

Seminar

Zulässigkeit in der Regressionsanalyse

Prof. Dr. Götz Trenkler

Inoncent Agueusop

8. Januar 2009

ÜBERSICHT

- 1 Definition
- 2 Motivationsbeispiel
- 3 Zulässigkeit beim multiplen linearen Regressionsmodell
- 4 Literatur

ÜBERSICHT

- 1 Definition
- 2 Motivationsbeispiel
- 3 Zulässigkeit beim multiplen linearen Regressionsmodell
- 4 Literatur

ÜBERSICHT

- 1 Definition
- 2 Motivationsbeispiel
- 3 Zulässigkeit beim multiplen linearen Regressionsmodell
- 4 Literatur

ÜBERSICHT

- 1 Definition
- 2 Motivationsbeispiel
- 3 Zulässigkeit beim multiplen linearen Regressionsmodell
- 4 Literatur

Entscheidungsregel

- Eine Entscheidungsregel ist eine Abbildung $\delta : \mathbb{R}^n \rightarrow D$, so dass für jede Stichprobenrealisation $y \mapsto \delta(y) = \tilde{\beta}$ gilt.
- Eine Entscheidungsregel kann ein Schätzer, ein Konfidenzintervall oder eine Teststatistik sein.
- Es wird in diesem Bericht nur die Schätzer als Entscheidungsregel betrachtet.

Entscheidungsregel

- Eine Entscheidungsregel ist eine Abbildung $\delta : \mathbb{R}^n \rightarrow D$, so dass für jede Stichprobenrealisation $y \mapsto \delta(y) = \tilde{\beta}$ gilt.
- Eine Entscheidungsregel kann ein Schätzer, ein Konfidenzintervall oder eine Teststatistik sein.
- Es wird in diesem Bericht nur die Schätzer als Entscheidungsregel betrachtet.

Entscheidungsregel

- Eine Entscheidungsregel ist eine Abbildung $\delta : \mathbb{R}^n \rightarrow D$, so dass für jede Stichprobenrealisation $y \mapsto \delta(y) = \tilde{\beta}$ gilt.
- Eine Entscheidungsregel kann ein Schätzer, ein Konfidenzintervall oder eine Teststatistik sein.
- Es wird in diesem Bericht nur die Schätzer als Entscheidungsregel betrachtet.

Verlustfunktion und Risiko

Verlustfunktion

- $L: (\Theta \times D) \rightarrow \mathbb{R}$
 $(\beta, \tilde{\beta}) \mapsto L(\beta, \tilde{\beta})$. β ist der wahre Parameter und $\tilde{\beta}$ eine mögliche Entscheidung für β .
- Häufig betrachtet man den erwarteten Verlust:
 $\rho(\beta, \tilde{\beta}) = E[L(\beta, \delta(y))]$, wird auch Risiko genannt.

Verlustfunktion und Risiko

Verlustfunktion

- $L: (\Theta \times D) \rightarrow \mathbb{R}$
 $(\beta, \tilde{\beta}) \mapsto L(\beta, \tilde{\beta})$. β ist der wahre Parameter und $\tilde{\beta}$ eine mögliche Entscheidung für β .
- Häufig betrachtet man den erwarteten Verlust:
 $\rho(\beta, \tilde{\beta}) = E[L(\beta, \delta(y))]$, wird auch Risiko genannt.

Verlustfunktion und Risiko

Verlustfunktion

- $L: (\Theta \times D) \rightarrow \mathbb{R}$
 $(\beta, \tilde{\beta}) \mapsto L(\beta, \tilde{\beta})$. β ist der wahre Parameter und $\tilde{\beta}$ eine mögliche Entscheidung für β .
- Häufig betrachtet man den erwarteten Verlust:
 $\rho(\beta, \tilde{\beta}) = E[L(\beta, \delta(y))]$, wird auch Risiko genannt.

Beispiele von Verlustfunktionen

- $L(\beta, \tilde{\beta}) = \beta - \tilde{\beta}$, $L(\beta, \tilde{\beta}) = |\beta - \tilde{\beta}|$, oder die quadratische Verlustfunktion in verschiedenen Formen:
 $L(\beta, \tilde{\beta}) = (\beta - \tilde{\beta})^2$, $L(\beta, \tilde{\beta}) = (\beta - \tilde{\beta})'(\beta - \tilde{\beta})$,
 $L(\beta, \tilde{\beta}) = (\beta - \tilde{\beta})(\beta - \tilde{\beta})'$.
- $\rho(\beta, \tilde{\beta}) = E[(\beta - \tilde{\beta})^2] = E[(\beta + E(\tilde{\beta}) - E(\tilde{\beta}) - \tilde{\beta})^2] =$
 $\text{Var}(\tilde{\beta}) + [E(\tilde{\beta}) - \beta]^2$

Beispiele von Verlustfunktionen

- $L(\beta, \tilde{\beta}) = \beta - \tilde{\beta}$, $L(\beta, \tilde{\beta}) = |\beta - \tilde{\beta}|$, oder die quadratische Verlustfunktion in verschiedenen Formen:
 $L(\beta, \tilde{\beta}) = (\beta - \tilde{\beta})^2$, $L(\beta, \tilde{\beta}) = (\beta - \tilde{\beta})'(\beta - \tilde{\beta})$,
 $L(\beta, \tilde{\beta}) = (\beta - \tilde{\beta})(\beta - \tilde{\beta})'$.
- $\rho(\beta, \tilde{\beta}) = E[(\beta - \tilde{\beta})^2] = E[(\beta + E(\tilde{\beta}) - E(\tilde{\beta}) - \tilde{\beta})^2] = \text{Var}(\tilde{\beta}) + [E(\tilde{\beta}) - \beta]^2$

Vergleich von Entscheidungsregeln bzw. Schätzern

- $\tilde{\beta}_1$ heißt **besser** als $\tilde{\beta}_2$, wenn
 $\rho(\beta, \tilde{\beta}_1) \leq \rho(\beta, \tilde{\beta}_2)$ für alle β in dem Parameter Raum gilt und
 $\rho(\beta_*, \tilde{\beta}_1) < \rho(\beta_*, \tilde{\beta}_2)$ für mindestens ein β_*
 - $\tilde{\beta}_1$ heißt **nicht schlechter** als $\tilde{\beta}_2$ falls
 $\rho(\beta, \tilde{\beta}_1) \leq \rho(\beta, \tilde{\beta}_2)$ für alle β in dem Parameter Raum Θ .
- $\tilde{\beta}_1$ und $\tilde{\beta}_2$ sind **unvergleichbar** falls β_1 und β_2 existieren mit
 $\rho(\beta_1, \tilde{\beta}_1) < \rho(\beta_1, \tilde{\beta}_2)$ und $\rho(\beta_2, \tilde{\beta}_1) > \rho(\beta_2, \tilde{\beta}_2)$.

Vergleich von Entscheidungsregeln bzw. Schätzern

- $\tilde{\beta}_1$ heißt **besser** als $\tilde{\beta}_2$, wenn
$$\rho(\beta, \tilde{\beta}_1) \leq \rho(\beta, \tilde{\beta}_2) \text{ für alle } \beta \text{ in dem Parameter Raum gilt und}$$
$$\rho(\beta_*, \tilde{\beta}_1) < \rho(\beta_*, \tilde{\beta}_2) \text{ für mindestens ein } \beta_*$$
- $\tilde{\beta}_1$ heißt **nicht schlechter** als $\tilde{\beta}_2$ falls
$$\rho(\beta, \tilde{\beta}_1) \leq \rho(\beta, \tilde{\beta}_2) \text{ für alle } \beta \text{ in dem Parameter Raum } \Theta.$$
- $\tilde{\beta}_1$ und $\tilde{\beta}_2$ sind **unvergleichbar** falls β_1 und β_2 existieren mit
$$\rho(\beta_1, \tilde{\beta}_1) < \rho(\beta_1, \tilde{\beta}_2) \text{ und } \rho(\beta_2, \tilde{\beta}_1) > \rho(\beta_2, \tilde{\beta}_2).$$

Vergleich von Entscheidungsregeln bzw. Schätzern

- $\tilde{\beta}_1$ heißt **besser** als $\tilde{\beta}_2$, wenn
$$\rho(\beta, \tilde{\beta}_1) \leq \rho(\beta, \tilde{\beta}_2) \text{ für alle } \beta \text{ in dem Parameter Raum gilt und}$$
$$\rho(\beta_*, \tilde{\beta}_1) < \rho(\beta_*, \tilde{\beta}_2) \text{ für mindestens ein } \beta_*$$
- $\tilde{\beta}_1$ heißt **nicht schlechter** als $\tilde{\beta}_2$ falls
$$\rho(\beta, \tilde{\beta}_1) \leq \rho(\beta, \tilde{\beta}_2) \text{ für alle } \beta \text{ in dem Parameter Raum } \Theta.$$
- $\tilde{\beta}_1$ und $\tilde{\beta}_2$ sind **unvergleichbar** falls β_1 und β_2 existieren mit
$$\rho(\beta_1, \tilde{\beta}_1) < \rho(\beta_1, \tilde{\beta}_2) \text{ und } \rho(\beta_2, \tilde{\beta}_1) > \rho(\beta_2, \tilde{\beta}_2).$$

Vergleich von Entscheidungsregeln bzw. Schätzern

- $\tilde{\beta}_1$ heißt **besser** als $\tilde{\beta}_2$, wenn
$$\rho(\beta, \tilde{\beta}_1) \leq \rho(\beta, \tilde{\beta}_2) \text{ für alle } \beta \text{ in dem Parameter Raum gilt und}$$
$$\rho(\beta_*, \tilde{\beta}_1) < \rho(\beta_*, \tilde{\beta}_2) \text{ für mindestens ein } \beta_*$$
- $\tilde{\beta}_1$ heißt **nicht schlechter** als $\tilde{\beta}_2$ falls
$$\rho(\beta, \tilde{\beta}_1) \leq \rho(\beta, \tilde{\beta}_2) \text{ für alle } \beta \text{ in dem Parameter Raum } \Theta.$$
- $\tilde{\beta}_1$ und $\tilde{\beta}_2$ sind **unvergleichbar** falls β_1 und β_2 existieren mit
$$\rho(\beta_1, \tilde{\beta}_1) < \rho(\beta_1, \tilde{\beta}_2) \text{ und } \rho(\beta_2, \tilde{\beta}_1) > \rho(\beta_2, \tilde{\beta}_2).$$

Zulässigkeit

- Ein Schätzer heißt zulässig (admissible), falls es keinen anderen gibt, der besser als dieser Schätzer ist. Andernfalls heißt er unzulässig.
- Zwei Zulässigen Schätzer sind entweder äquivalent oder unvergleichbar.
- Ist ein Schätzer bezüglich der ungewichteten quadratischen Verlustfunktion zulässig, dann ist er auch bezüglich jeder gewichteten quadratischen Verlustfunktion zulässig.

Zulässigkeit

- Ein Schätzer heißt zulässig (admissible), falls es keinen anderen gibt, der besser als dieser Schätzer ist. Andernfalls heißt er unzulässig.
- Zwei Zulässigen Schätzer sind entweder äquivalent oder unvergleichbar.
- Ist ein Schätzer bezüglich der ungewichteten quadratischen Verlustfunktion zulässig, dann ist er auch bezüglich jeder gewichteten quadratischen Verlustfunktion zulässig.

Zulässigkeit

- Ein Schätzer heißt zulässig (admissible), falls es keinen anderen gibt, der besser als dieser Schätzer ist. Andernfalls heißt er unzulässig.
- Zwei Zulässigen Schätzer sind entweder äquivalent oder unvergleichbar.
- Ist ein Schätzer bezüglich der ungewichteten quadratischen Verlustfunktion zulässig, dann ist er auch bezüglich jeder gewichteten quadratischen Verlustfunktion zulässig.

Zulässigkeit

- Zulässigkeit bezüglich der reellwertigen quadratischen Verlustfunktion $L(\beta, \tilde{\beta}) = (\beta - \tilde{\beta})'(\beta - \tilde{\beta})$ impliziert die Zulässigkeit bezüglich der Matrixwertigen quadratischen Verlustfunktion
$$L(\beta, \tilde{\beta}) = (\beta - \tilde{\beta})(\beta - \tilde{\beta})'.$$
- Zulässigkeit bezüglich $(\beta - \tilde{\beta})'(\beta - \tilde{\beta})$ impliziert auch die Zulässigkeit bezüglich jeder gewichteten quadratischen Verlustfunktion
$$L(\beta, \tilde{\beta}) = (\beta - \tilde{\beta})W(\beta - \tilde{\beta})'.$$

Zulässigkeit

- Zulässigkeit bezüglich der reellwertigen quadratischen Verlustfunktion $L(\beta, \tilde{\beta}) = (\beta - \tilde{\beta})'(\beta - \tilde{\beta})$ impliziert die Zulässigkeit bezüglich der Matrixwertigen quadratischen Verlustfunktion
$$L(\beta, \tilde{\beta}) = (\beta - \tilde{\beta})(\beta - \tilde{\beta})'.$$
- Zulässigkeit bezüglich $(\beta - \tilde{\beta})'(\beta - \tilde{\beta})$ impliziert auch die Zulässigkeit bezüglich jeder gewichteten quadratischen Verlustfunktion
$$L(\beta, \tilde{\beta}) = (\beta - \tilde{\beta})W(\beta - \tilde{\beta})'.$$

Lineares Regressionsmodell

Gegeben sei das lineare Modell : $\mathbf{y} = \mathbf{X}\beta + \epsilon$ mit Annahmen (i)-(iv).

- (i) \mathbf{X} ist nicht stochastisch.
- (ii) Die ϵ_i sind stochastisch.
- (iii) $E(\epsilon_i) = 0$ für alle $i = 1, \dots, n$.
- (iv) Die Kovarianzmatrix $\text{Cov}(\epsilon) = \sigma^2 \mathbf{I}_n$, $\sigma^2 > 0$.

Man möchte den Parameter Vektor β schätzen.

Homogene Schätzer

$$L^h(\beta) = \{\tilde{\beta} = Ay : A \in \mathbb{R}^{p \times n}\}.$$

Nicht homogene Schätzer

$$L(\beta) = \{\tilde{\beta} = Ay + a : A \in \mathbb{R}^{p \times n}, a \in \mathbb{R}^p\}$$

Die ungewichtete quadratische Verlustfunktion

$L(\beta, \tilde{\beta}) = (\beta - \tilde{\beta})'(\beta - \tilde{\beta})$ wird hier verwendet und als Risiko
manchmal $\rho(\beta, \tilde{\beta}) = \text{tr}[\text{MSE}(\beta, \tilde{\beta})]$.

Homogene Schätzer

$$L^h(\beta) = \{\tilde{\beta} = Ay : A \in \mathbb{R}^{p \times n}\}.$$

Nicht homogene Schätzer

$$L(\beta) = \{\tilde{\beta} = Ay + a : A \in \mathbb{R}^{p \times n}, a \in \mathbb{R}^p\}$$

Die ungewichtete quadratische Verlustfunktion

$L(\beta, \tilde{\beta}) = (\beta - \tilde{\beta})'(\beta - \tilde{\beta})$ wird hier verwendet und als Risiko
manchmal $\rho(\beta, \tilde{\beta}) = \text{tr}[\text{MSE}(\beta, \tilde{\beta})]$.

Homogene Schätzer

$$L^h(\beta) = \{\tilde{\beta} = Ay : A \in \mathbb{R}^{p \times n}\}.$$

Nicht homogene Schätzer

$$L(\beta) = \{\tilde{\beta} = Ay + a : A \in \mathbb{R}^{p \times n}, a \in \mathbb{R}^p\}$$

Die ungewichtete quadratische Verlustfunktion

$L(\beta, \tilde{\beta}) = (\beta - \tilde{\beta})'(\beta - \tilde{\beta})$ wird hier verwendet und als Risiko
manchmal $\rho(\beta, \tilde{\beta}) = \text{tr}[\text{MSE}(\beta, \tilde{\beta})]$.

Motivationsbeispiel.

Sei das Modell: $y = \mu 1_n + \epsilon$.

Homogener Schätzer

$\tilde{\beta} = a' y$ und $\rho(\beta, \tilde{\beta}) = E[(a' y - \mu)^2] = \sigma^2 a' a + \mu^2 (1_n' a - 1)^2$ Wenn es kein α gibt, so dass $a = \alpha 1_n$ gilt, dann gibt es andere Schätzer, die besser als $a' y$ sind.

Z.B. $b' y$ mit $b = \psi 1_n$ und $\psi = \frac{a' 1_n}{n}$

$$\alpha = \frac{1}{n}$$

Für $\alpha = \frac{1}{n}$, haben wir $\rho(\mu, \alpha 1_n' y) = \frac{\sigma^2}{n}$ und alle andere Schätzer haben größerer Risiko, nämlich

$$\rho(\mu, \psi 1_n' y) = \sigma^2 n \psi^2 + \mu^2 (\psi n - 1)^2 > \frac{\sigma^2}{n}$$

Motivationsbeispiel.

Sei das Modell: $y = \mu 1_n + \epsilon$.

Homogener Schätzer

$\tilde{\beta} = a' y$ und $\rho(\beta, \tilde{\beta}) = E[(a' y - \mu)^2] = \sigma^2 a' a + \mu^2 (1_n' a - 1)^2$ Wenn es kein α gibt, so dass $a = \alpha 1_n$ gilt, dann gibt es andere Schätzer, die besser als $a' y$ sind.

Z.B. $b' y$ mit $b = \psi 1_n$ und $\psi = \frac{a' 1_n}{n}$

$$\alpha = \frac{1}{n}$$

Für $\alpha = \frac{1}{n}$, haben wir $\rho(\mu, \alpha 1_n' y) = \frac{\sigma^2}{n}$ und alle andere Schätzer haben größerer Risiko, nämlich

$$\rho(\mu, \psi 1_n' y) = \sigma^2 n \psi^2 + \mu^2 (\psi n - 1)^2 > \frac{\sigma^2}{n}$$

Motivationsbeispiel

Der Schätzer $\tilde{\beta} = a'y = \alpha 1_n y$ ist im allgemeinen in der Klasse der homogenen Schätzer zulässig, wenn $0 \leq \alpha \leq \frac{1}{n}$ ist.

Kompakter Schreibweise: $1_n a' a 1_n' \leq 1_n a'$.

Nicht homogene Schätzer für β

Ein nicht homogener Schätzer $\tilde{\beta} = a'y + d$ ist genau dann zulässig, wenn der homogene Teil zulässig ist und $d = 1 - a'1_n$.

Motivationsbeispiel

Der Schätzer $\tilde{\beta} = a'y = \alpha 1_n y$ ist im allgemeinen in der Klasse der homogenen Schätzer zulässig, wenn $0 \leq \alpha \leq \frac{1}{n}$ ist.

Kompakter Schreibweise: $1_n a' a 1_n' \leq 1_n a'$.

Nicht homogene Schätzer für β

Ein nicht homogener Schätzer $\tilde{\beta} = a'y + d$ ist genau dann zulässig, wenn der homogene Teil zulässig ist und $d = 1 - a'1_n$.

Allgemeines Modell $Y = X\beta + \epsilon$

- Die Zulässigkeit des Schätzers $\tilde{\beta} = Ay + a$ bezieht sich auf A und a .

Implizite Voraussetzungen

- Theorem 1:
 $\tilde{\beta} = Ay + a$ ist bezüglich der ungewichteten Verlustfunktion zulässig genau dann, wenn $XA A' X' \leq_L XA$ und $a \in C(I_p - AX)$ gilt.
- Dann folgt, dass die Matrix XA symmetrisch ist und a liegt in dem selben Unterraum als $\text{Bias}(Ay) = E(Ay) - \beta = (AX - I_p)\beta$.

Allgemeines Modell $Y = X\beta + \epsilon$

- Die Zulässigkeit des Schätzers $\tilde{\beta} = Ay + a$ bezieht sich auf A und a .

Implizite Voraussetzungen

- Theorem 1:
 $\tilde{\beta} = Ay + a$ ist bezüglich der ungewichteten Verlustfunktion zulässig genau dann, wenn $XA A' X' \leq_L XA$ und $a \in C(I_p - AX)$ gilt.
- Dann folgt, dass die Matrix XA symmetrisch ist und a liegt in dem selben Unterraum als $\text{Bias}(Ay) = E(Ay) - \beta = (AX - I_p)\beta$.

Allgemeines Modell $Y = X\beta + \epsilon$

- Die Zulässigkeit des Schätzers $\tilde{\beta} = Ay + a$ bezieht sich auf A und a .

Implizite Voraussetzungen

- Theorem 1:
 $\tilde{\beta} = Ay + a$ ist bezüglich der ungewichteten Verlustfunktion zulässig genau dann, wenn $XA A' X' \leq_L XA$ und $a \in C(I_p - AX)$ gilt.
- Dann folgt, dass die Matrix XA symmetrisch ist und a liegt in dem selben Unterraum als $\text{Bias}(Ay) = E(Ay) - \beta = (AX - I_p)\beta$.

Allgemeines Modell $Y = X\beta + \epsilon$

Alternative zum Theorem 1

- $\tilde{\beta} = Ay + a$ ist genau dann zulässig, wenn $XA = A'X'$, $\sigma(XA) \subset [0, 1]$ und $a \in C(I_p - AX)$, wobei $\sigma(XA)$ die Menge aller Eigenwerte der Matrix XA bezeichnet.

Ist $\tilde{\beta}$ zulässig in der Menge $L(\beta)$ und es gilt zusätzlich $p-2 \leq \text{Rang}(I_p - AX)$ dann ist $\tilde{\beta}$ unter allen Schätzern für β zulässig.

zurück zum Beispiel $y = 1_n \mu + \epsilon$

$\tilde{\beta} = \frac{1}{n+k} 1_n' y$ ist für alle $k \geq 0$ zulässig, da

$XA = 1_n \frac{1}{n+k} 1_n' = \frac{1}{n+k} 1_n 1_n' = A'X'$ und $\sigma(XA) = \left\{ \frac{1}{n+k} \right\} \subset [0, 1]$.

Allgemeines Modell $Y = X\beta + \epsilon$

Alternative zum Theorem 1

- $\tilde{\beta} = Ay + a$ ist genau dann zulässig, wenn $XA = A'X'$, $\sigma(XA) \subset [0, 1]$ und $a \in C(I_p - AX)$, wobei $\sigma(XA)$ die Menge aller Eigenwerte der Matrix XA bezeichnet.

Ist $\tilde{\beta}$ zulässig in der Menge $L(\beta)$ und es gilt zusätzlich $p-2 \leq \text{Rang}(I_p - AX)$ dann ist $\tilde{\beta}$ unter allen Schätzern für β zulässig.

zurück zum Beispiel $y = 1_n \mu + \epsilon$

$\tilde{\beta} = \frac{1}{n+k} 1'_n y$ ist für alle $k \geq 0$ zulässig, da

$XA = 1_n \frac{1}{n+k} 1'_n = \frac{1}{n+k} 1_n 1'_n = A'X'$ und $\sigma(XA) = \left\{ \frac{1}{n+k} \right\} \subset [0, 1]$.

Allgemeines Modell $Y = X\beta + \epsilon$

Alternative zum Theorem 1

- $\tilde{\beta} = Ay + a$ ist genau dann zulässig, wenn $XA = A'X'$, $\sigma(XA) \subset [0, 1]$ und $a \in C(I_p - AX)$, wobei $\sigma(XA)$ die Menge aller Eigenwerte der Matrix XA bezeichnet.

Ist $\tilde{\beta}$ zulässig in der Menge $L(\beta)$ und es gilt zusätzlich $p-2 \leq \text{Rang}(I_p - AX)$ dann ist $\tilde{\beta}$ unter allen Schätzern für β zulässig.

zurück zum Beispiel $y = 1_n \mu + \epsilon$

$\tilde{\beta} = \frac{1}{n+k} 1_n' y$ ist für alle $k \geq 0$ zulässig, da

$XA = 1_n \frac{1}{n+k} 1_n' = \frac{1}{n+k} 1_n 1_n' = A'X'$ und $\sigma(XA) = \{\frac{1}{n+k}\} \subset [0, 1]$.

Theorem 2: Explizite Voraussetzungen

$\tilde{\beta} = Ay + a$ ist genau dann zulässig, wenn A und a
 $A = (X'X)^{-\frac{1}{2}}G(X'X)^{-\frac{1}{2}}X'$ und $a = (X'X)^{-\frac{1}{2}}(I_p - G)(X'X)^{\frac{1}{2}}g$ gilt,
 wobei $\sigma(G) \subset [0,1]$ und $g \in \mathbb{R}^p$ gilt.

Vom Theorem 2 folgt, dass jeder zulässige Schätzer sich in form:
 $\tilde{\beta} = S^{-1/2}GS^{1/2}\hat{\beta} + S^{-1/2}(I_p - G)S^{1/2}g$ schreiben lässt, mit
 $S = X'X$ und $\hat{\beta}$ der KQ-Schätzer.

Ist $Ay+a$ zulässig, dann folgt $\text{Cov}(Ay+a) \leq \text{Cov}(\hat{\beta})$. Aber
 $\text{Cov}(Ay+a) \leq \text{Cov}(\hat{\beta})$ bedeutet nicht unbedingt, dass $Ay+a$ für β
 zulässig ist.

Theorem 2: Explizite Voraussetzungen

$\tilde{\beta} = Ay + a$ ist genau dann zulässig, wenn A und a
 $A = (X'X)^{-\frac{1}{2}}G(X'X)^{-\frac{1}{2}}X'$ und $a = (X'X)^{-\frac{1}{2}}(I_p - G)(X'X)^{\frac{1}{2}}g$ gilt,
 wobei $\sigma(G) \subset [0,1]$ und $g \in \mathbb{R}^p$ gilt.

Vom Theorem 2 folgt, dass jeder zulässige Schätzer sich in form:
 $\tilde{\beta} = S^{-1/2}GS^{1/2}\hat{\beta} + S^{-1/2}(I_p - G)S^{1/2}g$ schreiben lässt, mit
 $S = X'X$ und $\hat{\beta}$ der KQ-Schätzer.

Ist $Ay+a$ zulässig, dann folgt $\text{Cov}(Ay+a) \leq \text{Cov}(\hat{\beta})$. Aber
 $\text{Cov}(Ay+a) \leq \text{Cov}(\hat{\beta})$ bedeutet nicht unbedingt, dass $Ay+a$ für β
 zulässig ist.

Theorem 2: Explizite Voraussetzungen

$\tilde{\beta} = Ay + a$ ist genau dann zulässig, wenn A und a
 $A = (X'X)^{-\frac{1}{2}}G(X'X)^{-\frac{1}{2}}X'$ und $a = (X'X)^{-\frac{1}{2}}(I_p - G)(X'X)^{\frac{1}{2}}g$ gilt,
 wobei $\sigma(G) \subset [0,1]$ und $g \in \mathbb{R}^p$ gilt.

Vom Theorem 2 folgt, dass jeder zulässige Schätzer sich in form:
 $\tilde{\beta} = S^{-1/2}GS^{1/2}\hat{\beta} + S^{-1/2}(I_p - G)S^{1/2}g$ schreiben lässt, mit
 $S = X'X$ und $\hat{\beta}$ der KQ-Schätzer.

Ist $Ay + a$ zulässig, dann folgt $\text{Cov}(Ay + a) \leq \text{Cov}(\hat{\beta})$. Aber
 $\text{Cov}(Ay + a) \leq \text{Cov}(\hat{\beta})$ bedeutet nicht unbedingt, dass $Ay + a$ für β
 zulässig ist.

Beispiel

Beispiel für einen unzulässigen Schätzer für β

$$\hat{\beta}^-(k) = \frac{-1}{1+k} \hat{\beta} \text{ mit } k \geq 0$$

$$\text{Cov}(\hat{\beta}^-(k)) = \frac{\sigma^2}{(1+k)^2} (X'X)^{-1} \text{ und}$$

$$\text{cov}(\hat{\beta}) - \text{cov}(\hat{\beta}^-(k)) = \frac{k(2+k)}{(1+k)^2} (X'X)^{-1} \text{ ist nicht negativ definit.}$$

$$E(\hat{\beta}^-(k)) = \frac{-1}{1+k} \beta$$

$\hat{\beta}^-(k)$ ist aber **unzulässig** da $\sigma(XA) = \{ \frac{-1}{1+k} \} \notin [0,1]$.

Beispiel

Beispiel für einen unzulässigen Schätzer für β

$$\hat{\beta}^-(k) = \frac{-1}{1+k} \hat{\beta} \text{ mit } k \geq 0$$

$$\text{Cov}(\hat{\beta}^-(k)) = \frac{\sigma^2}{(1+k)^2} (X'X)^{-1} \text{ und}$$

$$\text{cov}(\hat{\beta}) - \text{cov}(\hat{\beta}^-(k)) = \frac{k(2+k)}{(1+k)^2} (X'X)^{-1} \text{ ist nicht negativ definit.}$$

$$E(\hat{\beta}^-(k)) = \frac{-1}{1+k} \beta$$

$$\hat{\beta}^-(k) \text{ ist aber } \mathbf{unzulässig} \text{ da } \sigma(XA) = \left\{ \frac{-1}{1+k} \right\} \notin [0,1].$$

Lineare Transformationen des KQ-Schätzers

- $L\hat{\beta}$ ist zulässig genau dann, wenn $X'XL = L'X'X$ und $\sigma(L) \subset [0,1]$ gelten.
- Ist $L = L'$, dann ist $X'XL = L'X'X$ genau dann erfüllt, wenn $X'X = U\Lambda U'$ und $L = U\Gamma U'$ gelten, mit U orthogonal
- $L^{ob} = \{L\hat{\beta} : L = L', X'XL = LX'X, \sigma(L) \subset [0,1]\}$
- L^{ob} ist die **Obenchain Klasse** von Schätzern für β

Lineare Transformationen des KQ-Schätzers

- $L\hat{\beta}$ ist zulässig genau dann, wenn $X'XL = L'X'X$ und $\sigma(L) \subset [0,1]$ gelten.
- Ist $L = L'$, dann ist $X'XL = L'X'X$ genau dann erfüllt, wenn $X'X = U\Lambda U'$ und $L = U\Gamma U'$ gelten, mit U orthogonal
- $L^{ob} = \{L\hat{\beta} : L = L', X'XL = LX'X, \sigma(L) \subset [0,1]\}$
- L^{ob} ist die **Obenchain Klasse** von Schätzern für β

Lineare Transformationen des KQ-Schätzers

- $L\hat{\beta}$ ist zulässig genau dann, wenn $X'XL = L'X'X$ und $\sigma(L) \subset [0,1]$ gelten.
- Ist $L = L'$, dann ist $X'XL = L'X'X$ genau dann erfüllt, wenn $X'X = U\Lambda U'$ und $L = U\Gamma U'$ gelten, mit U orthogonal
- $L^{ob} = \{L\hat{\beta} : L = L', X'XL = LX'X, \sigma(L) \subset [0,1]\}$
- L^{ob} ist die **Obenchain Klasse** von Schätzern für β

Lineare Transformationen des KQ-Schätzers

- **Alle Schätzer aus L^{ob} sind zulässig und es gilt**
$$\|L\hat{\beta}\|^2 \leq \|\hat{\beta}\|^2$$
- Jedes Element aus L^{ob} kann sich unter der zusätzlichen Bedingung $\sigma(L) \subset (0, 1]$ in der Form
$$\tilde{\beta} = (X'X + K)^{-1}X'y = \hat{\beta}_K$$
 darstellen, für eine symmetrisch nichtnegativ definite Matrix K .
- Schätzer der Obenchain Klasse werden bezüglich der Spektralzerlegung
$$X'X = U\Lambda U' = U_1\Lambda_1U_1' + U_2\Lambda_2U_2'$$
 in der Tabelle 1 zusammengestellt.

Lineare Transformationen des KQ-Schätzers

- **Alle Schätzer aus L^{ob} sind zulässig und es gilt**
 $\|L\hat{\beta}\|^2 \leq \|\hat{\beta}\|^2$
- Jedes Element aus L^{ob} kann sich unter der zusätzlichen Bedingung $\sigma(L) \subset (0, 1]$ in der Form
 $\tilde{\beta} = (X'X + K)^{-1}X'y = \hat{\beta}_K$ darstellen, für eine symmetrisch nichtnegativ definite Matrix K .
- Schätzer der Obenchain Klasse werden bezüglich der Spektralzerlegung
 $X'X = U\Lambda U' = U_1\Lambda_1U_1' + U_2\Lambda_2U_2'$ in der Tabelle 1 zusammengestellt.

Lineare Transformationen des KQ-Schätzers

- **Alle Schätzer aus L^{ob} sind zulässig und es gilt**
$$\|L\hat{\beta}\|^2 \leq \|\hat{\beta}\|^2$$
- Jedes Element aus L^{ob} kann sich unter der zusätzlichen Bedingung $\sigma(L) \subset (0, 1]$ in der Form
$$\tilde{\beta} = (X'X + K)^{-1}X'y = \hat{\beta}_K$$
 darstellen, für eine symmetrisch nichtnegativ definite Matrix K .
- Schätzer der Obenchain Klasse werden bezüglich der Spektralzerlegung
$$X'X = U\Lambda U' = U_1\Lambda_1U_1' + U_2\Lambda_2U_2'$$
 in der Tabelle 1 zusammengestellt.

Lineare Transformationen des KQ-Schätzers

Tabelle 1. Schätzer aus der Obenchain Klasse.

Schätzer	$L\hat{\beta} \in L^{ob}(\beta)$
KQ-Schätzer	$L = I_p$
Hauptkomponentenschätzer	$L = U_1 U_1'$
Marquardt	$L = U_1 U_1' + c U_2 U_2'$
Bayes-Parker	$L = U_1 (\Lambda_1 + k I_r)^{-1} \Lambda U_1'$
Immer E-treuer Ridge Schätzer	$L = U (\Lambda^2 + 2k\Lambda) (\Lambda + k I_p)^{-2} U'$
Iterative Schätzer	$L = U (I_p - (I_p - \delta \Lambda)^{m+1}) U'$
Shrinkage Schätzer	$L = (1 + \rho)^{-1} I_p$

Generalisierte Ridge Schätzer

Nicht homogen generalisierte Ridge Schätzer haben die Gestalt:

$$\hat{\beta}_{K, \beta_o} = \beta_o + (X'X + K)^{-1}X'(y - X\beta_o).$$

Theorem 3:

$\hat{\beta}_{K, \beta_o}$ ist zulässig genau dann, wenn K symmetrisch nicht stochastisch und nicht negativ definit ist und β_o beliebig.

$$\hat{\beta}_{K, \beta_o} = Ay + a \text{ mit } A = (X'X + K)^{-1}X' \text{ und } a = \beta_o - (X'X + K)^{-1}X'X\beta_o = (X'X + K)^{-1}K\beta_o.$$

Generalisierte Ridge Schätzer

Nicht homogen generalisierte Ridge Schätzer haben die Gestalt:

$$\hat{\beta}_{K, \beta_o} = \beta_o + (X'X + K)^{-1}X'(y - X\beta_o).$$

Theorem 3:

$\hat{\beta}_{K, \beta_o}$ ist zulässig genau dann, wenn K symmetrisch nicht stochastisch und nicht negativ definit ist und β_o beliebig.

$$\hat{\beta}_{K, \beta_o} = Ay + a \text{ mit } A = (X'X + K)^{-1}X' \text{ und } a = \beta_o - (X'X + K)^{-1}X'X\beta_o = (X'X + K)^{-1}K\beta_o.$$

Generalisierte Ridge Schätzer

Nicht homogen generalisierte Ridge Schätzer haben die Gestalt:

$$\hat{\beta}_{K, \beta_o} = \beta_o + (X'X + K)^{-1}X'(y - X\beta_o).$$

Theorem 3:

$\hat{\beta}_{K, \beta_o}$ ist zulässig genau dann, wenn K symmetrisch nicht stochastisch und nicht negativ definit ist und β_o beliebig.

$$\hat{\beta}_{K, \beta_o} = Ay + a \text{ mit } A = (X'X + K)^{-1}X' \text{ und } a = \beta_o - (X'X + K)^{-1}X'X\beta_o = (X'X + K)^{-1}K\beta_o.$$

Generalisierte Ridge Schätzer

Tabelle 2. Generalisierte Ridge Schätzer.

Schätzer	$\hat{\beta}_{K, \beta_o}$
KQ-Schätzer	$K = 0, \beta_o = 0$
Marquardt	$K = ((1 - c)/c) U_2 \Lambda_2 U_2', \beta_o = 0$
Ridge	$K = kI_p, \beta_o = 0$
Ridge	$K = U \text{diag}(k_1, \dots, k_p) U', \beta_o = 0$
Farebrother	$K = kR'R, \beta_o = R'(RR')^{-1}r$
Immer E-Treuer Ridge	$K = k^2(X'X + 2kI_p)^{-1}, \beta_o = 0$
Iterative Schätzer	$K = X'X(T^{-1} - I_p), \beta_o = 0$ $T = I_p - (I_p - \delta X'X)^{m+1}$
Schrinkage Schätzer	$K = \rho X'X, \beta_o = 0$
Richtungsmodifizierte Ridge	$K = \rho X'X$

Der restringierte KQ-Schätzer

- Annahme: $R\tilde{\beta}=r$ gilt mit $R \neq 0$ und r gegeben.

theorem 4

Wenn die $m \times p$ Matrix R vollen Zeilenrang hat, dann ist der restringierte KQ-Schätzer

$$\hat{\beta}_R = \hat{\beta} - (X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}(R\hat{\beta} - r) \text{ zulässig.}$$

Beweis Idee

$\hat{\beta}_R = Ay + a$ mit $A = (X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}R(X'X)^{-1}X'$ und $a = (X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}r$.

XA ist symmetrisch und idempotent. D.h $XA - XAA'X' = 0$. Es gilt auch $a = (I_p - AX)(R'R)^{-1}r$.

$\hat{\beta}$ ist also nach Theorem 1 zulässig.

Der restringierte KQ-Schätzer

- Annahme: $R\tilde{\beta}=r$ gilt mit $R \neq 0$ und r gegeben.

theorem 4

Wenn die $m \times p$ Matrix R vollen Zeilenrang hat, dann ist der restringierte KQ-Schätzer

$\hat{\beta}_R = \hat{\beta} - (X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}(R\hat{\beta} - r)$ zulässig.

Beweis Idee

$\hat{\beta}_R = Ay + a$ mit $A = (X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}R(X'X)^{-1}X'$ und $a = (X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}r$.

XA ist symmetrisch und idempotent. D.h $XA - XAA'X' = 0$. Es gilt auch $a = (I_p - AX)(R'R)^{-1}r$.

$\hat{\beta}$ ist also nach Theorem 1 zulässig.

Zulässigkeit und Shrinkage Eigenschaften

Theorem 5

Ein linearer homogener Schätzer Ay hat die Shrinkage Eigenschaft $\|Ay\|^2 \leq \|\hat{\beta}\|^2$ genau dann, wenn $A=L(X'X)^{-1}X'$ mit $\sigma(L'L) \subset [0, 1]$ gilt.

- Die homogenen Schätzer, die die Shrinkage Eigenschaft haben, sind nicht immer zulässig
- Umgekehrt zulässige homogene Schätzer müssen nicht die Shrinkage Eigenschaft haben, da $\sigma(L) \subset [0, 1]$ impliziert nicht $\sigma(L'L) \subset [0, 1]$.

Zulässigkeit und Shrinkage Eigenschaften

Theorem 5

Ein linearer homogener Schätzer Ay hat die Shrinkage Eigenschaft $\|Ay\|^2 \leq \|\hat{\beta}\|^2$ genau dann, wenn $A=L(X'X)^{-1}X'$ mit $\sigma(L'L) \subset [0, 1]$ gilt.

- Die homogenen Schätzer, die die Shrinkage Eigenschaft haben, sind nicht immer zulässig
- Umgekehrt zulässige homogene Schätzer müssen nicht die Shrinkage Eigenschaft haben, da $\sigma(L) \subset [0, 1]$ impliziert nicht $\sigma(L'L) \subset [0, 1]$.

Zulässigkeit und Shrinkage Eigenschaften

Theorem 5

Ein linearer homogener Schätzer Ay hat die Shrinkage Eigenschaft $\|Ay\|^2 \leq \|\hat{\beta}\|^2$ genau dann, wenn $A=L(X'X)^{-1}X'$ mit $\sigma(L'L) \subset [0, 1]$ gilt.

- Die homogenen Schätzer, die die Shrinkage Eigenschaft haben, sind nicht immer zulässig
- Umgekehrt zulässige homogene Schätzer müssen nicht die Shrinkage Eigenschaft haben, da $\sigma(L) \subset [0, 1]$ impliziert nicht $\sigma(L'L) \subset [0, 1]$.

Zulässigkeit und Shrinkage Eigenschaften

$\{L\tilde{\beta} : X'XL = LX'X, \sigma(L) \subset [0, 1], \sigma(L'L) \subset [0, 1]\}$ ist die Menge aller linearen homogenen Schätzer, die Shrinkage Eigenschaft haben und zulässig sind.

Nicht alle generalisierte Ridge haben die Shrinkage Eigenschaft.

Zulässigkeit und Shrinkage Eigenschaften

$\{L\tilde{\beta} : X'XL = LX'X, \sigma(L) \subset [0, 1], \sigma(L'L) \subset [0, 1]\}$ ist die Menge aller linearen homogenen Schätzer, die Shrinkage Eigenschaft haben und zulässig sind.

Nicht alle generalisierte Ridge haben die Shrinkage Eigenschaft.

Konvexe Kombination von Schätzern

- $\tilde{\beta}_3 = \alpha\tilde{\beta}_1 + (1 - \alpha)\tilde{\beta}_2$ mit $0 \leq \alpha \leq 1$. $\tilde{\beta}_3$ heißt konvexe Kombination von $\tilde{\beta}_1$ und $\tilde{\beta}_2$.

Theorem 6

$L(\beta, \tilde{\beta}) = (\beta - \tilde{\beta})' W (\beta - \tilde{\beta})$ sei jetzt die Verlustfunktion. Es gilt $L(\beta, \tilde{\beta}_3) \leq \max(L(\beta, \tilde{\beta}_1), L(\beta, \tilde{\beta}_2))$, wobei W eine symmetrische nicht negativ definite Matrix ist.

Konvexe Kombination von Schätzern

- $\tilde{\beta}_3 = \alpha\tilde{\beta}_1 + (1 - \alpha)\tilde{\beta}_2$ mit $0 \leq \alpha \leq 1$. $\tilde{\beta}_3$ heißt konvexe Kombination von $\tilde{\beta}_1$ und $\tilde{\beta}_2$.

Theorem 6

$L(\beta, \tilde{\beta}) = (\beta - \tilde{\beta})' W(\beta - \tilde{\beta})$ sei jetzt die Verlustfunktion. Es gilt $L(\beta, \tilde{\beta}_3) \leq \max(L(\beta, \tilde{\beta}_1), L(\beta, \tilde{\beta}_2))$, wobei W eine symmetrische nicht negativ definite Matrix ist.

Konvexe Kombination von Schätzern

- $\tilde{\beta} = \sum_{i=1}^q \alpha_i \tilde{\beta}_i$, wobei $\sum_{i=1}^q \alpha_i = 1$ und $\alpha_i \geq 0$ gilt.
- Sind die $\tilde{\beta}_i$ zulässig, dann ist $\tilde{\beta}$ auch zulässig.

Beispiel

- $\hat{\beta}_{R,\alpha} = (1 - \alpha)\hat{\beta}_R + \alpha\hat{\beta}$

Konvexe Kombination von Schätzern

- $\tilde{\beta} = \sum_{i=1}^q \alpha_i \tilde{\beta}_i$, wobei $\sum_{i=1}^q \alpha_i = 1$ und $\alpha_i \geq 0$ gilt.
- Sind die $\tilde{\beta}_i$ zulässig, dann ist $\tilde{\beta}$ auch zulässig.

Beispiel

- $\hat{\beta}_{R,\alpha} = (1 - \alpha)\hat{\beta}_R + \alpha\hat{\beta}$

Konvexe Kombination von Schätzern

- $\tilde{\beta} = \sum_{i=1}^q \alpha_i \tilde{\beta}_i$, wobei $\sum_{i=1}^q \alpha_i = 1$ und $\alpha_i \geq 0$ gilt.
- Sind die $\tilde{\beta}_i$ zulässig, dann ist $\tilde{\beta}$ auch zulässig.

Beispiel

- $\hat{\beta}_{R,\alpha} = (1 - \alpha)\hat{\beta}_R + \alpha\hat{\beta}$

Lineare Bayes Schätzer

Die Schätzer der Form $\hat{\beta}^B = \mu + T[T + (X'X)^{-1}](\hat{\beta} - \mu)$ können verwendet werden, wenn β die Realisation eines Zufallvektors b mit $E(b)=\mu$ und $Cov(b)=\sigma^2 T$ ist.

- $Ay+a$ ist linearer Bayes Schätzer genau dann, wenn $A=(X'X)^{-1/2}G(X'X)^{-1/2}X'$ und $a=(X'X)^{-1/2}(I_p - G)(X'X)^{1/2}g$ für eine symmetrische $p \times p$ Matrix G mit $\sigma(G) \subset [0, 1)$
- Die linearen Bayes Schätzer sind zulässig für β und stimmen unter $\sigma(G) \subset [0, 1)$ mit den linearen zulässigen Schätzern überein.

Lineare Bayes Schätzer

Die Schätzer der Form $\hat{\beta}^B = \mu + T[T + (X'X)^{-1}](\hat{\beta} - \mu)$ können verwendet werden, wenn β die Realisation eines Zufallvektors b mit $E(b)=\mu$ und $Cov(b)=\sigma^2 T$ ist.

- $Ay+a$ ist linearer Bayes Schätzer genau dann, wenn $A=(X'X)^{-1/2}G(X'X)^{-1/2}X'$ und $a=(X'X)^{-1/2}(I_p - G)(X'X)^{1/2}g$ für eine symmetrische $p \times p$ Matrix G mit $\sigma(G) \subset [0, 1)$
- Die linearen Bayes Schätzer sind zulässig für β und stimmen unter $\sigma(G) \subset [0, 1)$ mit den linearen zulässigen Schätzern überein.

Lineare Bayes Schätzer

Die Schätzer der Form $\hat{\beta}^B = \mu + T[T + (X'X)^{-1}](\hat{\beta} - \mu)$ können verwendet werden, wenn β die Realisation eines Zufallvektors b mit $E(b)=\mu$ und $Cov(b)=\sigma^2 T$ ist.

- $Ay+a$ ist linearer Bayes Schätzer genau dann, wenn $A=(X'X)^{-1/2}G(X'X)^{-1/2}X'$ und $a=(X'X)^{-1/2}(I_p - G)(X'X)^{1/2}g$ für eine symmetrische $p \times p$ Matrix G mit $\sigma(G) \subset [0, 1)$
- Die linearen Bayes Schätzer sind zulässig für β und stimmen unter $\sigma(G) \subset [0, 1)$ mit den linearen zulässigen Schätzern überein.

Zulässigkeit unter elliptischer Restriktion

- $\beta' T \beta \leq \sigma^2$, wobei T eine nicht negativ definite Matrix ist und $\sigma^2 \in (0, \infty)$ beliebig aber fest.

- Wenn $(\beta - \beta_o)' T (\beta - \beta_o) \leq \sigma^2$ vorliegt, betrachtet man das Modell:

$$\bar{y} = X\bar{\beta} - \epsilon, \bar{y} = y - X\beta_o, \bar{\beta} = \beta - \beta_o \text{ und } \bar{\beta}' T \bar{\beta} \leq \sigma^2.$$

$A\bar{y}$ für $\bar{\beta}$ im neuen Modell entspricht dem Schätzer $A\bar{y} + \beta_o$ im alten Modell.

Theorem 7

- Unter der Restriktion $\beta' T \beta \leq \sigma^2$, ist ein Schätzer Ay für β zulässig genau dann, wenn $XA = A' X'$, $\sigma(XA) \subset [0, 1)$ und $tr[(I_p - AX)^{-1}(X'X)^{-1}T] \leq 1 + tr[(X'X)^{-1}T]$

Zulässigkeit unter elliptischer Restriktion

- $\beta' T \beta \leq \sigma^2$, wobei T eine nicht negativ definite Matrix ist und $\sigma^2 \in (0, \infty)$ beliebig aber fest.

- Wenn $(\beta - \beta_o)' T (\beta - \beta_o) \leq \sigma^2$ vorliegt, betrachtet man das Modell:

$$\bar{y} = X\bar{\beta} - \epsilon, \bar{y} = y - X\beta_o, \bar{\beta} = \beta - \beta_o \text{ und } \bar{\beta}' T \bar{\beta} \leq \sigma^2.$$

$A\bar{y}$ für $\bar{\beta}$ im neuen Modell entspricht dem Schätzer $A\bar{y} + \beta_o$ im alten Modell.

Theorem 7

- Unter der Restriktion $\beta' T \beta \leq \sigma^2$, ist ein Schätzer Ay für β zulässig genau dann, wenn $XA = A'X'$, $\sigma(XA) \subset [0, 1)$ und $tr[(I_p - AX)^{-1}(X'X)^{-1}T] \leq 1 + tr[(X'X)^{-1}T]$

Zulässigkeit unter elliptischer Restriktion

- $\beta' T \beta \leq \sigma^2$, wobei T eine nicht negativ definite Matrix ist und $\sigma^2 \in (0, \infty)$ beliebig aber fest.

- Wenn $(\beta - \beta_o)' T (\beta - \beta_o) \leq \sigma^2$ vorliegt, betrachtet man das Modell:

$$\bar{y} = X\bar{\beta} - \epsilon, \bar{y} = y - X\beta_o, \bar{\beta} = \beta - \beta_o \text{ und } \bar{\beta}' T \bar{\beta} \leq \sigma^2.$$

$A\bar{y}$ für $\bar{\beta}$ im neuen Modell entspricht dem Schätzer $A\bar{y} + \beta_o$ im alten Modell.

Theorem 7

- Unter der Restriktion $\beta' T \beta \leq \sigma^2$, ist ein Schätzer Ay für β zulässig genau dann, wenn $XA = A'X'$, $\sigma(XA) \subset [0, 1)$ und $tr[(I_p - AX)^{-1}(X'X)^{-1}T] \leq 1 + tr[(X'X)^{-1}T]$

Zulässigkeit unter elliptischer Restriktion

- Jeder zulässige Schätzer für β sich in folgender Form schreiben lässt: $\tilde{\beta} = M(M + (X'X)^{-1})^{-1}\hat{\beta}$
- Die Menge der linearen zulässigen Schätzer stimmt mit der Menge der Bayes-Schätzer überein für jede kompakte Teilmenge von \mathbb{R}^p Parameterraum.
- Der generalisierte Ridge Schätzer $\hat{\beta}_K = (X'X + K)^{-1}X'y$ ist genau dann zulässig, wenn $\text{tr}(K^{-1}T) \leq 1$ ist.

Literatur

Groß, J., (2003): Linear Regression. Springer-Verlag, Berlin Heidelberg New York.

Vielen Dank für ihre Aufmerksamkeit