

Kapitel 3

Das multiple Regressionsmodell

Zusätzliche Regressoren

Vorteil 1: größerer Erklärungsgehalt

z.B. Erfahrung & Bildung $\rightarrow$ Lohn

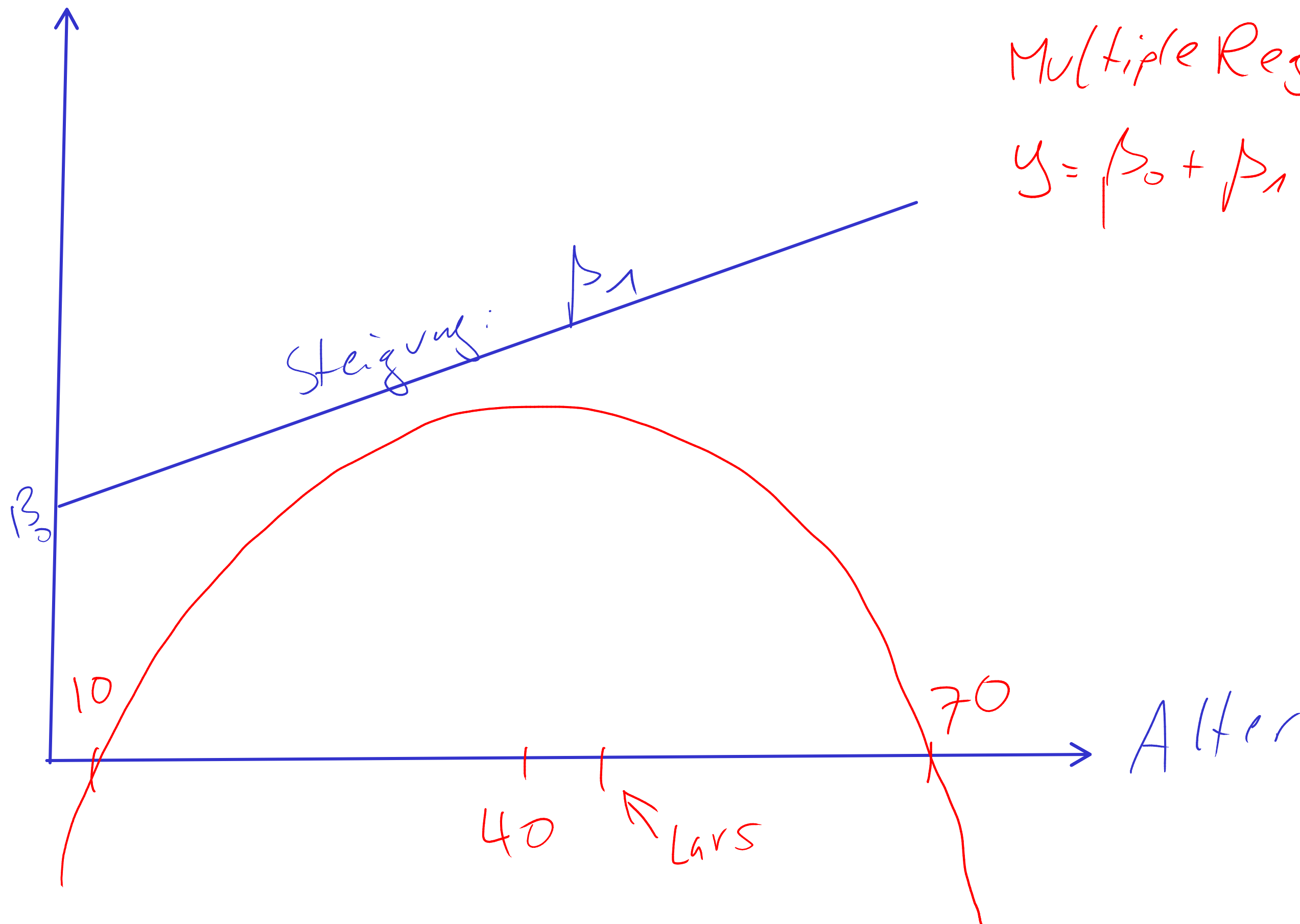
Vorteil 2: nicht-lineare Darstellung
möglich

Produktivität

Einfachregression

Multiple Regression

$$y = \beta_0 + \beta_1 x + \beta_2 x^2$$



Vorteil 3: Annahme der strikten Exogenität
lässt sich besser stützen

Nachteil: Rechenaufwand steigt bei
Summennotation erheblich

→ Wechsel auf Matrix Notation

kleinste Quadrate bei zwei Regressoren und Summennotation

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 z_i + v_i$$

Diagram illustrating the components of the regression equation:

- y_i is labeled **Regressand** (Dependent Variable).
- $\beta_0, \beta_1, \beta_2$ are collectively labeled **Parameter**.
- x_i, z_i are collectively labeled **Regressoren** (Independent Variables).
- v_i is labeled **Störterm** (Error Term).

Abkürzungen

$$\overline{ab} = \frac{1}{n} \sum_{i=1}^n a_i \cdot b_i$$

$$\bar{a} = \frac{1}{n} \sum_{i=1}^n a_i$$

$$\overline{aa} = \frac{1}{n} \sum_{i=1}^n a_i \cdot a_i$$

$$S_{aa} = \frac{1}{n} \sum_{i=1}^n (a_i - \bar{a})^2 = \overline{aa} - \bar{a} \cdot \bar{a}$$

$$S_{ab} = \overline{ab} - \bar{a} \cdot \bar{b}$$

Bedingungen

$$S_{xx} > 0, \quad S_{zz} > 0$$

$$|r_{xz}| \neq 1$$

$$r_{xz} = \frac{S_{xz}}{\sqrt{S_{xx}} \cdot \sqrt{S_{zz}}}$$

Prognose

Vermutungen über $\beta_0, \beta_1, \beta_2$: b_0, b_1, b_2

$$\hat{y}_i = b_0 + b_1 x_i + b_2 z_i$$

Prognosefehler (Residuum)

$$\hat{U}_i = y_i - \hat{y}_i = y_i - b_0 - b_1 x_i - b_2 z_i$$

Summe der quadrierten Residuen

$$S(b_0, b_1, b_2) = \sum_{i=1}^n \hat{U}_i^2 = \sum_{i=1}^n (y_i - b_0 - b_1 x_i - b_2 z_i)^2$$

„kleinste Quadrate“

$$\min_{b_0, b_1, b_2} S(b_0, b_1, b_2)$$

$$S(b_0, b_1, b_2) = \sum_{i=1}^n (y_i - b_0 - b_1 x_i - b_2 z_i)^2$$

$$\frac{\partial S(b_0, b_1, b_2)}{\partial b_0} = \sum_{i=1}^n -1 \cdot 2 \cdot (y_i - b_0 - b_1 x_i - b_2 z_i) \stackrel{!}{=} 0$$

$$\frac{\partial S(b_0, b_1, b_2)}{\partial b_1} = \sum_{i=1}^n -x_i \cdot 2 \cdot (y_i - b_0 - b_1 x_i - b_2 z_i) \stackrel{!}{=} 0$$

$$\frac{\partial S(b_0, b_1, b_2)}{\partial b_2} = \sum_{i=1}^n -z_i \cdot 2 \cdot (y_i - b_0 - b_1 x_i - b_2 z_i) \stackrel{!}{=} 0$$

$$\sum_{i=1}^n -1 \cdot 2 \cdot (y_i - b_0 - b_1 x_i - b_2 z_i) \stackrel{!}{=} 0 \quad | : -2$$

$$\sum_{i=1}^n -x_i \cdot 2 (y_i - b_0 - b_1 x_i - b_2 z_i) \stackrel{!}{=} 0 \quad | : -2$$

$$\sum_{i=1}^n -z_i \cdot 2 (y_i - b_0 - b_1 x_i - b_2 z_i) \stackrel{!}{=} 0 \quad | : -2$$

$$\sum_{i=1}^n y_i - b_0 - b_1 x_i - b_2 z_i = 0$$

$$\sum_{i=1}^n x_i (y_i - b_0 - b_1 x_i - b_2 z_i) = 0$$

$$\sum_{i=1}^n z_i (y_i - b_0 - b_1 x_i - b_2 z_i) = 0$$

$$\sum_{i=1}^n y_i - b_0 - b_1 x_i - b_2 z_i = 0$$

$$\sum_{i=1}^n x_i (y_i - b_0 - b_1 x_i - b_2 z_i) = 0$$

$$\sum_{i=1}^n z_i (y_i - b_0 - b_1 x_i - b_2 z_i) = 0$$

$$\sum_{i=1}^n y_i - \sum_{i=1}^n b_0 - \sum_{i=1}^n b_1 x_i - \sum_{i=1}^n b_2 z_i = 0 \quad | : n$$

$$\bar{y} - b_0 - b_1 \bar{x} - b_2 \bar{z} = 0$$

$$\sum_{i=1}^n x_i y_i - b_0 \sum_{i=1}^n x_i - b_1 \sum_{i=1}^n x_i x_i - b_2 \sum_{i=1}^n x_i z_i = 0 \quad | : n$$

$$\bar{x} \bar{y} - b_0 \bar{x} - b_1 \bar{x} \bar{x} - b_2 \bar{x} \bar{z} = 0$$

$$\sum_{i=1}^n z_i y_i - b_0 \sum_{i=1}^n z_i - b_1 \sum_{i=1}^n z_i x_i - b_2 \sum_{i=1}^n z_i z_i = 0 \quad | : n$$

$$\bar{z} \bar{y} - b_0 \bar{z} - b_1 \bar{x} \bar{z} - b_2 \bar{z} \bar{z} = 0$$

$$b_0 = \bar{y} - b_1 \bar{x} - b_2 \bar{z}$$

$$\bar{y} = b_0 + b_1 \bar{x} + b_2 \bar{z}$$

$$x\bar{y} = b_0 \bar{x} + b_1 \bar{x}\bar{x} + b_2 \bar{x}\bar{z}$$

$$\cancel{\bar{z}y} = b_0 \bar{z} + b_1 \bar{x}\bar{z} + b_2 \bar{z}\bar{z}$$

Lösung "KQ-Schätzer"

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} - \hat{\beta}_2 \bar{z}$$

$$\hat{\beta}_1 = (S_{zz} \cdot S_{xy} - S_{xz} S_{zy}) / (S_{xx} S_{zz} - S_{xz} \cdot S_{xz})$$

$$\hat{\beta}_2 = (S_{xx} S_{zy} - S_{xz} S_{xy}) / (S_{xx} S_{zz} - S_{xz} S_{xz})$$

Matrizen und Vektoren

Eine Matrix der Ordnung $m \times n$
ist eine Tabelle mit m Zeilen
und n Spalten.

Die Einträge der Matrix sind Zahlen.

nur eine Zeile ($m=1$) $\rightarrow$ Zeilenvektor

nur eine Spalte ($n=1$) $\rightarrow$ Spaltenvektor

nur eine Zeile und eine Spalte $\rightarrow$ Skalar

Matrix ist quadratisch, falls $m=n$

$$A_{m \times n} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

a_{ij} Element der i -ten Zeile
 und j -ten Spalte

Transponierte Matrix

$$A' = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

(Note: In the original image, the element a_{m1} in the matrix above is circled in red, and a red arrow points from it to the a_{m1} element in the matrix below.)

$$= \begin{pmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{pmatrix}$$

Eine Matrix A ist symmetrisch, falls

$$A' = A$$

→ A muss quadratisch sein

Besondere Matrizen

Einheitsmatrix I_n
 $n \times n$

$$= \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

Nullmatrix N
 $m \times n$

$$= \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}$$

Einsvektor

$$L = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

iota

Addition von Matrizen

für A, B gilt
 $m \times n$ $m \times n$

$$C = A + B = B + A$$

$m \times n$

$$c_{ij} = a_{ij} + b_{ij}$$

Multiplikation mit einem Skalar

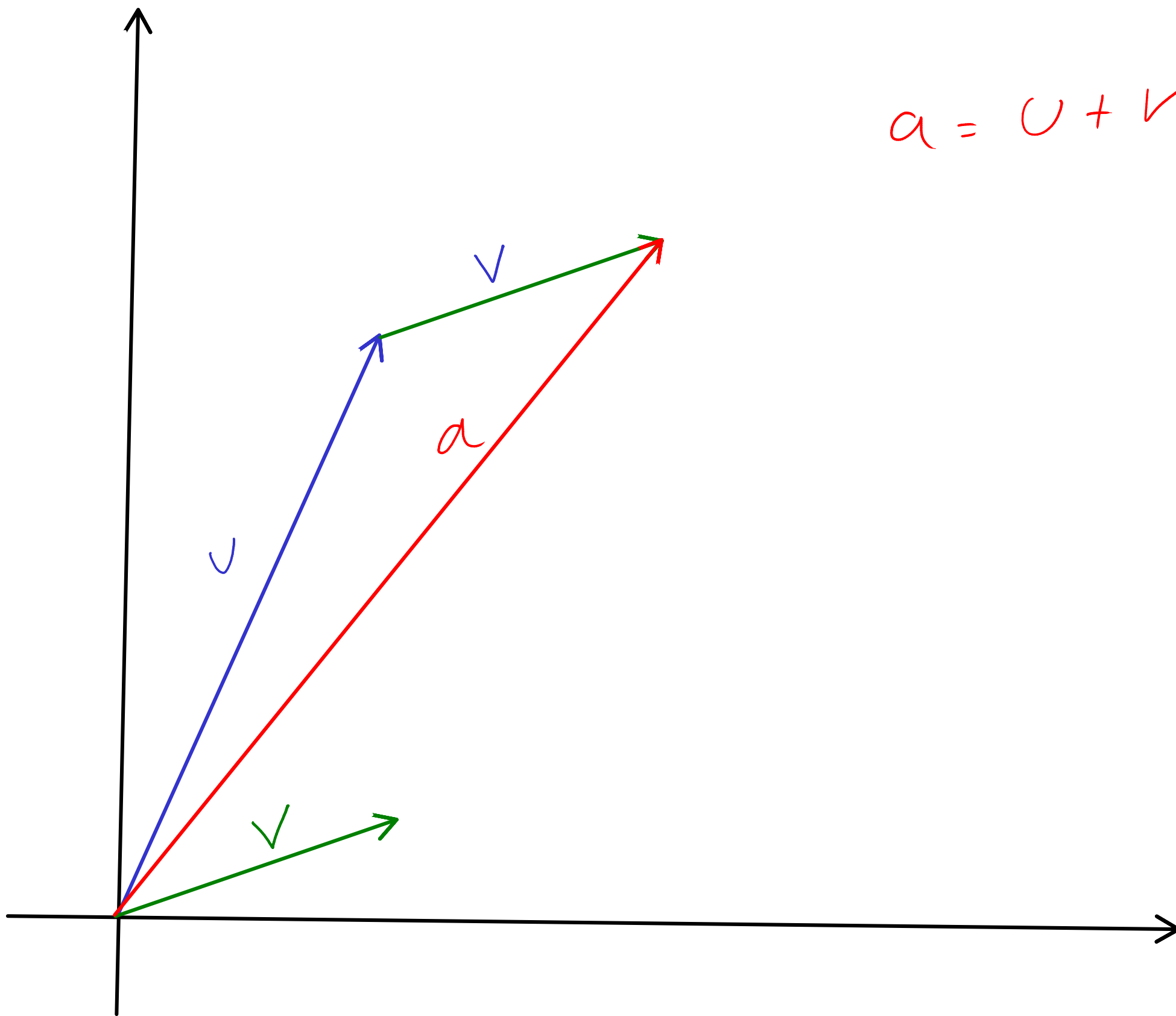
$$b \in \mathbb{R}, A \in \mathbb{R}^{m \times n}$$

$$C = b \cdot A = A \cdot b \quad \text{mit } c_{ij} = a_{ij} \cdot b$$

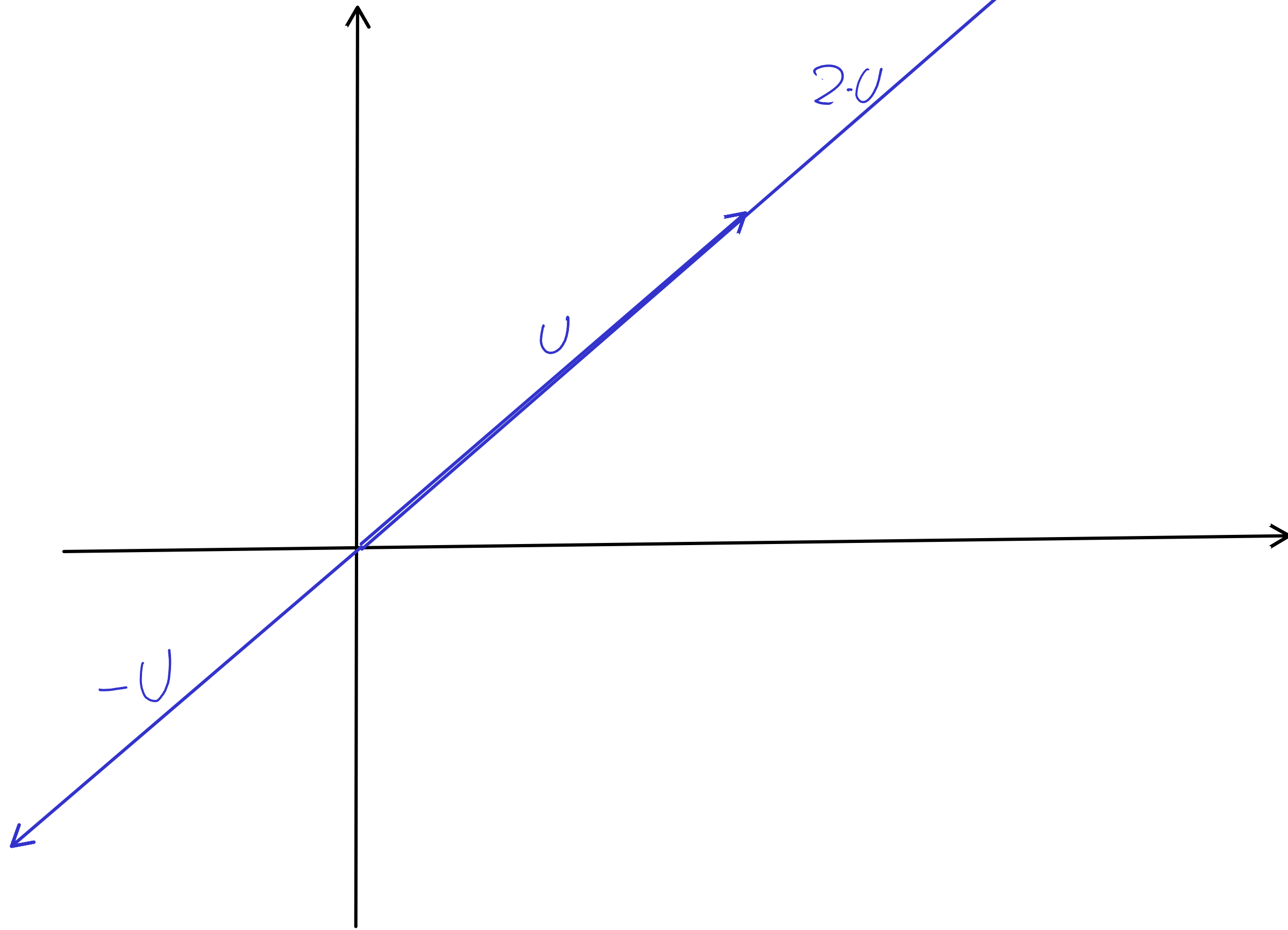
$m \times n$

Linear kombinationen (Addition)

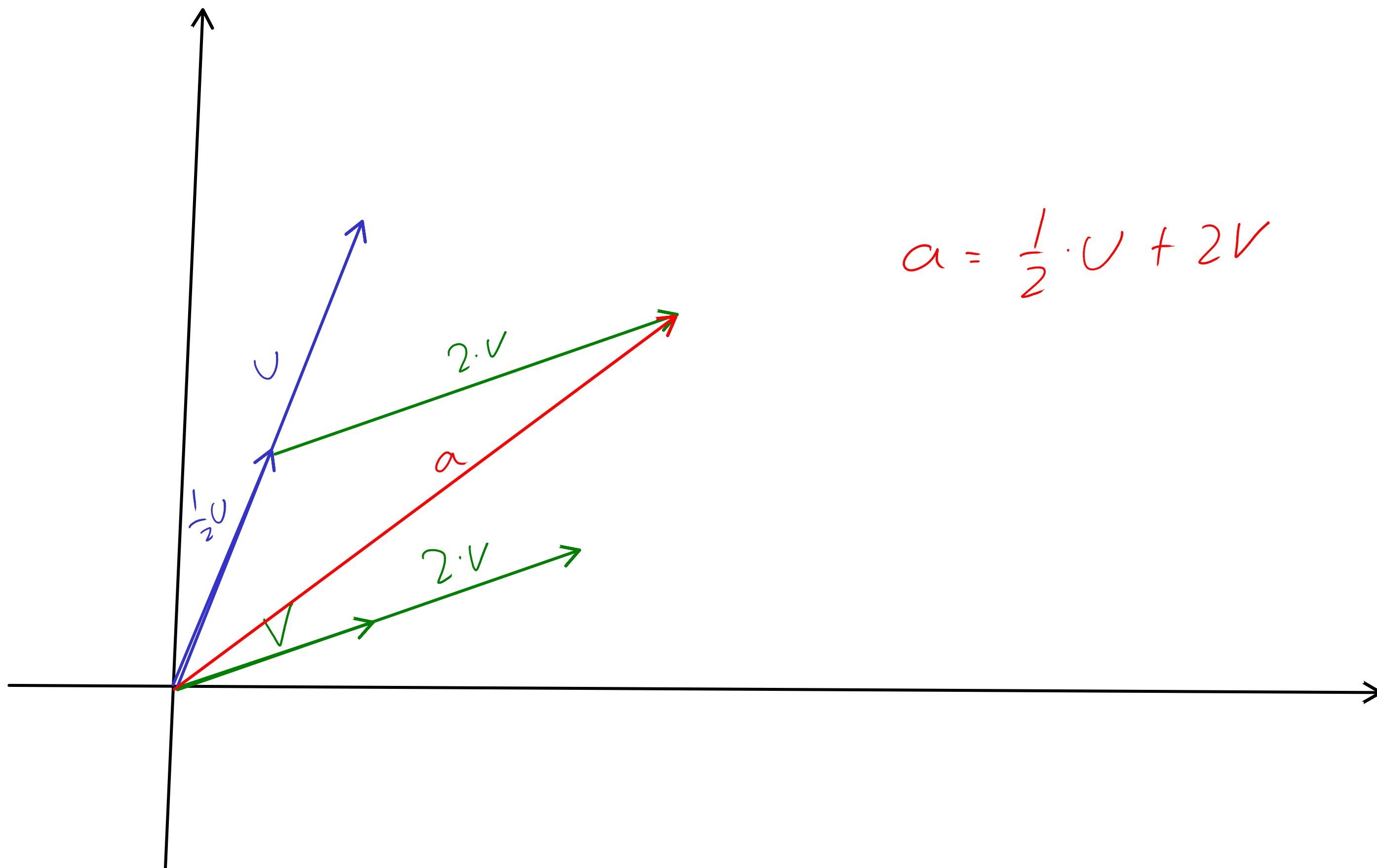
$$a = u + v$$



Linearkombination (Multiplikation mit Skalar)



Linearkombination



$$a = \frac{1}{2} \cdot u + 2v$$

Linear Kombination (allgemein)

Gegeben seien $v, w \in \mathbb{R}^n$, $a, b \in \mathbb{R}$

Dann heißt $v \cdot a + w \cdot b$ Linear Kombination

$$v \cdot a + w \cdot b = \begin{pmatrix} v_1 a + w_1 b \\ v_2 a + w_2 b \\ \vdots \\ v_n a + w_n b \end{pmatrix}$$

Skalarprodukt

für zwei Vektoren $v, w \in \mathbb{R}^n$

gilt
$$v \cdot w = \sum_{i=1}^n v_i \cdot w_i$$

$$v \cdot v = \sum_{i=1}^n \underbrace{v_i^2}_{\geq 0} \geq 0$$

$$= 0 \Leftrightarrow v_i = 0 \quad i = 1, \dots, n$$

Norm eines Vektors

$$\|v\| = \sqrt{v \cdot v} = \sqrt{\sum_{i=1}^n v_i^2}$$

Multiplikation zweier Matrizen

$$\begin{matrix} A & B & = & C \\ \underline{m \times n} & \underline{n \times o} & & \underline{m \times o} \end{matrix}$$

$$C_{ij} = \sum_{t=1}^n a_{it} b_{tj}$$

Skalarprodukt der
 i -ten Zeile von A mit der
 j -ten Spalte von B .

$$A \cdot B \neq B \cdot A$$

Spezialfall: B : Spaltenvektor

$$\rightarrow V = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

$$A \cdot V = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & & a_{mn} \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

$m \times n$ $n \times 1$

$$= \begin{pmatrix} a_{11} \cdot v_1 + \dots + a_{1n} \cdot v_n \\ \vdots \\ a_{m1} \cdot v_1 + \dots + a_{mn} \cdot v_n \end{pmatrix} = \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} v_1 + \dots + \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} v_n$$

Spezialfall Multiplikation einer Matrix
mit Zeilenvektor von links

$$\underbrace{(v_1, \dots, v_m)}_{\substack{1 \times m}} \underbrace{\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}}_{\substack{m \times n}} = \underbrace{\hspace{10em}}_{1 \times n}$$

1.4 Stochastische Matrizen & Vektoren

Erwartungswert

a_{ij} Zufallsvariable mit Dichte f_{ij}

$$E[a_{ij}] = \int_{-\infty}^{\infty} a_{ij} \cdot f_{ij}(a_{ij}) da_{ij}$$

Rechenregeln:

$$E[b \cdot a_{ij} + c] = b \cdot E[a_{ij}] + c$$

$$A = \begin{pmatrix} a_{11} & a_{1n} \\ a_{m1} & a_{mn} \end{pmatrix}$$

Zufallsmatrix

$$E[A] = \begin{pmatrix} E[a_{11}] & E[a_{1n}] \\ E[a_{m1}] & E[a_{mn}] \end{pmatrix}$$

B, C fest

$$\underbrace{E[B \cdot A]}_{0 \times n} + \underbrace{C}_{0 \times n} = B E[A] + C$$

Bedingter Erwartungswert

Zufallsvektor v

Matrix A

$E[v | A]$ bedingter E-Wert von v
(bedingt auf A)

$$E[A \cdot v | A] = A \cdot E[v | A]$$

Teleskop eigenschaft

$$E_A[E_v[v | A]] = E_v[v]$$

Varianzen & Kovarianzen von Zufallsvektoren

Zufallsvariablen v_i, w_j

$$\text{Cor}(v_i, w_j) = E[(v_i - E[v_i])(w_j - E[w_j])]$$

$$\text{Var}(v_i) = \text{Cor}(v_i, v_i) = E[(v_i - E[v_i])^2]$$

Rechenregeln

$$\text{Cor}(av_i + b, cw_j + d) = \text{Cor}(av_i, cw_j)$$

$$= a \cdot c \cdot \text{Cor}(v_i, w_j)$$

$$\text{Var}(av_i + b) = a^2 \text{Var}(v_i)$$

Zufallsvektoren $V = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$, $w = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$

Varianz-Kovarianz-Matrix von V & w .

$$\Sigma(V, w) = E \left[\underbrace{(V - E[V])}_{n \times 1} \underbrace{(w - E[w])' }_{1 \times n} \right]$$

$n \times n$

$$= E \left[\begin{pmatrix} v_1 - E[v_1] \\ \vdots \\ v_n - E[v_n] \end{pmatrix} (w_1 - E[w_1], \dots, w_n - E[w_n]) \right]$$

$$= E \left[\begin{array}{c} (v_1 - E[v_1])(w_1 - E[w_1]) \\ (v_n - E[v_n])(w_1 - E[w_1]) \\ (v_n - E[v_n])(w_n - E[w_n]) \end{array} \right]$$

$$= \begin{pmatrix} \text{Cor}(v_1, w_1) & \text{Cor}(v_1, w_n) \\ \text{Cor}(v_n, w_1) & \text{Cor}(v_n, w_n) \end{pmatrix}$$

$$\Sigma(v, v) = \begin{pmatrix} \text{Cor}(V_1, V_1) & & \text{Cor}(V_1, V_n) \\ & & \\ \text{Cor}(V_n, V_1) & & \text{Cor}(V_n, V_n) \end{pmatrix}$$

Varianz - Kovarianz - Matrix

$$\Sigma(v) = \begin{pmatrix} \text{Var}(V_1) & \text{Cor}(V_1, V_2) & \dots & \text{Cor}(V_1, V_n) \\ \text{Cor}(V_2, V_1) & \text{Var}(V_2) & \dots & \text{Cor}(V_2, V_n) \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cor}(V_n, V_1) & \text{Cor}(V_n, V_2) & \dots & \text{Var}(V_n) \end{pmatrix}$$

$\Sigma(v)$ ist symmetrisch

Rechenregeln

$$\nabla (Av + a, Bw + b) = A \nabla(v, w) \cdot B'$$


$$\nabla(Av + a) = A \nabla(v) \cdot A'$$

1.5 Das lineare Regressionsmodell in Matrixnotation

$$y_i = \beta_0 \cdot 1 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + u_i \quad i=1, \dots, n$$

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

"Regressanden"

Teil einer Zufallsstichprobe
y beobachtbar.

"iota"

$$1 = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

konstante

$$X_j = \begin{pmatrix} x_{1j} \\ x_{2j} \\ \vdots \\ x_{nj} \end{pmatrix}$$

Regressor j

$$j = 1, \dots, k < n$$

Teil von Zufallsstichprobe

beobachtbar

→ erklären y

$$\beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{pmatrix}$$

Parametervektor

linear im Regressionsmodell

nicht-stochastisch (fest)

unbekannt

unspezifiziert

$$U = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}$$

Störterm

zufällig

unbeobachtbar

$$y_1 = \beta_0 \cdot 1 + \beta_1 x_{11} + \beta_2 x_{12} + \dots + \beta_k x_{1k} + u_1$$

⋮

$$y_n = \beta_0 \cdot 1 + \beta_1 x_{n1} + \beta_2 x_{n2} + \dots + \beta_k x_{nk} + u_n$$

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \beta_0 \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} + \beta_1 \begin{pmatrix} x_{11} \\ \vdots \\ x_{n1} \end{pmatrix} + \beta_2 \begin{pmatrix} x_{12} \\ \vdots \\ x_{n2} \end{pmatrix} + \dots + \beta_k \begin{pmatrix} x_{1k} \\ \vdots \\ x_{nk} \end{pmatrix} + \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$$

Regressor matrix

$$X = \begin{pmatrix} 1 & x_{11} & x_{12} & \dots & x_{1k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nk} \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & x_{1k} \\ \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{nk} \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_k \end{pmatrix} + \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$$

$$\Leftrightarrow y = X \cdot \beta + u$$

$n \times k+1$

Lineare Unabhängigkeit von Vektoren

Eine Gruppe von Vektoren heißt linear unabhängig (l.u.)

falls, keiner dieser Vektoren durch eine Linear kombination der anderen Vektoren dargestellt werden kann.

Die Spalten der Matrix A sind l.u. falls

$$A \cdot v = 0 \Rightarrow v = 0$$

$$v \neq 0 \Rightarrow A \cdot v \neq 0$$

1.7 Modellannahmen

MLR 1 $y = \underline{X} \beta + u$

Linearität

y ist linear in den Parametern

$$\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_k \end{pmatrix}$$

$$y \in \mathbb{R}^n, \underline{X} \in \mathbb{R}^{n \times k+1}, \beta \in \mathbb{R}^{k+1}, u \in \mathbb{R}^n$$

β unbekannt, unspezifiziert

MLR 2
Zufallsstichprobe

$y, \underline{X}$ werden zufällig erhoben

$y, \underline{X}$ beobachtbar

MLR 3
voller Spaltenrang

Alle spalten von $\underline{X}$ linear unabhängig

$$\text{rk}(\underline{X}) = k+1$$

MLR 4
Strikte Exogenität

$$E[u | X] \underset{\text{Annahme}}{=} E[u] = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

ohne Beschränkung
der Allgemeinheit

MLR 5
Homoskedastizität
Serielle Unkorreliertheit

$$\Sigma(u) = \begin{pmatrix} \sigma^2 & 0 & \dots & 0 \\ 0 & \sigma^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma^2 \end{pmatrix} = \sigma^2 \cdot I_n$$

→ $\text{Var}(u_i) = \sigma^2$ für alle $i=1, \dots, n$

→ $\text{Cov}(u_i, u_j) = 0$ für $i \neq j$

MLR 6

$$u \sim \mathcal{N}(0, \sigma^2 \cdot I_n)$$

Störterme normalverteilt

Anmerkungen zu den Annahmen

MLR 1 linear in den Parametern

Regressoren dürfen nicht-linear sein

z.B. Alter, Alter²

MLR 2 zufallsstichprobe

ex ante: X zufällig

Stichprobe wurde noch nicht erhoben

$$\rightarrow E[\hat{\beta}]$$

ex post: X fest

Stichprobe wurde bereits beobachtet

$$\rightarrow E[\hat{\beta} | X]$$

MLR 3 Regressormatrix $\underline{X}$ hat
vollen Spaltenrang

→ Alle Spalten von $\underline{X}$ lin. unabh.

Angenommen letzte Spalte k könnte durch
eine lineare Kombination der anderen Spalten
dargestellt werden.

Alle Infos der letzten Spalte sind dann
in den übrigen Spalten enthalten; wir können
die letzte Spalte entfernen.

MLR4 $E[u | X] = E[u] \stackrel{(*)}{=} \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$

Strikte Exogenität

(*) ist ohne Beschränkung der Allgemeinheit

Verletzung der Modellannahmen

MLR 1 $y = \beta_0 + \beta_1 x + 1 \cdot z + u$

Verletzung: $\beta_2 = 1$ spezifiziert
 im Vorhinein müssen alle Werte für die Parameter erlaubt sein.

MCR 1

$$y = \beta_0 (1 + \beta_1 x) + u$$

$$= \beta_0 + \beta_0 \cdot \beta_1 x + u$$

nicht-linear!

$$= \tilde{\beta}_0 + \tilde{\beta}_1 x + u$$

$$y = \beta_0 \cdot x^{\beta_1} z^{\beta_2} \cdot u$$

nicht-linear

$$\begin{aligned} \ln(y) &= \ln(\beta_0 \cdot x^{\beta_1} \cdot z^{\beta_2} \cdot u) \\ &= \ln(\beta_0) + \ln(x^{\beta_1}) + \ln(z^{\beta_2}) + \ln(u) \\ &= \underbrace{\ln(\beta_0)}_{\tilde{\beta}_0} + \beta_1 \ln(x) + \beta_2 \ln(z) + \underbrace{\ln(u)}_{\tilde{u}} \end{aligned}$$

MLR 3

$$x_i = \begin{cases} 1 & , \text{ falls } i: \text{Katze} \\ 0 & , \text{ falls } i: \text{keine Katze} \end{cases}$$

$$z_i = \begin{cases} 0 & , \text{ falls } i: \text{Katze} \\ 1 & , \text{ falls } i: \text{keine Katze} \end{cases}$$

$$y = \beta_0 L + \beta_1 x + \beta_2 z + u$$

$$x + z = L \quad \Rightarrow \quad \text{rk}(L, x, z) = 2 < 3$$

MLR 4

Wahrer Zusammenhang Berufserfahrung

Lohn $\nearrow$

$$y = \beta_0 + \beta_1 \underset{\substack{\uparrow \\ \text{Bildung}}}{x} + \beta_2 \overset{\nwarrow}{z} + u$$

geschätztes Modell:

$$y = \beta_0 + \beta_1 x + \tilde{u}$$

plausibel: $\text{Cor}(x, z) \neq 0$

$$E[z|x] \neq E[z]$$

mit $\tilde{u} = \beta_2 z + u \Rightarrow E[\tilde{u}|x] = E[\beta_2 z + u|x] = \beta_2 \overbrace{E[z|x]}^{+ E[z]} + E[u|x]$

$$\neq E[\beta_2 z + u] = E[\tilde{u}]$$

$$MLR 5 \quad \text{Var}(U_i) = \sigma^2 \quad \text{für } i=1, \dots, n$$

$$\text{Cor}(U_i, U_j) = 0 \quad \text{für } i \neq j$$

$$\Sigma(u) = \sigma^2 \cdot I_n$$

$$y = \beta_0 + \beta_1 \underset{\uparrow}{x} + u$$

↗

Konsum

Einkommen

je größer x , desto größer $\text{Var}(y|x)$

$$\text{Var}(y|x) = \text{Var}(\beta_0 + \beta_1 x + u|x) = \text{Var}(u|x)$$

→ desto größer $\text{Var}(u|x)$

$$\text{MLR 6} : U \sim \mathcal{N} \left(\underset{\substack{\uparrow \\ \text{MLR 4}}}{\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}}, \underset{\substack{\uparrow \\ \text{MLR 5}}}{\sigma^2 \cdot I_n} \right)$$

Unbeobachtbare Faktoren:

z.B. Haushaltsgröße

Kinder

Monate Arbeitslosigkeit

→ ganze Zahlen ≥ 0

Normalverteilung lässt aber reelle Zahlen zu.
 „Lösung“ sehr große Stichproben

Transponieren von Matrixprodukten

$$\underbrace{A \cdot B}_{m \times o}$$

A is $m \times n$, B is $n \times o$

$$(A \cdot B)' = \underbrace{B' \cdot A'}_{o \times m}$$

B' is $o \times n$, A' is $n \times m$

2 Schätzer für β

2.1 Lineare Funktion

$$f: \mathbb{R}^n \rightarrow \mathbb{R}^m$$

Additivität: $f(x_1 + x_2) = f(x_1) + f(x_2)$



Homogenität: $f(t \cdot x_1) = t \cdot f(x_1) \quad t \in \mathbb{R}$

Matrix $A_{m \times n} \Rightarrow f(x) = A \cdot x$ Linear

2.2. Jacobi-Matrix:

(Matrix der 1. Ableitungen)

$$J_f(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \dots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \dots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}$$

Falls $f(x) = A \cdot x$: $J_f(x) = A'$

2.3. Quadratische Formen und positiv Definitheit

$$f: \mathbb{R}^n \rightarrow \mathbb{R}$$

Zeilenvektor

$$f(x) = x^T \cdot A \cdot x$$

Spaltenvektor

Beispiel

$$\text{für } A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

$$(x_1, x_2) \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$\begin{aligned} &= (x_1, x_2) \begin{pmatrix} ax_1 + bx_2 \\ cx_1 + dx_2 \end{pmatrix} = \underbrace{(x_1, x_2)}_{2 \times 2} \underbrace{\begin{pmatrix} ax_1 + bx_2 \\ cx_1 + dx_2 \end{pmatrix}}_{2 \times 1} \\ &= ax_1^2 + (b+c) \cdot x_1 x_2 + dx_2^2 \end{aligned}$$

eine quadratische Form $x'Ax$
heißt positiv semi definit, falls

$$x'Ax \geq 0 \quad \text{für alle } x$$

heißt positiv definit, falls

$$x'Ax > 0 \quad \text{für alle } x \neq \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Der Gradient von $f(x) = x'Ax$

lautet $\nabla_f(x) = (A + A') \cdot x$

Die Hesse matrix von f lautet

$$H_f(x) = A + A'$$

H_f ist symmetrisch:

$$(A + A')' = A' + (A')' = A' + A = A + A'$$

Falls A positiv definit ist,
dann ist auch $A + A'$ positiv definit

2.3 Löwner kleiner

$A, B \in \mathbb{R}^{n \times n}$. A löwner kleiner als B , $A \leq B$

falls: $x'Ax \leq x'Bx$ für alle $x \in \mathbb{R}^n$

$$\begin{aligned} (\Rightarrow) \quad 0 &\leq x'Bx - x'Ax \\ &= x'(B - A)x \end{aligned}$$

$(\Rightarrow) B - A$ positiv semi definit

2.4 Inverse einer Matrix

A Matrix B Matrix
 $m \times n$ $n \times m$

Falls $B \cdot A = \underset{n \times n}{I}$

nicht eindeutig
definiert

B "Pseudoinverse" von A

A B Matrizen
 $n \times n$ $n \times n$

Falls $A \cdot B = B \cdot A = I_n$: B Inverse von A

Inverse ist eindeutig definiert $\leadsto "A^{-1}"$

Die Inverse muss nicht immer existieren.

Falls A positiv definit ist, existiert A^{-1}
und A^{-1} ist auch pos. def.

$$(A^{-1})' = (A')^{-1}$$

2.5 Schätzer

Ein Schätzer ist eine Funktion der

Zufallsstichprobe (y, X)

Daher ist auch der Schätzer zufällig

Der Wert (die Werte) sind Schätzer für
die unbekannten Parameter $\beta_0, \beta_1, \dots, \beta_k, \sigma^2$

Allgemeine Bezeichnung $\theta \in \Theta$.

Der Schätzer $f(y, X)$ ist linear in y , falls

es eine Matrix A gibt mit $f(y, X) = A \cdot y$

Ein Schätzer f ist erwartungstreu für θ
falls $E[f(y, X)] = \theta$ für alle $\theta \in \Theta$

Ein Schätzer f ist effizient, falls
 f ist erwartungstreu

$$\text{Var}(f(y, X)) \stackrel{\text{L\"owner}}{\leq} \text{Var}(g(y, X))$$

andere Schätzer

für alle erwartungstreuen
Schätzer $g(y, X)$

2.6 Lineare und erwartungstreue Schätzer für β

f Schätzfunktion für β

Für ein gegebenes X (Regressor matrix)

$$f(y) = A \cdot y, \text{ wobei } A \text{ von } X \text{ abhängt}$$

MLR 1 $y = X\beta + u$

$$\Rightarrow f(y) = A(X\beta + u) = AX\beta + Au$$

$$\begin{aligned} E[f(y)|X] &= E[AX\beta + Au|X] \\ &= AX\beta + A \underbrace{E[u|X]}_{\text{MR4: } = 0} = AX\beta \end{aligned}$$

$$E[f(y) | X] = AX\beta$$

$$= \beta, \text{ falls } AX = I$$

Falls $AX = I$, dann ist $f(y) = A \cdot y$

erwartungstreu für β

Varianzen und Kovarianzen von $A \cdot y$

$$\sigma^2(f(y) | X) = \sigma^2(A \cdot y | X)$$

$$\text{MLR 1} = \sigma^2(A(X\beta + u) | X)$$

$$\mathbb{P}(f(y) | x) = \dots = \mathbb{P}(Ax\beta + Au | x)$$

$$= \mathbb{P}(Au | x) \quad (Ax, \beta \text{ ist konstant})$$

$$= A \underbrace{\mathbb{P}(u | x)} \cdot A'$$

MLR 5

$$= A \cdot \sigma^2 \cdot \underbrace{I}_{= A'} \cdot A'$$

$$= \sigma^2 A \cdot A'$$

Suche ein A sodass $AX = I$ und
Löwner

$$\sigma^2 \cdot A \cdot A' \leq \sigma^2 \cdot B \cdot B' \quad \text{mit } BX = I$$

Kleinste Quadrate Schätzung

Vermutung $b \in \mathbb{R}^{k+1} \rightsquigarrow$ Prognose $\hat{y} = Xb$

$\rightsquigarrow$ Residuum $\hat{u} = y - \hat{y} = y - Xb$

Summe der Quadrate

$$\sum_{i=1}^n \hat{u}_i^2 = \hat{u}' \hat{u} = (y - Xb)'(y - Xb) = S(b)$$

Ziel: minimiere $S(b)$ $\rightsquigarrow$ kleinste Quadrate Schätzer

Multivariate Optimierung

$f: \mathbb{R}^{k+1} \rightarrow \mathbb{R}$ stetig & differenzierbar

Falls $x^* \in \mathbb{R}^{k+1}$ Minimumstelle $\Rightarrow x^*$ stationär

Falls f streng konvex $\Rightarrow$ stationäres x^* ist Minimum

$$S(b) = (y - Xb)'(y - Xb)$$

$$= y'y - y'Xb - (Xb)'y + (Xb)'Xb$$

$$= y'y - y'Xb - b'X'y + b'X'Xb$$

$b'X'y$ ist eine Zahl $\leadsto b'X'y = (b'X'y)'$
 $= y'(X')'(b')' = y'Xb$

$$\Rightarrow S(b) = y'y - 2 \overset{A}{y'X}b + b'X'Xb$$

Falls $f(b) = A \cdot b \Rightarrow J_f = A'$

Falls $f(b) = b'Ab \Rightarrow \nabla_f = (A + A')b$

$$\nabla_S(b) = -2(y'X)' + (X'X + (X'X)')b$$

$$V_S(b) = -2(y'X)' + (X'X + (X'X)')b$$

$$= -2X'y + \underbrace{(X'X + (X'X)')}_{= X'X} b$$

Momente Matrix $X'X$

$$X = \begin{pmatrix} 1 & X_{11} & \dots & X_{1k} \\ 1 & X_{21} & \dots & X_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & \dots & X_{nk} \end{pmatrix}$$

$$X' = \begin{pmatrix} 1 & 1 & \dots & 1 \\ X_{11} & X_{21} & \dots & X_{n1} \\ \vdots & \vdots & \ddots & \vdots \\ X_{1k} & X_{2k} & \dots & X_{nk} \end{pmatrix}$$

$$\begin{aligned} V_S(b) &= -2X'y + 2X'Xb \\ &= 0 \quad (\Rightarrow) \end{aligned}$$

$$2X'Xb = 2X'y$$

$$X'Xb = X'y$$

$$X'X = \begin{pmatrix} L' L & L' x_1 & \dots & L' x_4 \\ x_1' L & x_1' x_1 & \dots & x_1' x_4 \\ \vdots & \vdots & \ddots & \vdots \\ x_k' L & x_k' x_1 & \dots & x_k' x_k \end{pmatrix}$$

$$= \begin{pmatrix} n & \sum_{i=1}^n x_{i1} & \dots & \sum_{i=1}^n x_{i4} \\ \sum_{i=1}^n x_{i1} & \sum_{i=1}^n x_{i1} x_{i1} & \dots & \sum_{i=1}^n x_{i1} x_{i4} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^n x_{ik} & \sum_{i=1}^n x_{ik} x_{i1} & \dots & \sum_{i=1}^n x_{ik} x_{ik} \end{pmatrix}$$

$$nX'X = \begin{pmatrix} 1 & \bar{x}_1 & \dots & \bar{x}_k \\ \bar{x}_1 & \overline{x_1 x_1} & \dots & \overline{x_1 x_k} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{x}_k & \overline{x_k x_1} & \dots & \overline{x_k x_k} \end{pmatrix}$$

$$S_{x_i x_j} = \overline{x_i x_j} - \bar{x}_i \cdot \bar{x}_j$$

$X'X$ ist symmetrisch

$$(X'X)' = X'(X')' = X'X$$

$$\text{Ist } S(\delta) = y'y - 2y'x\delta + \delta'x'x\delta$$

streng konvex?

$$D_S(\delta) = -2x'y + 2x'x\delta$$

Hessematrix

$$H_S = (2x'x)' = 2x'x$$

S ist streng konvex, falls H_S positiv definit

$2X'X$ ist positiv semidefinit, falls

$$b' 2X'X b \geq 0 \text{ für alle } b$$

$2X'X$ ist positiv definit, falls

$$b' 2X'X b > 0 \text{ für alle } b \neq 0$$

$$\hat{y} = X\beta$$

$$\Rightarrow \beta' X' X \beta = (X\beta)' (X\beta) = \hat{y}' \cdot \hat{y}$$

$$= \sum_{i=1}^n \underbrace{(\hat{y}_i)^2}_{\geq 0} \geq 0 \Rightarrow X'X \text{ pos. sem. def.}$$

$$\text{MLR 3} \quad \text{rk}(X) = k+1$$

$$X\beta = 0 \Rightarrow \beta = 0$$

$$\beta \neq 0 \Rightarrow X\beta \neq 0$$

$$\beta \neq 0 \Rightarrow \hat{y} \neq 0 \Rightarrow \hat{y}' \hat{y} > 0 \Rightarrow X'X \text{ pos def.}$$

Zusammenfassung

$f(s)$ ist streng konvex

Falls s $X'Xs = X'y$ erfüllt,

dann ist s eine strikte Minimumstelle
von $f(s)$

Momente Methode (method of moments)

MLR 2 : y, X zufallsstichprobe

$$E[X'y] = E[X'(X\beta + u)]$$

MLR 1 $\nearrow$

$$= E[X'X\beta + X'u]$$

$$= E[X'X]\beta + \underbrace{E[X'u]}_{=0}$$

MLR 4: $E[u|X] = 0$ $\nearrow$

theoretische Annahme:

$$E[X'y] = E[X'X]\beta$$

$$E[X'y] = E[X'X]\beta$$

Empirische Momente: (wähle b für β , so dass gilt)

$$X'y = X'Xb$$

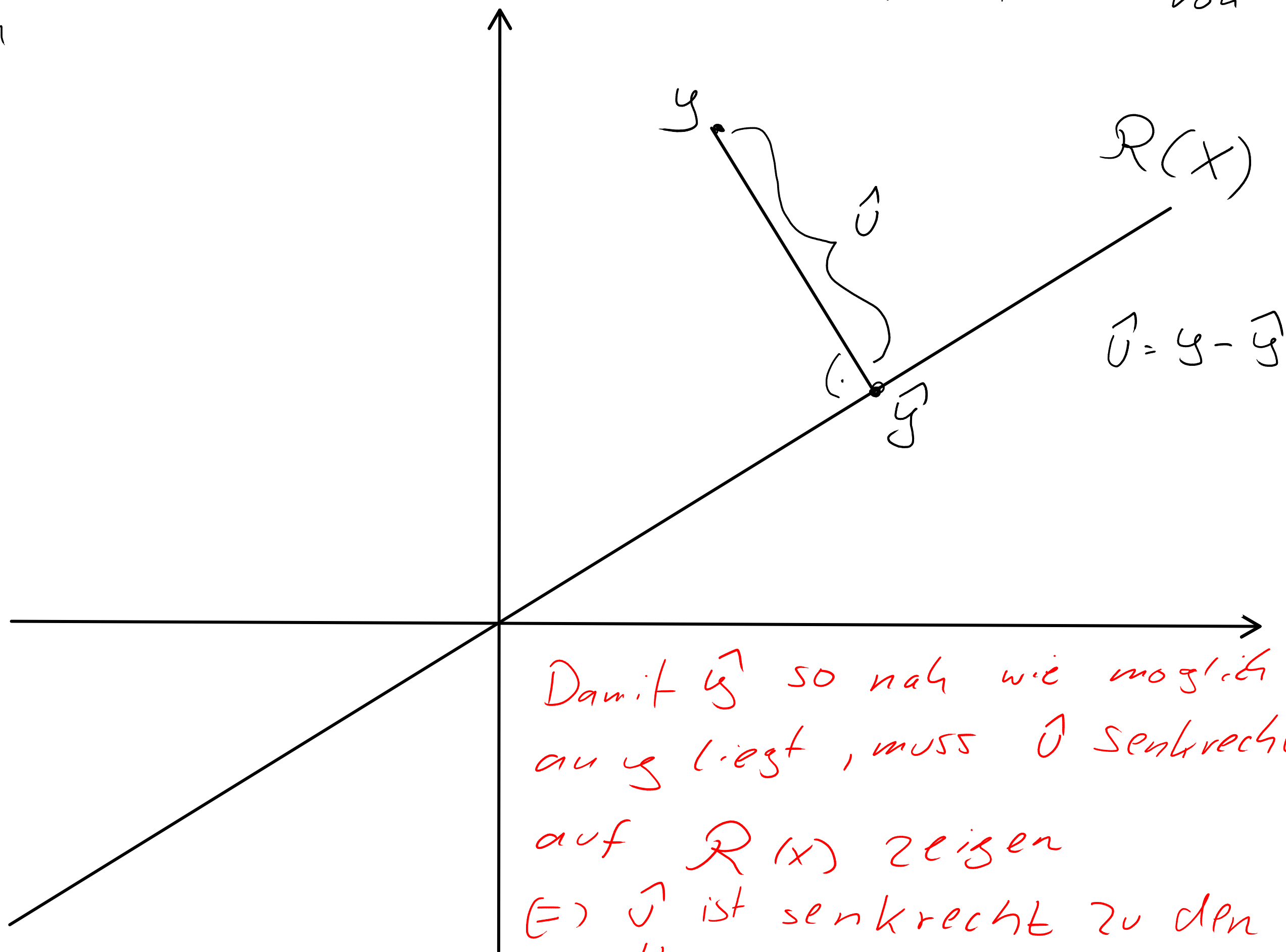
Identische Gleichung wie notw. Bed. der
kleinsten Quadrate Methode

Orthogonale Projektion

"Ebene": Spaltenraum von X

$\mathbb{R}^4$

Spaltenraum von X



Damit $\hat{y}$ so nah wie möglich
an y liegt, muss $\vec{u}$ senkrecht
auf $R(X)$ zeigen
($\Rightarrow$) $\vec{u}$ ist senkrecht zu den
Spalten von X

$$\hat{U}'X = 0 \quad (\Rightarrow) \quad X' \hat{U} = 0$$

mit $\hat{U} = y - \hat{y} = y - X\beta$:

$$X' \overbrace{(y - X\beta)}^{\hat{U}} = 0$$

$$(\Rightarrow) X'y - X'X\beta = 0$$

$$(\Rightarrow) \boxed{X'y = X'X\beta}$$

Lösung von $X'Xb = X'y$

$X'X$ positiv definit (symmetrisch) - pds

$\Rightarrow$ es existiert die Inverse $(X'X)^{-1}$

$$\underbrace{(X'X)^{-1} X'X}_I b = (X'X)^{-1} X'y$$

$$\boxed{\hat{\beta} = b = (X'X)^{-1} X'y}$$

Gauss Markov Theorem

Falls MLR1, MLR2, MLR3, MLR4, MLR5
gelten, dann ist $\hat{\beta}$ Varianzminimal
(effizient)

unter allen erwartungstreuen linearen
Schätzern für β

β best effizient

L linear

U unbiased unverzerrt

E estimator Schätzer

$$\hat{\beta} = \underbrace{(X'X)^{-1}X'}_A y$$

$$AX = (X'X)^{-1}X' \cdot X = I \Rightarrow E[\hat{\beta}] = \beta$$

$$\text{var}(\hat{\beta}) = \sigma^2 \cdot A \cdot A'$$

$$= \sigma^2 (X'X)^{-1} X' \left(\underbrace{(X'X)^{-1} X'}_{C'} \right)'$$

$$= \sigma^2 (X'X)^{-1} X' (X')' \cdot \left((X'X)^{-1} \right)'$$

$$= \sigma^2 \underbrace{(X'X)^{-1} X' X}_I \cdot \left((X'X)' \right)^{-1} = \sigma^2 (X'X)^{-1}$$

Einfachregression $k=1$

$$X = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix} \quad X'X = \begin{pmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{pmatrix}$$

$$\hookrightarrow X'X = \begin{pmatrix} 1 & \bar{x} \\ \bar{x} & \overline{x^2} \end{pmatrix}$$

$$\begin{aligned} \sigma^2(\hat{\beta}) &= \sigma^2 (X'X)^{-1} = \sigma^2 \begin{pmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{pmatrix}^{-1} \\ &= \sigma^2 \frac{1}{n \sum x_i^2 - (\sum x_i)^2} \begin{pmatrix} \sum x_i^2 & -\sum x_i \\ -\sum x_i & n \end{pmatrix} \end{aligned}$$

Schätzer $\hat{\beta} = (X'X)^{-1}X'y$

Estimator ✓ (Schätzer für $\beta \in \mathbb{R}^{k+1}$)

Dimension $\hat{\beta}$?

$$\underbrace{\underbrace{(X'X)^{-1}}_{\substack{k+1 \times n \quad n \times k+1 \\ k+1 \times k+1}}}_{k+1 \times k+1} \underbrace{X'y}_{\substack{k+1 \times n \quad n \times 1 \\ k+1 \times 1}} \in \mathbb{R}^{k+1}$$

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

Linear? ✓

$$A = (X'X)^{-1} X'$$

$$\Rightarrow \hat{\beta} = A \cdot y$$

$$\hat{\beta} = \underbrace{(X'X)^{-1}X'}_A y$$

Unbiased ✓

Variante lang

$$\text{MLR 1: } y = X\beta + u$$

Variante kurz:

$$\text{falls } A \cdot X = I \Rightarrow E[\hat{\beta}] = \beta$$

$$A \cdot X = (X'X)^{-1}X' \cdot X = I \quad \checkmark$$

$$\begin{aligned} E[\hat{\beta} | X] &= E[(X'X)^{-1}X'(X\beta + u) | X] \\ &= E[\underbrace{(X'X)^{-1}X'X}_{I}\beta + (X'X)^{-1}X'u | X] \\ &= \beta + E[(X'X)^{-1}X'u | X] \\ &= \beta + (X'X)^{-1}X' \underbrace{E[u | X]}_{\text{MLR 4: } = 0} = \beta \end{aligned}$$

$$E[\hat{\beta}] = E_x[E[\hat{\beta} | X]] = E_x[\beta] = \beta \quad \checkmark$$

$$\hat{\beta} = \underbrace{(x'x)^{-1}x'y}_{= A \cdot y}$$

$$\sigma^2(\hat{\beta}|x) = \sigma^2(x'x)^{-1}$$

$$\sigma^2(\hat{\beta}|x) = \sigma^2 \cdot A \cdot A'$$

zu zeigen

$$\sigma^2(\hat{\beta}|x) \leq \sigma^2(\hat{\beta}^v|x) \quad \text{für alle } \hat{\beta}^v = B \cdot y$$

mit $Bx = \underline{1}$

$$\Leftrightarrow \sigma^2 \cdot A \cdot A' \leq \sigma^2 \cdot B \cdot B'$$

$$\Leftrightarrow A \cdot A' \leq B \cdot B'$$

$$\Leftrightarrow B \cdot B' - A \cdot A' \quad \text{positiv semi definit}$$

$$\Leftrightarrow \overset{\text{zeilenvektor}}{x'} (B \cdot B' - A \cdot A') \overset{\text{spaltenvektor}}{x} \geq 0 \quad \text{für alle } x \in \mathbb{R}^n$$

$$(B - A)(B - A)'$$

$$A = (X'X)^{-1}X'$$

$$= B \cdot B' - B A' - A B' + A A'$$

$$B \cdot A' = B \cdot (X'X)^{-1}X' = B \cdot (X')' \cdot ((X'X)^{-1})'$$

$$= B \cdot X \cdot ((X'X)')^{-1} = \underbrace{BX}_{=I} (X'X)^{-1} \Rightarrow B A' = (X'X)^{-1}$$

Analog:

$$A B' = (X'X)^{-1}$$

$$A \cdot A' = (X'X)^{-1}$$

$$\rightarrow = BB' - (X'X)^{-1} - (X'X)^{-1} + (X'X)^{-1} = BB' - (X'X)^{-1} = BB' - AA'$$

Wegen $(B-A)(B-A)' = BB' - AA'$ gilt

$BB' - AA'$ pos. sem. def $\Leftrightarrow (B-A)(B-A)'$ pos sem def.

$$\underbrace{x'(B-A)}_{z'} \underbrace{(B-A)'}_z \cdot x = z' \cdot z, \quad x \in \mathbb{R}^n$$

$$z' = ((B-A)' \cdot x)' = x' ((B-A)')' = x' (B-A)$$

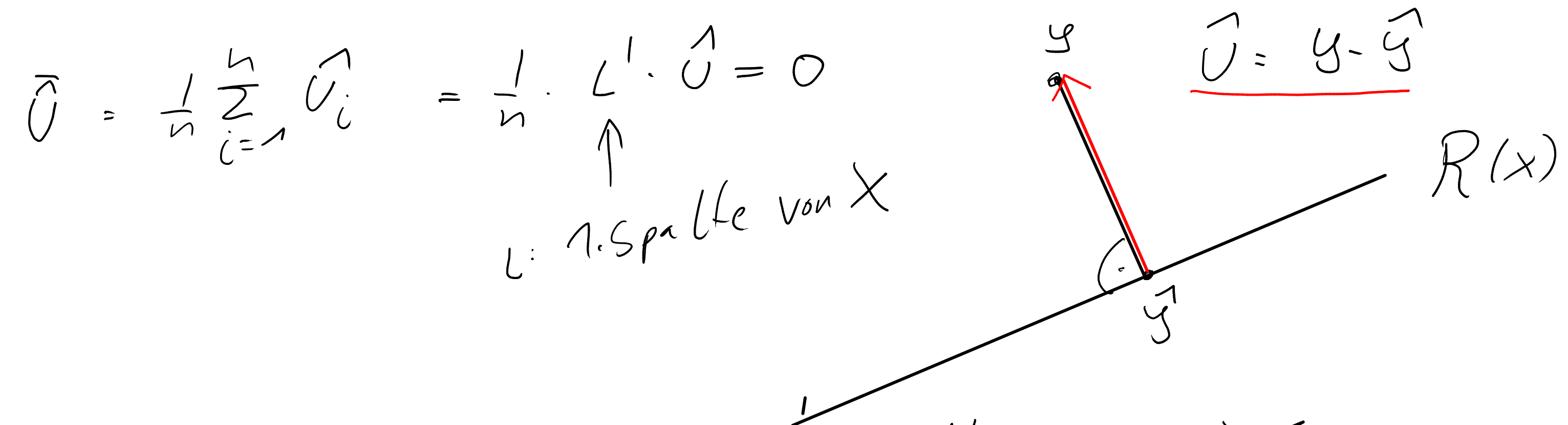
$$z' \cdot z = \sum \underbrace{z_i^2}_{\geq 0} \geq 0 \quad \square$$

Schätzer für $\sigma^2 = \text{Var}(u_i) \quad i = 1, \dots, n$

$k=1$ (Einfachregression)

$$s^2 = \frac{1}{n-2} \sum_{i=1}^n (\hat{u}_i - \bar{\hat{u}})^2 = \frac{1}{n-2} \sum_{i=1}^n \hat{u}_i^2 = \frac{1}{n-2} \hat{U}'\hat{U}$$

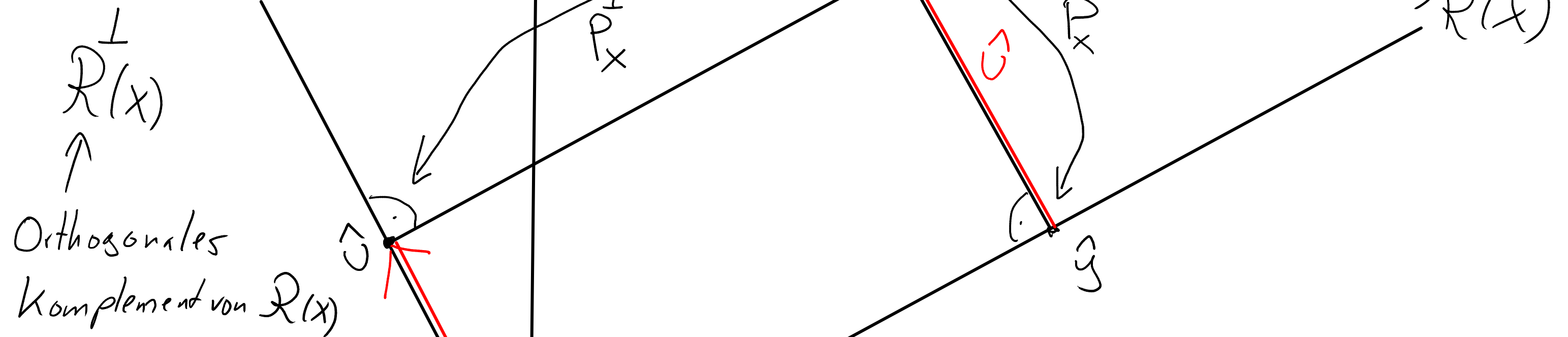
$$\bar{\hat{U}} = \frac{1}{n} \sum_{i=1}^n \hat{u}_i = \frac{1}{n} \cdot \underset{\substack{\uparrow \\ L: 1.\text{Spalte von } X}}{L'} \cdot \hat{U} = 0$$



$\hat{U} \perp X$: $\hat{U}$ ist senkrecht zu allen Spalten von X

Orthogonale Projektion

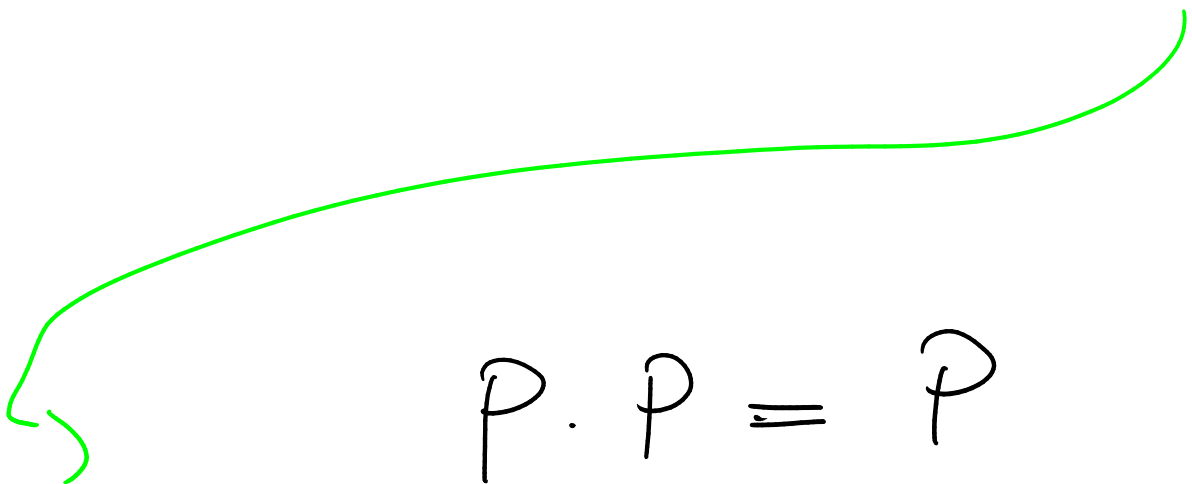
Spaltenraum von X



$$\hat{y} = X \cdot \hat{\beta} = \underbrace{X \cdot (X'X)^{-1} X'}_{P_X} \cdot y$$

$$\hat{u} = y - \hat{y} = y - P_X \cdot y = (\underbrace{I - P_X}_{P_X^\perp}) y$$

Eine orthogonale Projektionsmatrix ist
symmetrisch und idempotent


$$P \cdot P = P$$

$$y \leadsto R(x) \quad \hat{y} = P \cdot y$$

$$\hat{y} \leadsto R(x) \quad \underline{P \cdot \hat{y}} = \underline{P \cdot P \cdot \hat{y}} \Rightarrow P = P \cdot P$$

$$P_X \cdot P_X = X \underbrace{(X^T X)^{-1} X^T}_{\underline{I}} \cdot X (X^T X)^{-1} X^T = X (X^T X)^{-1} X^T = P_X$$

$$P_X^\perp P_X^\perp = (I - P_X)(I - P_X) = \underline{I \cdot I} - P_X - P_X + \underbrace{P_X P_X}_{P_X} = I - P_X = P_X^\perp$$

$$\hat{U}'\hat{U} = (P_x^\perp \cdot y)' (P_x^\perp \cdot y)$$

$$\hat{U} = P_x^\perp \cdot y$$

$$= y' (P_x^\perp)' P_x^\perp \cdot y$$

$$= U' (P_x^\perp)' P_x^\perp \cdot U$$

$$= U' P_x^\perp P_x^\perp \cdot U$$

$$= U' P_x^\perp \cdot U$$

$$\text{MLR 1: } y = X\beta + u$$

$$\Rightarrow \hat{U} = P_x^\perp \cdot (X\beta + u)$$

$$= (I - X(X'X)^{-1}X')X\beta$$

$$+ (I - X(X'X)^{-1}X') \cdot u$$

$$= \overbrace{X\beta - X \underbrace{(X'X)^{-1}X'}_I \cdot \underbrace{X'\beta}_I}_{\text{0}} + P_x^\perp \cdot u$$

$$\hat{U}'\hat{U} = U' P_x^\perp \cdot U$$

Spur einer Matrix (trace)

$$\text{tr}(A) = \sum_{i=1}^n a_{ii}$$

$n \times n$

z.B. $\text{tr}(I_n) = n$

Falls A idempotent $\Rightarrow \text{tr}(A) = \text{rk}(A)$

Es gilt $\text{rk}(P_X) = \text{rk}(X) = k+1$

$$\text{rk}(P_X^\perp) = n - (k+1)$$

Satz zyklische Vertauschungen

$$\text{tr}(ABC) = \text{tr}(C \cdot AB) = \text{tr}(BCA)$$

$$\text{tr}(\hat{\beta}'\hat{\beta}) = \hat{\beta}'\hat{\beta}$$

$$= \text{tr}(U' P_X^\perp U)$$

$$= \text{tr}(U U' P_X^\perp)$$

$$E[\hat{\beta}'\hat{\beta} | X] = E[\text{tr}(U U' P_X^\perp) | X]$$

$$= \text{tr}(E[U U' P_X^\perp | X])$$

$$= \text{tr}(E[U U' | X] P_X^\perp)$$

mit $E[u|x] = 0$ (MLR 4)

gilt

$$E[uu'|x] = E[(u - E[u|x])(u - E[u|x])']]$$
$$= \sigma^2(x) = \sigma^2 \cdot I_n \quad \text{MLR 5}$$

$$E[\hat{\beta}'\hat{\beta}|x] = \dots = \text{tr} \left(\underbrace{E[uu'|x]}_{\sigma^2 \cdot I_n} P_x^\perp \right)$$
$$= \text{tr}(\sigma^2 \cdot I_n \cdot P_x^\perp) = \sigma^2 \cdot \text{tr}(P_x^\perp)$$
$$= \sigma^2 \cdot \text{rk}(P_x^\perp) = \sigma^2 \cdot (n - k - 1)$$

Also für

$$\hat{\sigma}^2 = \frac{1}{n-k-1} \hat{u}' \hat{u}$$

$$E[\hat{\sigma}^2 | X] = E\left[\frac{1}{n-k-1} \hat{u}' \hat{u} | X\right]$$

$$= \frac{1}{n-k-1} E[\hat{u}' \hat{u} | X]$$

$$= \frac{1}{\cancel{n-k-1}} \sigma^2 (\cancel{n-k-1}) = \sigma^2$$

$$E[E[\hat{\sigma}^2 | X]] = E[\sigma^2] = \sigma^2 \quad \checkmark$$

Streuungszerlegung & Bestimmtheitsmaß

sum of total squares

$$SST = \sum_{i=1}^n (y_i - \bar{y})^2$$

explained sum of squares

$$SSE = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2$$

unexplained sum of squares

$$SSU = \sum_{i=1}^n \hat{u}_i^2$$

$$\hat{u} = y - \hat{y}$$

$$\Rightarrow \hat{y} = y - \hat{u}$$

$$\Rightarrow \bar{\hat{y}} = \bar{y} - \bar{\hat{u}} = \bar{y} \quad \leftarrow 0$$

$$\underline{\underline{(\bar{\hat{y}} = \bar{y})}}$$

$$\bar{\hat{u}} = 0$$

Es gilt

$$\hat{u}_i = y_i - \hat{y}_i \Leftrightarrow y_i = \hat{u}_i + \hat{y}_i$$

$$SST = SSE + SSU$$

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (\hat{y}_i - \bar{y} + \hat{u}_i)^2 = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 + 2(\hat{y}_i \hat{u}_i) + \hat{u}_i^2$$

$$= \underbrace{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}_{SSE} + 2 \underbrace{\sum_{i=1}^n \hat{y}_i \cdot \hat{u}_i}_{=0} + \underbrace{\sum_{i=1}^n \hat{u}_i^2}_{SSU}$$

$$\hat{y} \in \mathcal{R}(X), \hat{u} \in \mathcal{R}^\perp(X) \Rightarrow \hat{y}' \cdot \hat{u} = 0$$

Bestimmtheitsmaß

Interpretation

R^2 : 100% der
Streuung von y
wird durch die
Regression erklärt

$$R^2 = \frac{SSE}{SST} \in [0, 1]$$

$$SSE = \sum_{i=1}^n \underbrace{(\hat{y}_i - \bar{y})^2}_{\geq 0} \geq 0$$

$$SSU \geq 0 \quad (\text{analog})$$

$$\Rightarrow SSE = SST - SSU \leq SST$$