldsymbol{x}$ und $\boldsymbol{c}$ $k \times 1$ Spaltenvektoren, dann ist die Ableitung der linearen Funktion

$$g(\boldsymbol{x}) = \boldsymbol{c}'\boldsymbol{x} = \sum_{i=1}^{k} c_i x_i$$

nach $\boldsymbol{x}$

$$\frac{\partial g(\boldsymbol{x})}{\partial \boldsymbol{x}} = \boldsymbol{c}'$$

Sei $\boldsymbol{A}$ eine symmetrische $k \times k$ Matrix, dann ist die Ableitung der quadratischen Form

$$f(\boldsymbol{x}) = \boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}$$

nach $\boldsymbol{x}$ gleich

$$\frac{\partial f(\boldsymbol{x})}{\partial \boldsymbol{x}} = 2\boldsymbol{x}'\boldsymbol{A}$$

ein $1 \times k$ Vektor.

A.3 Eigenwerte und Eigenvektoren

Frage: Existiert für eine $k \times k$ Matrix $\boldsymbol{V}$ eine Skalar λ und ein Vektor $\boldsymbol{x}$ der die folgende Gleichung erfüllt?

$$\boldsymbol{V}\boldsymbol{x} = \lambda\boldsymbol{x}$$

Wenn ein solcher Skalar λ existiert wird er Eigenwert (auch charakteristische oder latente Wurzel) der Matrix $\boldsymbol{V}$ genannt, und $\boldsymbol{x}$ heißt Eigenvektor oder charakteristischer Vektor der Matrix $\boldsymbol{V}$.

Die Matrixgleichung $\boldsymbol{V}\boldsymbol{x} = \lambda\boldsymbol{x}$ kann auch als $\boldsymbol{V}\boldsymbol{x} - \lambda\boldsymbol{x} = \boldsymbol{0}$ oder

$$(\boldsymbol{V} - \lambda\boldsymbol{I})\boldsymbol{x} = \boldsymbol{0}$$

geschrieben werden ($\boldsymbol{I}$ ist die Einheitsmatrix), bzw. ausgeschrieben

$$(\boldsymbol{V} - \lambda\boldsymbol{I})\boldsymbol{x} = \begin{pmatrix} v_{11} - \lambda & v_{12} & \cdots & v_{1k} \\ v_{21} & v_{22} - \lambda & \cdots & v_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ v_{k1} & v_{k2} & \cdots & v_{kk} - \lambda \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Dies ist ein lineares *homogenes* Gleichungssystem.[1]

Eine nicht-triviale Lösung für $\boldsymbol{x}$ existiert nur, wenn die Matrix $(\boldsymbol{V} - \lambda \boldsymbol{I})$, auch charakteristische Matrix von $\boldsymbol{V}$ genannt, *singulär* ist, das heißt, wenn die Determinante Null ist

$$|\boldsymbol{V} - \lambda \boldsymbol{I}| = \begin{vmatrix} v_{11} - \lambda & v_{12} & \cdots & v_{1k} \\ v_{21} & v_{22} - \lambda & \cdots & v_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ v_{k1} & v_{k2} & \cdots & v_{kk} - \lambda \end{vmatrix} = 0$$

Dies liefert nach einer Laplace-Expansion ein Polynom vom Grad k in der Variable λ. Dieses Polynom hat k Nullstellen, die die Eigenwerte (charakteristischen Wurzeln) $(\lambda_1, \lambda_2, \ldots, \lambda_k)$ genannt werden.

Wenn $\boldsymbol{V}$ symmetrisch ist sind die Eigenwerte immer reelle Zahlen, die positiv oder negativ sein können.

Da jeder Eigenwert λ_i eine singuläre Matrix $|\boldsymbol{V} - \lambda \boldsymbol{I}|$ erzeugt ist die Lösung nicht eindeutig, es existieren unendlich viele Lösungen.

Für eine eindeutige Lösung muss zusätzlich eine *Normalisierung* erfolgen, d.h. eine zusätzliche Restriktion auferlegt werden. Diese zusätzliche Bedingung ist

$$\sum_{i=1}^{k} x_i^2 = 1$$

Durch Einsetzen von λ und *Normalisierung* kann für jeden Eigenwert λ_i ein eindeutiger Eigenvektor $\boldsymbol{x}|_{\lambda=\lambda_i}$ berechnet werden.

Die solcherart bestimmten Eigenvektoren $\boldsymbol{c}_i$ besitzen zwei wichtige Eigenschaften

1. Aus der Normalisierung folgt

$$\boldsymbol{c}_i'\boldsymbol{c}_i = \begin{pmatrix} x_1 & x_2 & \cdots & x_k \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_k \end{pmatrix} = \sum_{i=1}^{k} x_i^2 = 1$$

2. Das skalare Produkt

$$\boldsymbol{c}_i'\boldsymbol{c}_j = 0 \qquad \text{für } i \neq j$$

oder kürzer

$$\boldsymbol{c}_i'\boldsymbol{c}_j = \begin{cases} 1 & \text{wenn } i = j \\ 0 & \text{wenn } i \neq j \end{cases}$$

Achtung: Dies gilt nur, wenn die Matrix $\boldsymbol{V}$ symmetrisch ist und vollen Rang hat.

Vektoren, deren skalares Produkt gleich Null ist, heißen *orthogonal*. Da das skalare Produkt unterschiedlicher Eigenvektoren gleich Null ist ($\boldsymbol{c}_i'\boldsymbol{c}_j = 0$) sind Eigenvektoren symmetrischer Matrizen immer orthogonal. In der graphischen Abbildung bedeutet dies, dass sie rechtwinklig aufeinander stehen. Aufgrund der Normalisierung sind sie auch *orthonormal*.

[1] Ein homogenes Gleichungssystem hat die Form $\boldsymbol{Ax} = \boldsymbol{0}$, wobei $\boldsymbol{A}$ die Datenmatrix ist; ein nicht homogenes Gleichungssystem hat die Form $\boldsymbol{Ax} = \boldsymbol{b}$ mit $\boldsymbol{b} \neq 0$.

Beweis: Wir betrachten zwei unterschiedliche Eigenwerte und die dazugehörigen Eigenvektoren

$$\boldsymbol{V}\boldsymbol{c}_1 = \lambda_1 \boldsymbol{c}_1 \qquad \text{und} \qquad \boldsymbol{V}\boldsymbol{c}_2 = \lambda_2 \boldsymbol{c}_2$$

Wir (vor-)multiplizieren die erste Gleichung mit $\boldsymbol{c}_2'$ und die zweite Gleichung mit $\boldsymbol{c}_1'$

$$\boldsymbol{c}_2'\boldsymbol{V}\boldsymbol{c}_1 = \lambda_1 \boldsymbol{c}_2'\boldsymbol{c}_1 \qquad \text{und} \qquad \boldsymbol{c}_1'\boldsymbol{V}\boldsymbol{c}_2 = \lambda_2 \boldsymbol{c}_1'\boldsymbol{c}_2$$

Transponieren der zweiten (rechten) Gleichung gibt $\boldsymbol{c}_2'\boldsymbol{V}\boldsymbol{c}_1 = \lambda_2 \boldsymbol{c}_2'\boldsymbol{c}_1$, da $\boldsymbol{V}$ annahmegemäß symmetrisch ist. Daraus folgt

$$\lambda_1 \boldsymbol{c}_2'\boldsymbol{c}_1 = \lambda_2 \boldsymbol{c}_2'\boldsymbol{c}_1$$

Da die Eigenwerte symmetrischer Matrizen mit vollem Rang alle unterschiedlich sind folgt

$$\boldsymbol{c}_2'\boldsymbol{c}_1 = 0$$

Dies gilt für beliebige Paare von Eigenvektoren, d.h. die Eigenvektoren stehen paarweise orthogonal aufeinander.

Beispiel: Berechnen Sie Eigenwerte und Eigenvektoren der Matrix

$$\boldsymbol{V} = \begin{pmatrix} 2 & 2 \\ 2 & -1 \end{pmatrix}$$

Substitution in $(\boldsymbol{V} - \lambda\boldsymbol{I})\boldsymbol{x} = \boldsymbol{0}$ gibt

$$(\boldsymbol{V} - \lambda\boldsymbol{I}) = \begin{pmatrix} 2-\lambda & 2 \\ 2 & -1-\lambda \end{pmatrix}$$

Um die Eigenwerte zu berechnen müssen wir die Determinante dieser Matrix berechnen und Null setzen

$$\begin{vmatrix} 2-\lambda & 2 \\ 2 & -1-\lambda \end{vmatrix} = (2-\lambda)(-1-\lambda) - 4 = \lambda^2 - \lambda - 6 = 0$$

Dieses Polynom 2. Grades hat zwei Lösungen[2], die Eigenwerte $\lambda_1 = 3$ und $\lambda_2 = -2$. Für den ersten Eigenwert $\lambda_1 = 3$ ist die Matrixgleichung $(\boldsymbol{V} - \lambda\boldsymbol{I})\boldsymbol{x} = \boldsymbol{0}$

$$\begin{pmatrix} 2-3 & 2 \\ 2 & -1-3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -1 & 2 \\ 2 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Da die zwei Zeilen linear abhängig sind (die Eigenwerte λ wurden ja berechnet um eine singuläre Matrix zu erzeugen!) existieren unendlich viele Lösungen für die Eigenvektor, die die Bedingung der obigen Matrixgleichung

$$x_1 = 2x_2$$

erfüllen.

[2]Wenn $x^2 + px + q = 0$ ist $x_{1,2} = -\frac{p}{2} \pm \sqrt{\frac{p^2}{4} - q}$

Für eine eindeutige Lösung wird mit $\sum_{i=1}^{2} x_i^2$ normalisiert, d.h. die Restriktion $x_1^2 + x_2^2 = 1$ auferlegt

$$x_1^2 + x_2^2 = (2x_2)^2 + x_2^2 = 5x_2^2 = 1$$

Daraus folgt $x_2 = 1/\sqrt{5}$ und $x_1 = 2x_2 = 2/\sqrt{5}$. Der erste Eigenvektor ist also

$$\boldsymbol{c}_1 = \begin{pmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{pmatrix}$$

Ganz ähnlich kann für den zweiten Eigenwert $\lambda_2 = -2$ der Eigenvektor

$$\boldsymbol{c}_2 = \begin{pmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{pmatrix}$$

berechnet werden.

Diese Eigenvektoren erfüllen die Bedingungen $\boldsymbol{c}_i'\boldsymbol{c}_i = 1$ und $\boldsymbol{c}_i'\boldsymbol{c}_j = 0$ für $i \neq j$

$$\boldsymbol{c}_1'\boldsymbol{c}_1 = \begin{pmatrix} 2/\sqrt{5} & 1/\sqrt{5} \end{pmatrix} \begin{pmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{pmatrix} = \frac{4}{5} + \frac{1}{5} = 1$$

$$\boldsymbol{c}_2'\boldsymbol{c}_2 = \begin{pmatrix} -1/\sqrt{5} & 2/\sqrt{5} \end{pmatrix} \begin{pmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{pmatrix} = \frac{1}{5} + \frac{4}{5} = 1$$

$$\boldsymbol{c}_1'\boldsymbol{c}_2 = \begin{pmatrix} 2/\sqrt{5} & 1/\sqrt{5} \end{pmatrix} \begin{pmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{pmatrix} = \frac{-2}{5} + \frac{2}{5} = 0$$

Wir haben nun Eigenwerte λ_i und Eigenvektoren $\boldsymbol{c}_i$, die die Gleichung

$$(\boldsymbol{V} - \lambda_i \boldsymbol{I})\boldsymbol{c}_i = \boldsymbol{0} \qquad \text{bzw.} \qquad \boldsymbol{V}\boldsymbol{c}_i = \lambda_i \boldsymbol{c}_i$$

erfüllen.

Wenn aus allen Eigenvektoren eine $k \times k$ Matrix $\boldsymbol{C} = (\boldsymbol{c}_1 \;\; \boldsymbol{c}_2 \;\; \cdots \boldsymbol{c}_k)$ geformt wird können alle Lösungen in der Matrizengleichung $\boldsymbol{V}\boldsymbol{C} = \boldsymbol{C}\boldsymbol{\Lambda}$ zusammengefasst werden, wobei $\boldsymbol{\Lambda}$ eine Diagonalmatrix mit den Eigenwerten λ_i auf der Hauptdiagonale ist.[3]

Dies kann einfach gezeigt werden:

$$\begin{aligned}
\boldsymbol{V} \left(\begin{pmatrix} c_{11} \\ c_{12} \\ \vdots \\ c_{1k} \end{pmatrix} \begin{pmatrix} c_{21} \\ c_{22} \\ \vdots \\ c_{2k} \end{pmatrix} \cdots \begin{pmatrix} c_{k1} \\ c_{k2} \\ \vdots \\ c_{kk} \end{pmatrix} \right) &= \left(\lambda_1 \begin{pmatrix} c_{11} \\ c_{12} \\ \vdots \\ c_{1k} \end{pmatrix} \lambda_2 \begin{pmatrix} c_{21} \\ c_{22} \\ \vdots \\ c_{2k} \end{pmatrix} \cdots \; \lambda_k \begin{pmatrix} c_{k1} \\ c_{k2} \\ \vdots \\ c_{kk} \end{pmatrix} \right) \\
&= \begin{pmatrix} c_{11} & c_{21} & \dots & c_{k1} \\ c_{12} & c_{22} & \dots & c_{k2} \\ \vdots & \vdots & \ddots & \vdots \\ c_{1k} & c_{2k} & \dots & c_{kk} \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_k \end{pmatrix}
\end{aligned}$$

oder kompakter

$$\boldsymbol{V}\boldsymbol{C} = \boldsymbol{C}\boldsymbol{\Lambda}$$

[3] $\boldsymbol{\Lambda}$ ist der griechische Großbuchstabe Lambda.

Wenn $\boldsymbol{C}$ nicht singulär ist kann diese Gleichung mit $\boldsymbol{C}^{-1}$ vormultipliziert werden um nach der Diagonalmatrix mit den Eigenwerten $\boldsymbol{\Lambda}$ zu lösen

$$\boldsymbol{\Lambda} = \boldsymbol{C}^{-1}\boldsymbol{V}\boldsymbol{C}$$

wobei $\boldsymbol{\Lambda}$ eine Diagonalmatrix mit den Eigenwerten auf der Hauptdiagonale ist

$$\boldsymbol{\Lambda} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \lambda_k \end{pmatrix}$$

Die Matrix der Eigenvektoren $\boldsymbol{C}$ kann also verwendet werden um die Matrix $\boldsymbol{V}$ zu diagonalisieren.

Wir erinnern uns, dass die Eigenvektoren symmetrischer Matrizen paarweise orthogonal sind $\boldsymbol{c}_2'\boldsymbol{c}_1 = 0$. Aufgrund der Normalisierung ($||\boldsymbol{c}_i|| = 1$) folgt zudem

$$\boldsymbol{C}'\boldsymbol{C} = \boldsymbol{I}$$

woraus durch Nachmultiplikation mit $\boldsymbol{C}^{-1}$ weiter folgt

$$\boldsymbol{C}' = \boldsymbol{C}^{-1}$$

Deshalb gilt für symmetrische Matrizen auch

$$\boldsymbol{\Lambda} = \boldsymbol{C}'\boldsymbol{V}\boldsymbol{C}$$

oder

$$\boldsymbol{V} = \boldsymbol{C}\boldsymbol{\Lambda}\boldsymbol{C}'$$

Dies kann verwendet werden um zu zeigen, dass für jede symmetrische und positiv definite Matrix eine nichtsinguläre Matrix $\boldsymbol{P}$ existiert, mit deren Hilfe $\boldsymbol{V}$ folgendermaßen zerlegt werden kann.

$$\boldsymbol{V} = \boldsymbol{P}\boldsymbol{P}'$$

Beweis: Wenn $\boldsymbol{V}$ positiv definit ist sind alle Eigenwerte positiv. Deshalb kann die Diagonalmatrix mit den Eigenwerten $\boldsymbol{\Lambda}$ faktorisiert werden

$$\boldsymbol{\Lambda} = \boldsymbol{\Lambda}^{1/2}\boldsymbol{\Lambda}^{1/2}$$

mit

$$\boldsymbol{\Lambda}^{1/2} = \begin{pmatrix} \sqrt{\lambda_1} & 0 & \cdots & 0 \\ 0 & \sqrt{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \sqrt{\lambda_k} \end{pmatrix}$$

Einsetzen gibt

$$\boldsymbol{V} = \boldsymbol{C}\boldsymbol{\Lambda}\boldsymbol{C}' = \boldsymbol{C}\boldsymbol{\Lambda}^{1/2}\boldsymbol{\Lambda}^{1/2}\boldsymbol{C}' = (\boldsymbol{C}\boldsymbol{\Lambda}^{1/2})(\boldsymbol{C}\boldsymbol{\Lambda}^{1/2})'$$

Für $\boldsymbol{P} = (\boldsymbol{C}\boldsymbol{\Lambda}^{1/2})$ folgt deshalb $\boldsymbol{V} = \boldsymbol{P}\boldsymbol{P}'$.

Eigenschaften von Eigenwerten und Eigenvektoren

Abschließend seien noch einige Eigenschaften von Eigenwerten und Eigenvektoren von positiv definiten Matrizen $\boldsymbol{V}$ ohne Beweis aufgezählt. Für eine ausführlichere Darstellung siehe jedes Lehrbuch für lineare Algebra oder Johnston/DiNardo 1997, S. 478ff.

1. Die Summe der Eigenwerte ist gleich der Spur von $\boldsymbol{V}$: $\sum \lambda_i = \operatorname{tr}(\boldsymbol{V})$
2. Das Produkt der Eigenwerte ist gleich der Determinante von $\boldsymbol{V}$.
3. Der Rang der Matrix $\boldsymbol{V}$ ist gleich der Anzahl der Eigenwerte ungleich Null.
4. Die Eigenwerte von $\boldsymbol{V}^2$ sind das Quadrat der Eigenwerte von $\boldsymbol{V}$; die Eigenvektoren sind gleich.
5. Die Eigenwerte von $\boldsymbol{V}^{-1}$ sind die Kehrwerte der Eigenwerte von $\boldsymbol{V}$; die Eigenvektoren sind gleich.
6. Jeder Eigenwert einer idempotenten Matrix hat entweder den Wert Null oder Eins.
7. Der Rang einer idempotenten Matrix ist gleich ihrer Spur.
8. Eine notwendige und hinreichende Bedingung dafür, dass die Matrix $\boldsymbol{V}$ positiv definit ist, ist dass alle Eigenwerte von $\boldsymbol{V}$ positiv sind.

A.4 Verteilung Quadratischer Formen

Angenommen $\boldsymbol{x}$ sein ein $k \times 1$ Vektor mit standardnormalverteilten Zufallsvariablen X

$$\boldsymbol{x} \sim \mathrm{N}(\boldsymbol{0}, \boldsymbol{I})$$

dann ist

$$\boldsymbol{x}'\boldsymbol{x} \sim \chi^2(k)$$

Wenn

$$\boldsymbol{x} \sim \mathrm{N}(\boldsymbol{0}, \sigma^2\boldsymbol{I})$$

ist

$$\frac{X_1^2}{\sigma_2} + \frac{X_2^2}{\sigma_2} + \cdots + \frac{X_k^2}{\sigma_2} \sim \chi^2(k)$$

Dies kann auch geschrieben werden als

$$\frac{1}{\sigma^2}\boldsymbol{x}'\boldsymbol{x} \sim \chi^2(k) \qquad \text{bzw.} \qquad \boldsymbol{x}'(\sigma^2\boldsymbol{I})^{-1}\boldsymbol{x} \sim \chi^2(k)$$

Dabei ist die Matrix der quadratischen Form die Inverse der Varianzmatrix.

Wenn die Zufallsvariablen normalverteilt mit Erwartungswert Null, aber nicht notwendigerweise unabhängig verteilt sind

$$x \sim \mathrm{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$$

wobei $\boldsymbol{\Sigma}$ eine positiv definite Matrix ist, gilt immer noch

$$\boldsymbol{x}'\boldsymbol{\Sigma}^{-1}\boldsymbol{x} \sim \chi^2(k)$$

aber der Beweis is komplizierter, da die X nicht unabhängig sind.

Dazu ist werden die X in standardnormalverteilte Y Variablen transformiert.

Da $\boldsymbol{\Sigma}$ positiv definit ist kann die Matrix zerlegt werden in $\boldsymbol{\Sigma} = \boldsymbol{P}\boldsymbol{P}'$ wobei $\boldsymbol{P}$ eine nichtsinguläre $k \times k$ Matrix ist.

Die Inverse ist

$$\boldsymbol{\Sigma}^{-1} = (\boldsymbol{P}^{-1})'\boldsymbol{P}^{-1} \quad \text{bzw.} \quad \boldsymbol{P}^{-1}\boldsymbol{\Sigma}(\boldsymbol{P}^{-1})' = \boldsymbol{I}$$

Wir definieren mit Hilfe der $\boldsymbol{P}$ Matrix einen $k \times 1$ $\boldsymbol{y}$ Vektor

$$\boldsymbol{y} = \boldsymbol{P}^{-1}\boldsymbol{x}$$

Diese Y sind multivariat normalverteilt, da $\mathrm{E}(\boldsymbol{y}) = \mathbf{0}$ und

$$\begin{aligned}
\mathrm{var}(\boldsymbol{y}) &= \mathrm{E}\left[\boldsymbol{P}^{-1}\boldsymbol{x}\boldsymbol{x}'(\boldsymbol{P}^{-1})'\right] \\
&= \boldsymbol{P}^{-1}\boldsymbol{\Sigma}(\boldsymbol{P}^{-1})' \\
&= \boldsymbol{I}
\end{aligned}$$

Da die Y standardnormalverteilt sind gilt

$$\boldsymbol{y}'\boldsymbol{y} \sim \chi^2(k)$$

und weil $\boldsymbol{y} = \boldsymbol{P}^{-1}\boldsymbol{x}$ folgt weiter

$$\begin{aligned}
\boldsymbol{y}'\boldsymbol{y} &= \boldsymbol{x}'(\boldsymbol{P}^{-1})'\boldsymbol{P}^{-1}\boldsymbol{x} \\
&= \boldsymbol{x}'\boldsymbol{\Sigma}^{-1}\boldsymbol{x} \sim \chi^2(k)
\end{aligned}$$

Wenn $\boldsymbol{A}$ eine symmetrische und idempotente Matrix vom Rang $r \leq k$ ist kann diese mit Hilfe der orthogonalen Matrix der Eigenvektoren $\boldsymbol{Q}$ zerlegt werden

$$\boldsymbol{Q}'\boldsymbol{A}\boldsymbol{Q} = \boldsymbol{\Lambda} = \begin{pmatrix} \boldsymbol{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$$

d.h. $\boldsymbol{\Lambda}$ hat r Einsen und $k - r$ Nullen auf der Hauptdiagonale.

Wir definieren wieder für $\boldsymbol{x} \sim \mathrm{N}(\mathbf{0}, \boldsymbol{I})$

$$\boldsymbol{y} = \boldsymbol{Q}'\boldsymbol{x} \quad \text{bzw.} \quad \boldsymbol{x} = \boldsymbol{Q}\boldsymbol{y}$$

Dann gilt wieder

$$\mathrm{E}(\boldsymbol{y}) = \mathbf{0}$$

und

$$\begin{aligned}
\mathrm{var}(\boldsymbol{y}) &= \mathrm{E}(\boldsymbol{y}\boldsymbol{y}') \\
&= \mathrm{E}\left[\boldsymbol{Q}'\boldsymbol{x}\boldsymbol{x}'\boldsymbol{Q}'\right] \\
&= \boldsymbol{Q}'\boldsymbol{I}\boldsymbol{Q} \\
&= \boldsymbol{I}
\end{aligned}$$

Da $\boldsymbol{x} = \boldsymbol{Q}\boldsymbol{y}$ und

$$\boldsymbol{Q}'\boldsymbol{A}\boldsymbol{Q} = \begin{pmatrix} \boldsymbol{I}_r & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}$$

kann die quadratische Form $\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}$ geschrieben werden als

$$\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}'\boldsymbol{Q}'\boldsymbol{A}\boldsymbol{Q}\boldsymbol{y} = Y_1^2 + Y_2^2 + \cdots + Y_r^2$$

woraus folgt

$$\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} \sim \chi^2(r)$$

oder allgemeiner:

Wenn $\boldsymbol{x} \sim \mathrm{N}(\boldsymbol{0}, \sigma^2\boldsymbol{I})$ und $\boldsymbol{A}$ eine symmetrische und idempotente Matrix vom Rang r, dann gilt

$$\frac{1}{\sigma^2}\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} \sim \chi^2(r)$$

Dieses Ergebnis kann verwendet werden, um die Verteilung der Quadratsumme der Residuen im OLS Modell zu bestimmen.

Wir erinnern uns

$$\hat{\boldsymbol{\varepsilon}} = \boldsymbol{M}\boldsymbol{y} \quad \text{mit} \quad \boldsymbol{M} = \boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'$$

wobei $\boldsymbol{M}$ symmetrisch und idempotent ist mit $\boldsymbol{M}\boldsymbol{X} = 0$. Deshalb ist

$$\hat{\boldsymbol{\varepsilon}} = \boldsymbol{M}\boldsymbol{y} = \boldsymbol{M}(\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}) = \boldsymbol{M}\boldsymbol{\varepsilon}$$

und

$$\hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}} = \boldsymbol{\varepsilon}'\boldsymbol{M}\boldsymbol{\varepsilon}$$

Wenn $\boldsymbol{\varepsilon} \sim \mathrm{N}(\boldsymbol{0}, \sigma^2\boldsymbol{I})$ folgt

$$\frac{\hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}}}{\sigma^2} \sim \chi^2(r)$$

Da der Rang einer symmetrischen und idempotenten Matrix gleich der Spur (tr) ist folgt schließlich

$$\begin{aligned} \mathrm{rk}(\boldsymbol{M}) &= \mathrm{tr}\left[\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\right] \\ &= n - \mathrm{tr}\left[(\boldsymbol{X}'\boldsymbol{X})^{-1}(\boldsymbol{X}'\boldsymbol{X})\right] \\ &= n - k \end{aligned}$$

weshalb

$$\frac{\hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}}}{\sigma^2} \sim \chi^2(n-k)$$

A.4.1 Unabhängigkeit Quadratischer Formen

Sei wieder $\boldsymbol{x} \sim \mathrm{N}(\mathbf{0}, \sigma^2 \boldsymbol{I})$ und $\boldsymbol{A}$ und $\boldsymbol{B}$ symmetrische idempotente Matrizen mit den quadratischen Formen

$$\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} = (\boldsymbol{A}\boldsymbol{x})'(\boldsymbol{A}\boldsymbol{x}) \quad \text{und} \quad \boldsymbol{x}'\boldsymbol{B}\boldsymbol{x} = (\boldsymbol{B}\boldsymbol{x})'(\boldsymbol{B}\boldsymbol{x})$$

Die Korrelation zwischen den Vektoren $\boldsymbol{A}\boldsymbol{x}$ und $\boldsymbol{B}\boldsymbol{x}$ ist Null, wenn sie statistisch unabhängig verteilt sind. Die Kovarianz zwischen $\boldsymbol{A}\boldsymbol{x}$ und $\boldsymbol{B}\boldsymbol{x}$ ist

$$\mathrm{E}\left[(\boldsymbol{A}\boldsymbol{x})(\boldsymbol{B}\boldsymbol{x})'\right] = \mathrm{E}(\boldsymbol{A}\boldsymbol{x}\boldsymbol{x}'\boldsymbol{B}) = \sigma^2 \boldsymbol{A}\boldsymbol{B}$$

d.h. die Kovarianzen sind Null wenn $\boldsymbol{A}\boldsymbol{B} = \boldsymbol{B}\boldsymbol{A} = \mathbf{0}$.

A.5 Kronecker Produkt

Sei $\boldsymbol{A}$ eine $m \times n$ Matrix und $\boldsymbol{B}$ eine $k \times l$ Matrix, dann ist das *Kronecker Produkt* $\boldsymbol{A} \otimes \boldsymbol{B}$ definiert als

$$\boldsymbol{A} \otimes \boldsymbol{B} = \begin{pmatrix} a_{11}\boldsymbol{B} & a_{12}\boldsymbol{B} & \cdots & a_{1n}\boldsymbol{B} \\ a_{21}\boldsymbol{B} & a_{22}\boldsymbol{B} & \cdots & a_{2n}\boldsymbol{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}\boldsymbol{B} & a_{m2}\boldsymbol{B} & \cdots & a_{mn}\boldsymbol{B} \end{pmatrix}$$

Jedes einzelne Element der 'linken' Matrix wird mit der kompletten 'rechten' Matrix multipliziert. Das Ergebnis ist eine $mk \times nl$ Matrix.

Eigenschaften des Kronecker Produkts

- $(\boldsymbol{A} \otimes \boldsymbol{B})(\boldsymbol{C} \otimes \boldsymbol{D}) = (\boldsymbol{A}\boldsymbol{C}) \otimes (\boldsymbol{B}\boldsymbol{D})$

  für $\boldsymbol{A}_{m\times n}$, $\boldsymbol{B}_{k\times l}$, $\boldsymbol{C}_{n\times p}$, $\boldsymbol{D}_{l\times q}$.

- $(\boldsymbol{A} \otimes \boldsymbol{B})' = \boldsymbol{A}' \otimes \boldsymbol{B}'$.

- Seien $\boldsymbol{A}_{m\times m}$ und $\boldsymbol{B}_{k\times k}$ nichtsinguläre Matrizen, dann ist $\boldsymbol{A} \otimes \boldsymbol{B}$ ebenfalls nichtsingulär und es gilt

$$(\boldsymbol{A} \otimes \boldsymbol{B})^{-1} = \boldsymbol{A}^{-1} \otimes \boldsymbol{B}^{-1}$$