

Mathematik 2

Dr. Thomas Zehrt

Matrizen in der Ökonomie

Benötigtes Vorwissen: Der Stoff der Vorlesung „Statistik“ und „Mathematik“ wird als bekannt vorausgesetzt, insbesondere das Kapitel „Regression“

Inhaltsverzeichnis

1	Einführung in die Ökonometrie	2
1.1	Einfache lineare Regression	2
1.2	Multiple lineare Regression	3
1.2.1	Das Modell	3
1.2.2	Annahmen im klassischen linearen Regressionsmodell	4
1.2.3	Der KQ-Schätzer	6
1.2.4	Die Vektoren \mathbf{y} , $\hat{\mathbf{y}} = \mathbf{X}\mathbf{b}$ und \mathbf{e}	8
1.2.5	Varianzzerlegung	10
2	Portfoliotheorie	12
3	Input-Output-Modelle	15
4	Mehrstufige Produktionsprozesse	18
5	Übungsaufgabe	21

1 Einführung in die Ökonometrie

1.1 Einfache lineare Regression

In einer grossen Grundgesamtheit werden zwei stetige Zufallsvariablen X und Y untersucht. Dazu werden beide Zufallsvariablen n mal gemessen (Stichprobe der Grösse n): $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$.

- Erinnerung:

Der Korrelationskoeffizient ρ ist ein dimensionsloses Mass, das die Stärke und die Richtung eines linearen Zusammenhangs zwischen X und Y misst.

- Ziel der Regressionsanalyse:

Der (hoffentlich real existierende) lineare Zusammenhang zwischen den beiden Zufallsvariablen soll durch ein (einfaches) Modell erfasst werden.

- Einfachstes (lineares) deterministisches Modell einer Ursache-Wirkungsbeziehung zwischen X und Y :

$$Y = f(X) = \beta_0 + \beta_1 X$$

Dabei sind β_0 und β_1 irgendwelche reelle Zahlen.

- Univariates (von einer Variablen abhängig) lineares Modell:

$$Y = f(X) = \beta_0 + \beta_1 X + \epsilon$$

- β_0, β_1 unbekannte Modellparameter
- ϵ zufällige Fehlervariable mit $E(\epsilon) = 0$ und $Var(\epsilon) = \sigma^2$
- X wird stets als gegeben (also **nicht** zufällig) angesehen, ist also eigentlich keine Zufallsvariable.

1.2 Multiple lineare Regression

1.2.1 Das Modell

Häufiges Problem in der Praxis: Die zufällige stetige Variable Y (auch Response genannt) hängt von k verschiedenen Einflussgrößen X_1, X_2, \dots, X_k ab (alle stetig und **nicht** zufällig).

Modell:

$$Y = f(X) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \epsilon$$

Alle Variablen seien n -mal beobachtet wurden. Daraus resultieren n lineare Gleichungen:

$$\begin{aligned} y_1 &= \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \dots + \beta_k x_{1k} + \epsilon_1 \\ y_2 &= \beta_0 + \beta_1 x_{21} + \beta_2 x_{22} + \dots + \beta_k x_{2k} + \epsilon_2 \\ &\vdots \\ y_n &= \beta_0 + \beta_1 x_{n1} + \beta_2 x_{n2} + \dots + \beta_k x_{nk} + \epsilon_n \end{aligned}$$

Matrixschreibweise:

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \beta_0 \begin{pmatrix} \mathbf{x}_0 \\ \downarrow \\ 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} + \beta_1 \begin{pmatrix} \mathbf{x}_1 \\ \downarrow \\ x_{11} \\ x_{21} \\ \vdots \\ x_{n1} \end{pmatrix} + \dots + \beta_k \begin{pmatrix} \mathbf{x}_k \\ \downarrow \\ x_{1k} \\ x_{2k} \\ \vdots \\ x_{nk} \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{pmatrix}$$

oder

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & \dots & x_{1k} \\ 1 & x_{21} & \dots & x_{2k} \\ \vdots & \vdots & \dots & \vdots \\ 1 & x_{n1} & \dots & x_{nk} \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{pmatrix}$$

$$\begin{array}{ccccccc} & & & & & & \\ & & & & & & \\ & & & & & & \\ & & & & & & \\ & & & & & & \\ & & & & & & \\ \uparrow & & & \uparrow & & \uparrow & \uparrow \\ \mathbf{y} & & & \mathbf{X} & & \boldsymbol{\beta} & \boldsymbol{\epsilon} \end{array}$$

Durch die Festlegung $\mathbf{x}_0 = \mathbf{1} = (1, 1, \dots, 1)^T$, wird eine konstante (Intercept) ins Modell integriert.

Für eine einfachere Notation geben wir auch die bisherige strenge Unterscheidung zwischen Zufallsvariablen Y und deren Realisierung y auf. Der Vektor \mathbf{y} soll beides (je nach Situation) ausdrücken.

Die Matrix \mathbf{X} ist natürlich **keine** Zufallsgrösse.

Beispiel 1.1 Ein Rechenbeispiel soll die Formeln verdeutlichen. Dabei sind die Zahlen bewusst einfach gewählt. Das Modell enthalte zwei (exogene) Variablen

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \epsilon$$

mit den beobachteten Werten:

$$\mathbf{y} = \begin{pmatrix} 3 \\ 1 \\ 8 \\ 3 \\ 5 \end{pmatrix} \text{ und } \mathbf{X} = \begin{pmatrix} 1 & 3 & 5 \\ 1 & 1 & 4 \\ 1 & 5 & 6 \\ 1 & 2 & 4 \\ 1 & 4 & 6 \end{pmatrix}$$

Das heisst also, dass Sie $n = 5$ Messungen durchgeführt haben, wobei Sie bei der ersten Messung $Y = y_1 = 3$, $X_1 = x_{11} = 3$ und $X_2 = x_{12} = 5$ erhalten haben.

Genauer glauben wir an die folgenden 5 Gleichungen:

$$\begin{aligned} 3 &= \beta_0 + \beta_1 3 + \beta_2 5 + \epsilon_1 \\ 1 &= \beta_0 + \beta_1 1 + \beta_2 4 + \epsilon_2 \\ 8 &= \beta_0 + \beta_1 5 + \beta_2 6 + \epsilon_3 \\ 3 &= \beta_0 + \beta_1 2 + \beta_2 4 + \epsilon_4 \\ 5 &= \beta_0 + \beta_1 4 + \beta_2 6 + \epsilon_5 \end{aligned}$$

1.2.2 Annahmen im klassischen linearen Regressionsmodell

1. $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$
2. \mathbf{X} nicht zufällig
3. $\text{Rang}(\mathbf{X}) = k + 1 \leq n$

Es existiert keine lineare Beziehung zwischen den Grössen X_1, X_2, \dots, X_k .

Die Matrix $\mathbf{X}^T \mathbf{X}$ ist quadratisch und hat $k+1$ Zeilen und Spalten. Unter den gegebenen Voraussetzungen ist sie regulär, also invertierbar. Insbesondere existiert die Inverse $(\mathbf{X}^T \mathbf{X})^{-1}$.

Im Spezialfall $k = 1$, also $Y = \beta_0 + \beta_1 X + \epsilon$ gilt

$$\mathbf{X}^T \cdot \mathbf{X} = \begin{pmatrix} 1 & 1 & \dots & 1 \\ x_{11} & x_{21} & \dots & x_{n1} \end{pmatrix} \begin{pmatrix} 1 & x_{11} \\ 1 & x_{21} \\ \vdots & \vdots \\ 1 & x_{n1} \end{pmatrix} = \begin{pmatrix} n & \sum_{i=1}^n x_{i1} \\ \sum_{i=1}^n x_{i1} & \sum_{i=1}^n x_{i1}^2 \end{pmatrix}$$

4. $E(\boldsymbol{\epsilon}) = \mathbf{0}$

$$5. E(\boldsymbol{\epsilon} \cdot \boldsymbol{\epsilon}^T) = \sigma^2 \cdot \mathbf{I}_n$$

Alle Fehlervariablen haben die Varianz σ^2 und sind unkorreliert.

$$\boldsymbol{\epsilon} \cdot \boldsymbol{\epsilon}^T = \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{pmatrix} \cdot (\epsilon_1 \ \epsilon_2 \ \dots \ \epsilon_n) = \begin{pmatrix} \epsilon_1^2 & \epsilon_1 \epsilon_2 & \dots & \epsilon_1 \epsilon_n \\ \epsilon_1 \epsilon_2 & \epsilon_2^2 & \dots & \epsilon_2 \epsilon_n \\ \vdots & \vdots & \dots & \vdots \\ \epsilon_1 \epsilon_n & \epsilon_2 \epsilon_n & \dots & \epsilon_n^2 \end{pmatrix}$$

Beispiel 1.2

Mit $\mathbf{X} = \begin{pmatrix} 1 & 3 & 5 \\ 1 & 1 & 4 \\ 1 & 5 & 6 \\ 1 & 2 & 4 \\ 1 & 4 & 6 \end{pmatrix}$ ergibt sich sofort die reguläre Matrix

$$\mathbf{X}^T \cdot \mathbf{X} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 3 & 1 & 5 & 2 & 4 \\ 5 & 4 & 6 & 4 & 6 \end{pmatrix} \cdot \begin{pmatrix} 1 & 3 & 5 \\ 1 & 1 & 4 \\ 1 & 5 & 6 \\ 1 & 2 & 4 \\ 1 & 4 & 6 \end{pmatrix} = \begin{pmatrix} 5 & 15 & 25 \\ 15 & 55 & 81 \\ 25 & 81 & 129 \end{pmatrix} \text{ und}$$

$$(\mathbf{X}^T \cdot \mathbf{X})^{-1} = \begin{pmatrix} 26.7 & 4.5 & -8 \\ 4.5 & 1 & -1.5 \\ -8 & -1.5 & 2.5 \end{pmatrix}$$

1.2.3 Der KQ-Schätzer

Wir sind nun an optimalen Schätzungen $\hat{\boldsymbol{\beta}}$ und $\hat{\sigma}^2$ der unbekannt Parameter $\boldsymbol{\beta}$ und σ^2 interessiert. Zunächst schreiben wir mit dem Residuenvektor \mathbf{e}

$$\mathbf{y} = \underbrace{\mathbf{X} \boldsymbol{\beta} + \boldsymbol{\epsilon}}_{\text{Realität}} \approx \underbrace{\mathbf{X} \hat{\boldsymbol{\beta}} + \mathbf{e}}_{\text{Schätzung}}$$

Schätzprinzip der kleinsten Quadrate:

Wähle \mathbf{b} so, dass die Summe der quadrierten Residuen

$$r(\mathbf{b}) = \mathbf{e}^T \mathbf{e} = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (\mathbf{y} - \mathbf{X} \mathbf{b})^T (\mathbf{y} - \mathbf{X} \mathbf{b})$$

kleinstmöglich wird.

Satz 1 Die Funktion $r = r(\mathbf{b})$ ist minimal, falls

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

gewählt wird. Diese Schätzung $\hat{\boldsymbol{\beta}}$ heisst KQ-Schätzung von $\boldsymbol{\beta}$.

Beweisskizze

Zunächst zeigt man, dass

$$r(\mathbf{b}) = \mathbf{y}^T \mathbf{y} - 2\mathbf{b}^T \mathbf{X}^T \mathbf{y} + \mathbf{b}^T \mathbf{X}^T \mathbf{X} \mathbf{b}$$

gilt. Ableiten nach \mathbf{b} ergibt dann die notwendige Bedingung erster Ordnung für das Vorliegen eines Minimums im Punkt $\mathbf{b} = \hat{\boldsymbol{\beta}}$

$$\frac{\partial r}{\partial \mathbf{b}}(\hat{\boldsymbol{\beta}}) = -2\mathbf{X}^T \mathbf{y} + 2\mathbf{X}^T \mathbf{X} \hat{\boldsymbol{\beta}} = \mathbf{0},$$

woraus die so genannte Normalengleichung

$$(\mathbf{X}^T \cdot \mathbf{X}) \cdot \hat{\boldsymbol{\beta}} = \mathbf{X}^T \cdot \mathbf{y}$$

die sich leicht zur Behauptung des Satzes umformen lässt.

Um die Notationen etwas zu vereinfachen, schreiben wir oft auch \mathbf{b} statt $\hat{\boldsymbol{\beta}}$.

Für die konkrete Berechnung des Vektors $\mathbf{b} = \hat{\boldsymbol{\beta}}$ ist natürlich die Normalengleichung besser geeignet als die Formel oben im Satz, denn man kann auf das ineffiziente Invertieren der Matrix verzichten. Das durch die Normalengleichung definierte lineare Gleichungssystem kann effizient mit dem Gauss-Verfahren gelöst werden. \square

Beispiel 1.3 Für unser Beispiel gilt nun

$$\begin{aligned}\hat{\boldsymbol{\beta}} = \mathbf{b} &= (\mathbf{X}^T \cdot \mathbf{X})^{-1} \cdot \mathbf{X}^T \cdot \mathbf{y} \\ &= \underbrace{\begin{pmatrix} 26.7 & 4.5 & -8 \\ 4.5 & 1 & -1.5 \\ -8 & -1.5 & 2.5 \end{pmatrix}}_{=(\mathbf{X}^T \mathbf{X})^{-1}} \cdot \underbrace{\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 3 & 1 & 5 & 2 & 4 \\ 5 & 4 & 6 & 4 & 6 \end{pmatrix}}_{=\mathbf{X}^T} \cdot \underbrace{\begin{pmatrix} 3 \\ 1 \\ 8 \\ 3 \\ 5 \end{pmatrix}}_{\mathbf{y}} \\ &= \begin{pmatrix} 4.0 \\ 2.5 \\ -1.5 \end{pmatrix}\end{aligned}$$

Regressionsgleichung $Y = 4 + 2.5X_1 - 1.5X_2$

Beispiel 1.4 Im Falle der einfachen Regression (d.h. $k = 1$) vereinfacht sich die Normalgleichung $(\mathbf{X}^T \cdot \mathbf{X}) \cdot \hat{\boldsymbol{\beta}} = \mathbf{X}^T \cdot \mathbf{y}$ zu:

$$\begin{pmatrix} n & \sum_{i=1}^n x_{i1} \\ \sum_{i=1}^n x_{i1} & \sum_{i=1}^n x_{i1}^2 \end{pmatrix} \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & \dots & 1 \\ x_{11} & x_{21} & \dots & x_{n1} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

bzw. zu den beiden aus der Statistik wohl bekannten linearen Gleichungen

$$\begin{aligned}n \cdot \hat{\beta}_0 + \left(\sum_{i=1}^n x_{i1} \right) \cdot \hat{\beta}_1 &= \sum_{i=1}^n y_i \\ \left(\sum_{i=1}^n x_{i1} \right) \cdot \hat{\beta}_0 + \left(\sum_{i=1}^n x_{i1}^2 \right) \cdot \hat{\beta}_1 &= \sum_{i=1}^n x_{i1} y_i\end{aligned}$$

1.2.4 Die Vektoren \mathbf{y} , $\hat{\mathbf{y}} = \mathbf{X}\mathbf{b}$ und \mathbf{e}

Aus der Normalengleichung $\mathbf{X}^T \mathbf{X} \mathbf{b} = \mathbf{X}^T \mathbf{y}$ folgt sofort:

$$\begin{aligned} \mathbf{X}^T \hat{\mathbf{y}} &= \mathbf{X}^T \mathbf{y} \\ \longrightarrow \mathbf{X}^T \hat{\mathbf{y}} &= \mathbf{X}^T (\hat{\mathbf{y}} + \mathbf{e}) = \mathbf{X}^T \hat{\mathbf{y}} + \mathbf{X}^T \mathbf{e} \\ \longrightarrow \mathbf{0} &= \mathbf{X}^T \mathbf{e} \\ \longrightarrow 0 &= \mathbf{b}^T \mathbf{X}^T \mathbf{e} = \hat{\mathbf{y}}^T \mathbf{e} \end{aligned}$$

Folgerungen

- Die vorletzte Zeile $\mathbf{0} = \mathbf{X}^T \mathbf{e}$ der obigen Umformungen bedeutet ausgeschrieben:

$$\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & \dots & x_{1k} \\ 1 & x_{21} & \dots & x_{2k} \\ \vdots & \vdots & \dots & \vdots \\ 1 & x_{n1} & \dots & x_{nk} \end{pmatrix}^T \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} 1 & 1 & \dots & 1 \\ x_{11} & x_{21} & \dots & x_{n1} \\ \vdots & \vdots & \dots & \vdots \\ x_{1k} & x_{2k} & \dots & x_{nk} \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix}.$$

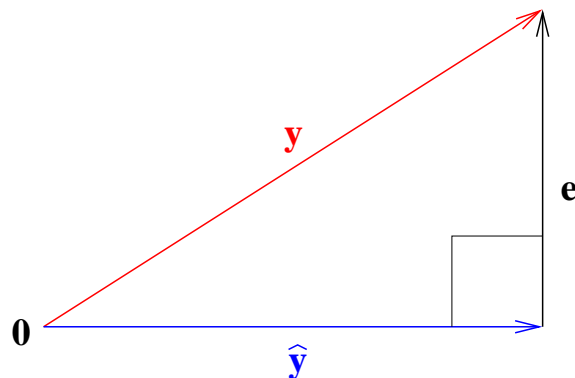
Das bedeutet, dass der Vektor \mathbf{e} auf allen Spaltenvektoren von \mathbf{X} (also auf allen Zeilenvektoren von \mathbf{X}^T) senkrecht steht.

Insbesondere gilt also auch (erste Zeile der Matrizenmultiplikation) die so genannte Zentraleigenschaft:

$$\mathbf{1}^T \mathbf{e} = e_1 + e_2 + \dots + e_n = 0$$

d.h. die Summe aller Residuen (Fehler) ist stets gleich 0.

- Aus der letzten Zeile $0 = \hat{\mathbf{y}}^T \mathbf{e}$ der obigen Umformungen ergibt sich sofort, dass die Vektoren $\hat{\mathbf{y}} = \mathbf{X}\mathbf{b}$ und \mathbf{e} orthogonal sind.



Beispiel 1.5 Für unser Beispiel gilt

$$\mathbf{y} = \begin{pmatrix} 3 \\ 1 \\ 8 \\ 3 \\ 5 \end{pmatrix}, \hat{\mathbf{y}} = \mathbf{X}\mathbf{b} = \begin{pmatrix} 1 & 3 & 5 \\ 1 & 1 & 4 \\ 1 & 5 & 6 \\ 1 & 2 & 4 \\ 1 & 4 & 6 \end{pmatrix} \begin{pmatrix} 4.0 \\ 2.5 \\ -1.5 \end{pmatrix} = \begin{pmatrix} 4 \\ 0.5 \\ 7.5 \\ 3 \\ 5 \end{pmatrix} \text{ und}$$

$$\mathbf{e} = \mathbf{y} - \hat{\mathbf{y}} = \begin{pmatrix} 3 \\ 1 \\ 8 \\ 3 \\ 5 \end{pmatrix} - \begin{pmatrix} 4 \\ 0.5 \\ 7.5 \\ 3 \\ 5 \end{pmatrix} = \begin{pmatrix} -1 \\ 0.5 \\ 0.5 \\ 0 \\ 0 \end{pmatrix}$$

Sind die Vektoren \mathbf{e} und $\hat{\mathbf{y}}$ orthogonal?

Steht \mathbf{e} senkrecht auf jeder Spalte von \mathbf{X} ?

Stimmt die Zentraleigenschaft?

1.2.5 Varianzzerlegung

Wir hatten gesehen, dass die drei Vektoren in der Gleichung $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{e}$ ein rechtwinkliges Dreieck bilden. Es folgt:

$$\begin{aligned} \mathbf{y}^T \mathbf{y} &= (\hat{\mathbf{y}} + \mathbf{e})^T (\hat{\mathbf{y}} + \mathbf{e}) \\ &= (\hat{\mathbf{y}}^T + \mathbf{e}^T) (\hat{\mathbf{y}} + \mathbf{e}) \\ &= \hat{\mathbf{y}}^T \hat{\mathbf{y}} + \underbrace{\hat{\mathbf{y}}^T \mathbf{e}}_{=0} + \underbrace{\mathbf{e}^T \hat{\mathbf{y}}}_{=0} + \mathbf{e}^T \mathbf{e} = \hat{\mathbf{y}}^T \hat{\mathbf{y}} + \mathbf{e}^T \mathbf{e} \end{aligned}$$

Was dort steht, ist nichts anderes als der Satz von Pythagoras, denn z.B. gilt $\mathbf{y}^T \mathbf{y} = y_1^2 + y_2^2 + \dots + y_n^2 = \|\mathbf{y}\|^2$ und das ist das Quadrat der Länge der Hypotenuse in dem Additionsdreieck $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{e}$.

Neben dieser geometrischen Deutung wollen wir versuchen, diese Eigenschaft der drei Vektoren auch statistisch zu deuten. Ein Ausdruck der Gestalt $y_1^2 + y_2^2 + \dots + y_n^2$ könnte uns auch an Varianzberechnung erinnern. Es gilt doch z.B.

$$\text{var}(\mathbf{y}) = \frac{1}{n} \sum_{j=1}^n (y_j - \bar{y})^2 = \frac{1}{n} \sum_{j=1}^n y_j^2 - \bar{y}^2 = \frac{1}{n} \mathbf{y}^T \mathbf{y} - \bar{y}^2.$$

Damit können wir die Relation zwischen den drei Vektoren auch wie folgt umschreiben. Dazu subtrahieren wir auf beiden Seiten der Gleichung das quadrierte arithmetische Mittel \bar{y}^2 (die Datenreihen \mathbf{y} und $\hat{\mathbf{y}}$ haben das selbe arithmetische Mittel) und fügen auf der rechten Seite noch $-\bar{e}^2$ hinzu (wegen der Zentraleigenschaft gilt $\bar{e} = 0$).

$$\begin{aligned} \mathbf{y}^T \mathbf{y} &= \hat{\mathbf{y}}^T \hat{\mathbf{y}} + \mathbf{e}^T \mathbf{e} \\ \longrightarrow \frac{1}{n} \mathbf{y}^T \mathbf{y} &= \frac{1}{n} \hat{\mathbf{y}}^T \hat{\mathbf{y}} + \frac{1}{n} \mathbf{e}^T \mathbf{e} \\ \longrightarrow \underbrace{\frac{1}{n} \mathbf{y}^T \mathbf{y} - \bar{y}^2}_{\text{var}(\mathbf{y})} &= \underbrace{\frac{1}{n} \hat{\mathbf{y}}^T \hat{\mathbf{y}} - \bar{y}^2}_{\text{var}(\hat{\mathbf{y}})} + \underbrace{\frac{1}{n} \mathbf{e}^T \mathbf{e} - \bar{e}^2}_{\text{var}(\mathbf{e})} \end{aligned}$$

Satz 2 (Varianzzerlegung/Streuungszerlegung) Die Gesamtvarianz $\text{var}(\mathbf{y})$ der echten Messwerte \mathbf{y} wird zerlegt in die durch das lineare Modell erklärte Varianz $\text{var}(\hat{\mathbf{y}})$ und die so genannte Residualvarianz $\text{var}(\mathbf{e})$:

$$\text{var}(\mathbf{y}) = \text{var}(\hat{\mathbf{y}}) + \text{var}(\mathbf{e})$$

Bemerkung 1.1 Oft schreibt man diese Relation auch ohne die Umformungen auf die Varianz vorzunehmen:

$$S_Y^2 = S_{\hat{Y}}^2 + S_E^2$$

mit z.B. $S_Y^2 = \mathbf{y}^T \mathbf{y} = y_1^2 + y_2^2 + \dots + y_n^2$.

Bemerkung 1.2 *Dividieren wir die Varianzzerlegungsgleichung durch $\text{var}(\mathbf{y})$ erhalten wir*

$$1 = \frac{\text{var}(\hat{\mathbf{y}})}{\text{var}(\mathbf{y})} + \frac{\text{var}(\mathbf{e})}{\text{var}(\mathbf{y})}$$

Der Term $\frac{\text{var}(\hat{\mathbf{y}})}{\text{var}(\mathbf{y})}$ drückt aus, welcher Anteil der Streuung in den Originaldaten durch das lineare Modell erklärbar ist. Umso grösser dieser Term ist, desto besser beschreibt unser lineares Modell die Messwerte!

Beispiel 1.6 *Für unser Beispiel gilt*

$$\mathbf{y}^T \mathbf{y} = (3 \ 1 \ 8 \ 3 \ 5) \begin{pmatrix} 3 \\ 1 \\ 8 \\ 3 \\ 5 \end{pmatrix} = 108, \quad \hat{\mathbf{y}}^T \hat{\mathbf{y}} = 106.5 \quad \text{und} \quad \mathbf{e}^T \mathbf{e} = 1.5$$

oder

$$\begin{aligned} \text{var}(\mathbf{y}) &= 108/5 - 4^2 = 5.6 \\ \text{var}(\hat{\mathbf{y}}) &= 106.5/5 - 4^2 = 5.3 \\ \text{var}(\mathbf{e}) &= 1.5/5 - 0^2 = 0.3 \end{aligned}$$

sowie

$$\frac{\text{var}(\hat{\mathbf{y}})}{\text{var}(\mathbf{y})} = \frac{5.3}{5.6} = 0.9464,$$

d.h. gut 94% der Originalstreuung wird durch das lineare Modell erklärt. Unser lineares Modell scheint also gut zu den Daten zu passen!

2 Portfoliotheorie

Einführung In der Praxis besitzen Investoren wohl Tausende von verschiedenen Anlagemöglichkeiten wie z.B. Aktien, Anleihen, Fonds, Edelmetalle und Währungen. Normalerweise wird ein (vernünftiger) Investor sein Geld auf mehrere Anlagen verteilen (Portfolio). Doch auf welche Art sollte ein Investor sein Portfolio zusammenstellen? Die zwei wesentlichen Begriffe für jede Investition sind dabei die (erwartete) **Rendite**, die man natürlich maximieren möchte, und das (erwartete) **Risiko**, das man minimieren möchte.

Die erwartete Rendite könnten wir durch das arithmetische Mittel der relativen Kursänderungen der vergangenen Jahre schätzen. Als Mass für das Risiko eignet sich die Varianz (Stärke der Ausschläge) der erwarteten Rendite. Auch die erwartete Varianz einer Anlage kann aus den Daten der vergangenen Jahre geschätzt werden.

Im Jahr 1952 publizierte der 25-jährige Doktorand *Harry Markowitz* von der University of Chicago eine 15-seitige Arbeit mit dem Titel „Portfolio Selection“ im „Journal of Finance“. Für diesen Meilenstein der Finanzmathematik wurde dem Verfasser 1990 der Nobelpreis für Wirtschaftswissenschaften verliehen.

Mathematisches Modell Wir betrachten n verschiedene Geldanlagen $1, 2, \dots, n$ mit zugehörigen Gewichten $\mathbf{x} = (x_1, x_2, \dots, x_n)$, d.h. $\sum_{i=1}^n x_i = 1$ und $x_i \geq 0$ für alle i . Aus diesen Geldanlagen soll ein Portfolio so zusammengestellt werden, dass ein gegebenes Guthaben den Gewichten x_1, x_2, \dots, x_n entsprechend in die n Geldanlagen investiert wird. Die Aufteilung wollen wir nach gewissen sinnvollen Vorgaben optimal gestalten. Weiterhin seien

- $\hat{\mu}_i$ bzw. $\hat{\mu}_P$ die erwarteten Renditen der Geldanlage i bzw. des Gesamtportfolios P (im kommenden Jahr);
- $\hat{\sigma}_i^2$ bzw. $\hat{\sigma}_P^2$ die erwarteten Varianzen der Renditen der Geldanlage i bzw. des Gesamtportfolios P (im kommenden Jahr);
- $\hat{\sigma}_{ij}$ die erwartete Kovarianz der Renditen der Geldanlagen i und j (im kommenden Jahr).

Die Kovarianz ist ein Mass für den Gleichlauf zweier Zahlenreihen. Eine grosse positive Kovarianz würde bedeuten, dass sich beide Geldanlagen ähnlich entwickeln (beide fallen oder steigen gleichzeitig), z.B. bei zwei Aktien aus derselben Branche. Eine kleine (also negative) Kovarianz bedeutet dagegen, dass sich beide Geldanlagen eher gegenläufig entwickeln (steigt die eine, so fällt die andere im Kurs).

Es scheint klar zu sein, dass diese Kovarianzen in die Berechnung der Varianz (Risiko) des Portfolios eingehen müssen. Entwickeln sich alle Geldanlagen ähnlich, so sollte das Risiko meines Portfolios steigen.

Für die erwartete Rendite und die Varianz der Portfoliorendite (die natürlich von der gewählten Aufteilung auf die Geldanlagen abhängt) gilt:

$$\hat{\mu}_P = \sum_{i=1}^n x_i \hat{\mu}_i = x_1 \hat{\mu}_1 + x_2 \hat{\mu}_2 + \dots + x_n \hat{\mu}_n$$

$$\hat{\sigma}_P^2 = \sum_{i=1}^n x_i^2 \hat{\sigma}_i^2 + \underbrace{\sum_{i \neq j} x_i x_j \hat{\sigma}_{ij}}_{=:C} = \mathbf{x}^T \begin{pmatrix} \hat{\sigma}_1^2 & \hat{\sigma}_{12} & \dots & \hat{\sigma}_{1n} \\ \hat{\sigma}_{21} & \hat{\sigma}_2^2 & \dots & \hat{\sigma}_{2n} \\ \vdots & & \ddots & \vdots \\ \hat{\sigma}_{n1} & \hat{\sigma}_{n2} & \dots & \hat{\sigma}_n^2 \end{pmatrix} \mathbf{x}$$

Wir könnten nun versuchen, ein besonders schönes Portfolio zusammenzustellen, z.B. eines, das eine vorgegebene erwartete Rendite $\hat{\mu}_P$ bei minimaler Varianz (interpretiert als Risiko) realisiert.

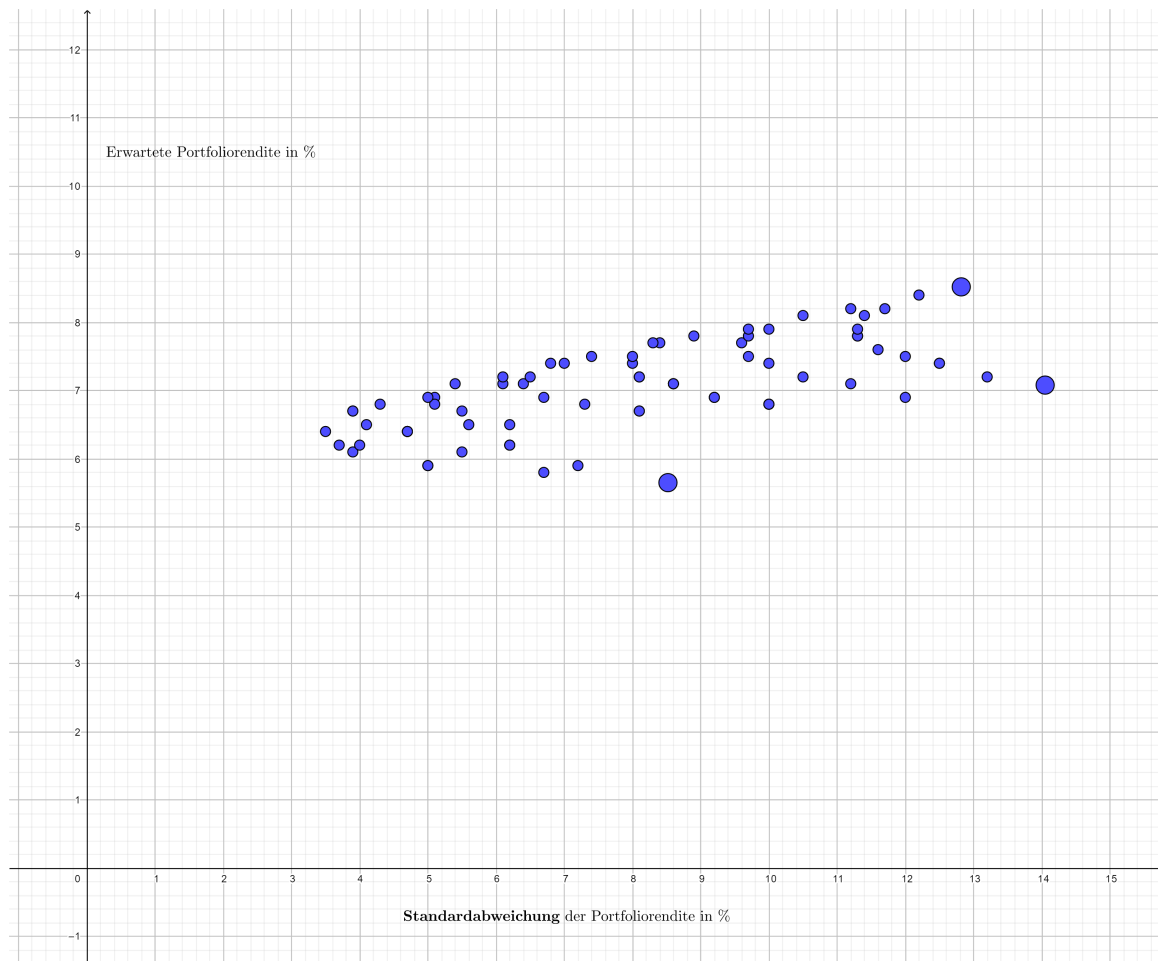
Beispiel 2.1 Wir betrachten drei verschiedenen Aktien, deren relevante Kennzahlen anhand gemessener Daten wie folgt abgeschätzt werden:

1. $\hat{\mu}_1 = 0.0852$ und $\hat{\sigma}_1^2 = 0.0164$,
2. $\hat{\mu}_2 = 0.0565$ und $\hat{\sigma}_2^2 = 0.0073$,
3. $\hat{\mu}_3 = 0.0708$ und $\hat{\sigma}_3^2 = 0.0197$,
4. $\hat{\sigma}_{12} = -0.0044$, $\hat{\sigma}_{13} = 0.0076$ und $\hat{\sigma}_{23} = -0.0091$.

In der Zeichnung sehen Sie einige Portfolios aus den drei Aktien. Die Beschriftung an jedem Punkt entspricht der Gewichtung der drei Aktien. Auf den Axen können Sie erwartete Rendite und Standardabweichung des Portfolios ablesen.



Hier noch ein weiteres Bild einiger Portfolios:



Für die erwartete Rendite und die Varianz der Portfoliorendite gilt:

$$\hat{\mu}_P = x_1 \cdot 0.0852 + x_2 \cdot 0.0565 + x_3 \cdot 0.0708$$

$$\hat{\sigma}_P^2 = (x_1 \ x_2 \ x_3) \begin{pmatrix} 0.0164 & -0.0044 & 0.0076 \\ -0.0044 & 0.0073 & -0.0091 \\ 0.0076 & -0.0091 & 0.0197 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

3 Input-Output-Modelle

Einführung Wir betrachten n verschiedene Industrien (z.B. Stahlindustrie, kohleförderung und Erzförderung) in einem geschlossenen System.

Dann hängt einerseits der Ausstoss (Output) z.B. der Stahlindustrie vom Eingang (Input) der erzeugten Güter aus allen n Industrien ab, denn um eine Tonne Stahl zu produzieren, benötigt man (mindestens) Erz, Kohle und auch Stahl einer bestimmten Menge. Andererseits wird natürlich auch der Output von Stahl von (fast) allen anderen Industrien als Input benötigt.

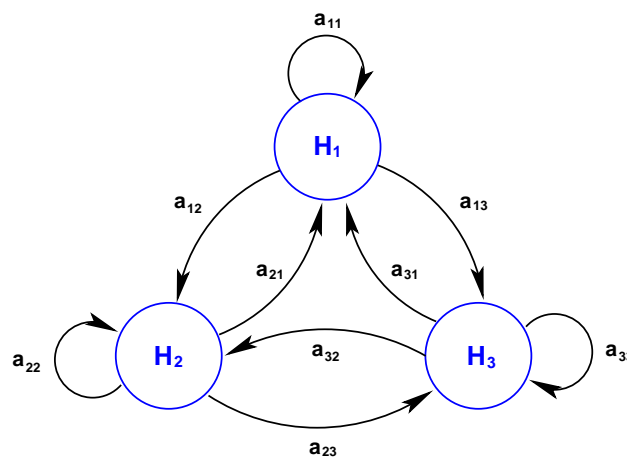
Im Allgemeinen sollte man hier natürlich darum bemüht sein, dass die Outputmengen aller n Industrien zu den Inputmengen „passen“, d.h. dass man unter anderem nicht mehr und nicht weniger Stahl produzieren sollte, wie sämtliche Industriezweige benötigen.

Wir wollen zunächst einige vereinfachende Annahmen treffen:

1. Jeder Industriezweig produziert genau ein Gut.
2. Jeder Industriezweig benötigt ein festes Verhältnis von Inputgütern für die Produktion eines bestimmten Outputs.
3. Die Produktion in jedem Industriezweig geschieht linear homogen, d.h. eine k -fache Erhöhung des (gesamten) Inputs bringt auch einen k -fachen Output des entsprechenden Gutes hervor.

Mathematisches Modell Es seien die n verschiedenen Industrien H_1, H_2, \dots, H_n gegeben. Die Produktionsmengen (z.B. gemessen in kg, Tonnen, Stück, ...) sämtlicher Industrien werden zunächst **einheitlich** in CHF (als Preis der produzierten Menge) angegeben. Weiterhin sei a_{ij} der Preis, für den Industrie H_j von H_i Waren einkaufen muss, um selbst seine eigene Ware im Wert von 1.– CHF (innerhalb einer festen Zeitperiode) produzieren zu können.

Für $n = 3$ Industrien könnte man den Warenaustausch durch das folgende Diagramm beschreiben:



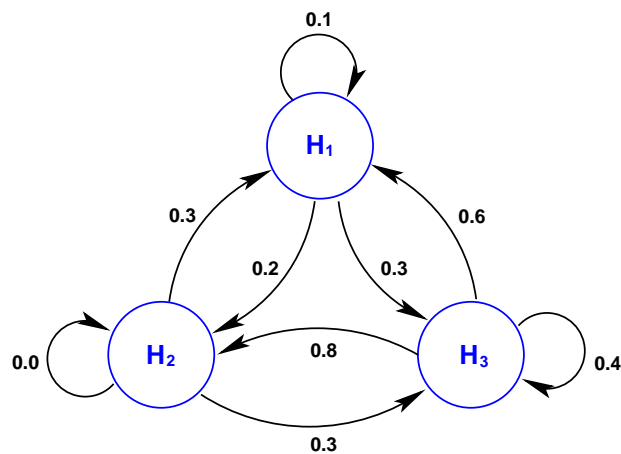
		Einkauf		
		H_1	H_2	H_3
Verkauf	H_1	a_{11}	a_{12}	a_{13}
	H_2	a_{21}	a_{22}	a_{23}
	H_3	a_{31}	a_{32}	a_{33}
		1	1	1

Die Matrix

$$A = (a_{ij}) = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

beschreibt die kompletten wirtschaftlichen Beziehungen der beteiligten Industrien.

Beispiel 3.1 Wir betrachten $n = 3$ Industrien, deren Warenaustausch durch das folgende Diagramm beschrieben wird:



		Einkauf		
		H_1	H_2	H_3
Verkauf	H_1	0.1	0.2	0.3
	H_2	0.3	0.0	0.3
	H_3	0.6	0.8	0.4
		1	1	1

Die den Produktaustausch beschreibende Matrix A hat damit die Gestalt

$$A = \begin{pmatrix} 0.1 & 0.2 & 0.3 \\ 0.3 & 0.0 & 0.3 \\ 0.6 & 0.8 & 0.4 \end{pmatrix}$$

In dem Vektor

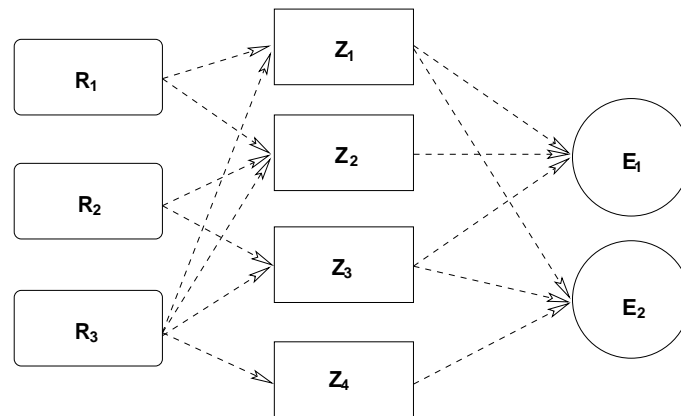
$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

bezeichne x_i die Gesamtproduktion der Industrie H_i . Damit der Wirtschaftskreislauf reibungslos und ohne Überschüsse zu produzieren, funktioniert, müssen wir die einzelnen Industrien passend dimensionieren. Mathematisch bedeutet das, zur Matrix A einen so genannten Eigenvektor \mathbf{x} zum Eigenwert 1 zu finden.

Wesentliche Teile der Input-Outputanalyse wurden vom russisch-amerikanischen Wirtschaftswissenschaftler *Wassily Leontief* in den 1930-er Jahren an der Harvard-University in Cambridge (Massachusetts) entwickelt. Wassily Leontief wurde im Jahr 1973 mit dem Nobelpreis für Wirtschaftswissenschaften ausgezeichnet.

4 Mehrstufige Produktionsprozesse

Einführung Bei mehrstufigen Produktionsprozessen werden aus Rohstoffen Zwischenprodukte und letztendlich Endprodukte hergestellt. Solche Prozesse lassen sich elegant und übersichtlich durch Matrizen beschreiben, was wir am Beispiel eines zweistufigen Produktionsprozesses erläutern wollen. Wir verwenden durchgängig die Abkürzung ME für Mengeneinheiten.



- Aus den drei Rohstoffen R_1 , R_2 und R_3 werden vier Zwischenprodukte Z_1 , Z_2 , Z_3 und Z_4 hergestellt. Dabei benötigt man

für Z_1 : 11 ME von R_1
7 ME von R_3

für Z_2 : 8 ME von R_1
3 ME von R_2
5 ME von R_3

für Z_3 : 12 ME von R_2
8 ME von R_3

für Z_4 : 10 ME von R_3

- Zur Herstellung der Endprodukte E_1 und E_2 benötigt man

für E_1 : 5 ME von Z_1
3 ME von Z_2
6 ME von Z_3

für E_2 : 3 ME von Z_1
2 ME von Z_3
10 ME von Z_4

Mathematisches Modell Wir wollen die Zusammenhänge zwischen Rohstoffen, Zwischenprodukten und Endprodukten durch geeignete Matrizen beschreiben.

Zunächst führen wir die folgenden Abkürzungen ein:

- r_i bezeichne die ME der benötigten Rohstoffe R_i ($i = 1, 2, 3$)
- z_j bezeichne die ME der Zwischenprodukte Z_j ($j = 1, 2, 3, 4$)
- e_k bezeichne die ME der Endprodukte E_k ($k = 1, 2$)

Daraus konstruieren wir die folgenden drei Vektoren:

- $\mathbf{r} = \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix}$ den Rohstoffvektor.
- $\mathbf{z} = \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{pmatrix}$ den Zwischenproduktvektor.
- $\mathbf{e} = \begin{pmatrix} e_1 \\ e_2 \end{pmatrix}$ den Endproduktvektor.

Für den Zusammenhang zwischen Rohstoffen und Zwischenprodukten gilt zunächst:

$$\begin{aligned} r_1 &= 11z_1 + 8z_2 + 0z_3 + 0z_4 \\ r_2 &= 0z_1 + 3z_2 + 12z_3 + 0z_4 \\ r_3 &= 7z_1 + 5z_2 + 8z_3 + 10z_4 \end{aligned}$$

bzw. in Matrixschreibweise:

$$\mathbf{r} = \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} = \underbrace{\begin{pmatrix} 11 & 8 & 0 & 0 \\ 0 & 3 & 12 & 0 \\ 7 & 5 & 8 & 10 \end{pmatrix}}_{=: A} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{pmatrix}$$

Die Matrix A trägt den Namen Rohstoff-Zwischenprodukt-Matrix und es gilt

$$\mathbf{r} = A\mathbf{z}$$

Entsprechend gilt für den Zusammenhang zwischen Zwischenprodukt und Endprodukt:

$$\begin{aligned} z_1 &= 5e_1 + 3e_2 \\ z_2 &= 3e_1 + 0e_2 \\ z_3 &= 6e_1 + 2e_2 \\ z_4 &= 0e_1 + 10e_2 \end{aligned}$$

bzw. in Matrizenschreibweise:

$$\mathbf{z} = \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{pmatrix} = \underbrace{\begin{pmatrix} 5 & 3 \\ 3 & 0 \\ 6 & 2 \\ 0 & 1 \end{pmatrix}}_{=: B} \begin{pmatrix} e_1 \\ e_2 \end{pmatrix}$$

Die Matrix B trägt den Namen Zwischenprodukt-Endprodukt-Matrix und es gilt

$$\mathbf{z} = B\mathbf{e}$$

Setzt man nun beide linearen Beziehungen ineinander ein, erhält man:

$$\mathbf{r} = A\mathbf{z} = A(B\mathbf{e}) = \underbrace{AB}_{=: C} \mathbf{e}$$

mit der Rohstoff-Endprodukt-Matrix $C = AB$. Es ist

$$C = AB = \begin{pmatrix} 11 & 8 & 0 & 0 \\ 0 & 3 & 12 & 0 \\ 7 & 5 & 8 & 10 \end{pmatrix} \begin{pmatrix} 5 & 3 \\ 3 & 0 \\ 6 & 2 \\ 0 & 10 \end{pmatrix} = \begin{pmatrix} 79 & 33 \\ 81 & 24 \\ 98 & 137 \end{pmatrix}$$

und allgemein

$$\mathbf{r} = C\mathbf{e}$$

5 Übungsaufgabe

1. In einem Kaufhauskonzern mit $n = 4$ Filialen sollen die Auswirkungen von Werbeausgaben X_1 (in 1'000.- CHF als Einheit) auf die Umsatzsteigerungen Y (in 10'000.- CHF als Einheit) untersucht werden.

Univariates lineares Regressionsmodell: $Y = \beta_0 + \beta_1 X_1 + \epsilon$

Messungen

i	y_i	x_i
1	2	1.5
2	3	2
3	6	3.5
4	5	2.5

Bestimmen Sie die folgenden Grössen:

$n = 4$ und $k = 1$

$$\mathbf{y} =$$

$$\mathbf{X} =$$

$$\mathbf{X}^T \mathbf{X} =$$

$$(\mathbf{X}^T \mathbf{X})^{-1} =$$

$$\mathbf{b} =$$

Regressionsgleichung: $Y =$

$$\hat{\mathbf{y}} =$$

$$\mathbf{e} =$$

$$\text{var}(\mathbf{y}) =$$

$$\text{var}(\hat{\mathbf{y}}) =$$

$$\text{var}(\mathbf{e}) =$$