

Univ.-Prof. Dr. Andreas Kleine
Univ.-Prof. Dr. Robinson Kruse-Becher
Univ.-Prof. Dr. Hermann Singer
Dipl.-Stat. Anja Bittner
Sebastian Förster, M.Sc.

Modul 32741

Vertiefung der Wirtschafts- mathematik und Statistik

Leseprobe

Fakultät für
**Wirtschafts-
wissenschaft**

Das Werk ist urheberrechtlich geschützt. Die dadurch begründeten Rechte, insbesondere das Recht der Vervielfältigung und Verbreitung sowie der Übersetzung und des Nachdrucks, bleiben, auch bei nur auszugsweiser Verwertung, vorbehalten. Kein Teil des Werkes darf in irgendeiner Form (Druck, Fotokopie, Mikrofilm oder ein anderes Verfahren) ohne schriftliche Genehmigung der FernUniversität reproduziert oder unter Verwendung elektronischer Systeme verarbeitet, vervielfältigt oder verbreitet werden.

Vertiefung der Analysis und Linearen Algebra

Fakultät für
**Wirtschafts-
wissenschaft**

1.2 Determinanten

1.2.1 Einführung von Determinanten

Eine Matrix kann durch verschiedene Kennzahlen charakterisiert werden. Für jede Matrix kann z.B. der Rang angegeben werden, welcher der maximalen Anzahl linear unabhängiger Zeilen bzw. Spalten entspricht.

Grundlagen 1.8. (Rang einer Matrix) [G: 5.2.3]

Die maximale Anzahl linear unabhängiger Spaltenvektoren bzw. Zeilenvektoren einer Matrix \mathbf{A} heißt **Rang der Matrix \mathbf{A}** und wird mit $rg(\mathbf{A})$ bezeichnet.

Die sogenannte Determinante ist eine weitere Kennzahl, welche jedoch nur für quadratische Matrizen existiert. Mit Hilfe der Determinante lässt sich z.B. prüfen, ob eine quadratische Matrix invertierbar ist. Gegebenenfalls kann dann die Inverse der Matrix bestimmen werden. Darüber hinaus bildet die Determinante eine wichtige Grundlage für die Lösung von Eigenwertproblemen (vgl. Abschnitt 1.3) bei quadratischen Matrizen sowie für die Differentialrechnung von reellen Funktionen mehrerer Variablen (vgl. Einheit 2).

Für jede quadratische Matrix $\mathbf{A} = (a_{ij})_{1 \leq i, j \leq n} \in \mathbb{R}^{n \times n}$ ist die n -reihige **Determinante** von \mathbf{A} definiert durch

$$\det \mathbf{A} = \det \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} := \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} \in \mathbb{R}.$$

Bei der Determinante von \mathbf{A} handelt es sich um eine reelle Zahl. Für den Fall $n = 1$ wird üblicherweise $\det(a_{11})$ anstelle von $|a_{11}|$ geschrieben, um eine klare Unterscheidung von den Betragsstrichen zu gewährleisten. Es gilt $\det(a_{11}) = a_{11}$ für alle $a_{11} \in \mathbb{R}$, also z.B. $\det(-4) = -4$.

Auf einen wichtigen Zusammenhang zwischen einer quadratische Matrix und ihrer Determinante sei bereits an dieser Stelle hingewiesen: Die Existenz einer Inversen kann mit Hilfe der Determinante überprüft werden.

Eine (n, n) -Matrix \mathbf{A} ist genau dann **invertierbar (regulär)**, wenn ihre Determinante ungleich Null ist: $\det \mathbf{A} \neq 0$.

Folglich gilt im Fall $\det \mathbf{A} \neq 0$ unmittelbar $rg(\mathbf{A}) = n$. Mit Hilfe der Determinante lässt sich somit ebenfalls prüfen, ob Vektoren linear unabhängig sind.

1.2.2 Berechnung von Determinanten

Wie die Determinante von einer quadratischen Matrix berechnet wird, wird in diesem Abschnitt zunächst für eine 2-reihige, dann für eine 3-reihige und schließlich allgemein für eine n -reihige Determinante gezeigt.

Die 2-reihige Determinante

Die Berechnung der Determinante für (2,2)-Matrizen ergibt sich aus der Differenz des Produkts der Elemente in der Hauptdiagonalen und des Produkts der Elemente in der Nebendiagonalen.

Für die Determinante einer (2,2)-Matrix \mathbf{A} gilt:

$$\det \mathbf{A} = \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11} \cdot a_{22} - a_{12} \cdot a_{21}.$$

Beispiel 1.1

- a) Gegeben sei die Matrix $\mathbf{A} = \begin{pmatrix} 3 & 1 \\ 5 & 2 \end{pmatrix}$. Für die Determinante gilt:

$$\det \mathbf{A} = \begin{vmatrix} 3 & 1 \\ 5 & 2 \end{vmatrix} = 3 \cdot 2 - 1 \cdot 5 = 1.$$

Somit hat die Matrix \mathbf{A} einen vollen Rang $rg(\mathbf{A}) = 2$ und ist invertierbar.

- b) Die Matrix $\mathbf{B} = \begin{pmatrix} 2 & 4 \\ 1 & 2 \end{pmatrix}$ ist nicht regulär, d.h. nicht invertierbar, denn für die Determinante gilt:

$$\det \mathbf{B} = \begin{vmatrix} 2 & 4 \\ 1 & 2 \end{vmatrix} = 2 \cdot 2 - 4 \cdot 1 = 0.$$

 L-13

Bestimmen Sie den Rang der Matrix \mathbf{B} (vgl. [G: 5.2.3]). Was fällt Ihnen auf?

- c) Gegeben sei die Matrix $\mathbf{C} = \begin{pmatrix} a & 3 \\ 2 & 6 \end{pmatrix}$ mit $a \in \mathbb{R}$. Für die Determinante gilt:

$$\det \mathbf{C} = \begin{vmatrix} a & 3 \\ 2 & 6 \end{vmatrix} = a \cdot 6 - 3 \cdot 2 = a \cdot 6 - 6.$$

Somit ist die Matrix \mathbf{C} für $a \neq 1$ invertierbar, denn es gilt:

$$\det \mathbf{C} = 0 \quad \iff \quad a \cdot 6 - 6 = 0 \quad \iff \quad a = 1.$$

2.2 Zulässige Lösungen linearer Programme

Die Menge der zulässigen Lösungen eines linearen Programms weist einige Besonderheiten auf. So ist der Zulässigkeitsbereich X eines LPs konvex und besitzt sogenannte Ecken. Was diese Eigenschaften beinhalten und warum diese zur Bestimmung einer optimalen Lösung eines LPs von Bedeutung sind, ist Gegenstand dieses Kapitels.

2.2.1 Konvexer Zulässigkeitsbereich

Der *Zulässigkeitsbereich* eines LPs kann äquivalent zur komprimierten Matrix-
Zulässigkeitsbereich
schreibweise in (2.3) auf Seite 57 auch geschrieben werden als:

$$X = \left\{ \mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n \left| \begin{array}{l} a_{11} x_1 + \dots + a_{1n} x_n \leq b_1 \\ \vdots \\ a_{m1} x_1 + \dots + a_{mn} x_n \leq b_m \\ x_1, \dots, x_n \geq 0 \end{array} \right. \right\}. \quad (2.8)$$

Zulässig sind demnach alle nichtnegativen Vektoren \mathbf{x} , die jeder der m Nebenbedingungen genügen. Diese Vektoren $\mathbf{x} \in X$ heißen *zulässige Lösungen von X* . Für die Lösungen müssen alle m Ungleichungen simultan und nicht nur
zulässige Lösung
einzelne Ungleichungen erfüllt sein. Graphisch kann diese Bedingung am Zulässigkeitsbereich eines LPs mit zwei Variablen veranschaulicht werden. Hierzu wird das Beispiel 2.1 erneut aufgegriffen.

Beispiel 2.5

Der Zulässigkeitsbereich für Beispiel 2.1, bei welchem zwei Produkte P1 und P2 herzustellen sind, lautet:

$$X = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \left| \begin{array}{l} x_1 + x_2 \leq 180 \quad \text{(MI)} \\ x_1 + 2x_2 \leq 240 \quad \text{(MII)} \\ 2x_2 \leq 180 \quad \text{(MIII)} \\ x_1, x_2 \geq 0 \quad \text{(NNB)} \end{array} \right. \right\}.$$

Gemäß dieser Darstellungsweise durchlaufen die beiden Produkte zunächst Maschine MI und dann Maschine MII. Produkt P2 durchläuft schließlich noch Maschine MIII. Es sind alle drei Nebenbedingungen zu berücksichtigen, da sonst die begrenzte Kapazität einer Maschine vernachlässigt wird.

Der Zulässigkeitsbereich der beiden Variablen x_1 und x_2 kann in einem entsprechenden Koordinatensystem graphisch dargestellt werden. Dazu werden in Abbildung 2.2 schrittweise die einzelnen Nebenbedingungen ergänzt. Aufgrund der NNB ist nur der erste Quadrant des Koordinatensystems zu betrachten (links oben). Einzeichnen der Restriktion MI schränkt die Menge der zulässigen Lösungen bereits

deutlich ein (rechts oben). Wie die beiden Nebenbedingungen für MII und MIII diese Menge noch weiter eingrenzen, zeigen die unteren beiden Abbildungen. Der Zulässigkeitsbereich X ist schließlich in der unteren rechten Abbildung dargestellt.

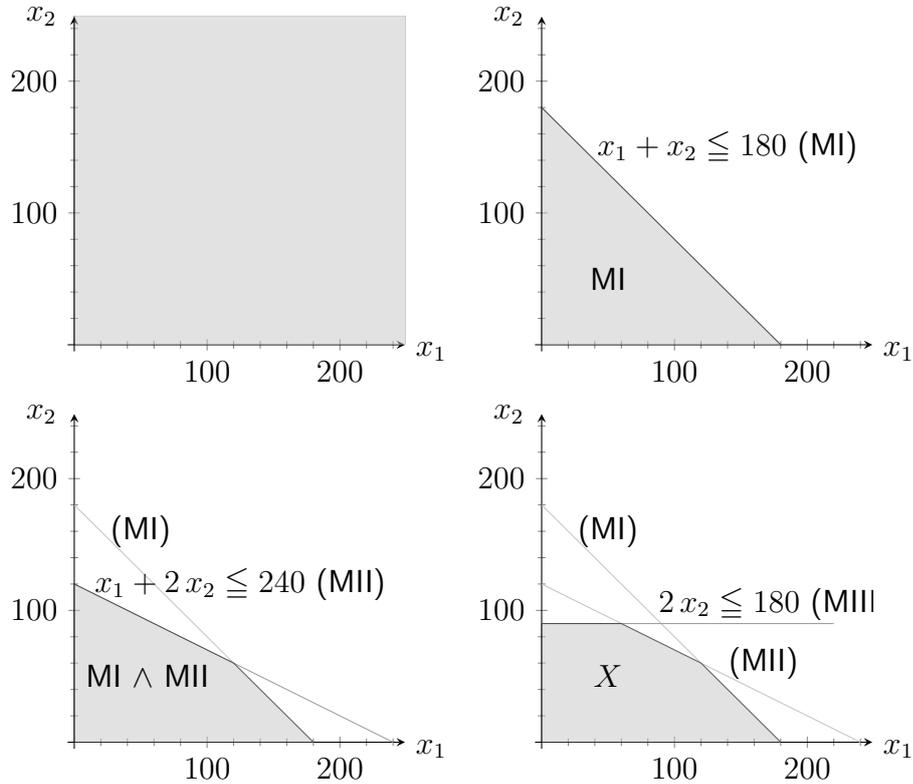


Abbildung 2.2: Nebenbedingungen des Zulässigkeitsbereichs X

Ergänzen Sie in der Abbildung 2.2 (rechts unten) die zusätzlichen Restriktion $x_1 \leq 200$. Was fällt Ihnen auf?

 L-25

Eine Nebenbedingung mit den Eigenschaften, wie beispielsweise die Restriktion $x_1 \leq 200$ in Abbildung 2.2 oder die Restriktion (3.) aus der Übungsaufgabe 2.4 aufweist, wird als *redundant* bezeichnet. Durch eine *redundante Nebenbedingung* i wird der Zulässigkeitsbereich X eines LPs nicht weiter eingeschränkt. D.h. für $i \in \{1, \dots, m\}$ gilt:

$$X \cap \left\{ (x_1, \dots, x_n)^T \in \mathbb{R}^n \mid \sum_{j=1}^n a_{ij} x_j \leq b_i \right\} = X. \quad (2.9)$$

Wie Beispiel 2.5 zeigt, kann der Zulässigkeitsbereich für Probleme mit zwei Variablen anschaulich graphisch dargestellt werden. Bereits bei drei Variablen wird dieses allerdings wesentlich aufwendiger (vgl. X_b in Übungsaufgabe 2.5). In höherdimensionalen Beispielen ist die graphische Visualisierung deshalb meist nicht mehr möglich.

redundante
Nebenbedingung

5 Einführung in Differential- und Differenzengleichungen

Vorbemerkungen

Wenn sich eine Krankheit ungehindert ausbreitet, so wird häufig von einem exponentiellen Wachstum gesprochen. Die Ermittlung einer Funktion zur Bestimmung der Anzahl an erkrankten Personen zu jedem möglichen Zeitpunkt führt zur Theorie der Differentialgleichungen. Bei vielen ökonomischen Modellen sind Funktionen implizit in Form von Differential- oder Differenzengleichungen gegeben, insbesondere im Zusammenhang mit Produktions- und Nutzenfunktionen sowie Wachstums- und Marktprozessen. Dabei handelt es sich um Gleichungen, die eine Beziehung zwischen einer Funktion und ihren Ableitungen oder zwischen einer Folge und ihren Differenzenfolgen darstellen. Gesucht sind somit die Funktionen bzw. Folgen, die diesen Gleichungen genügen.

In diesem Kapitel wird in die Theorie der Differential- und Differenzengleichungen eingeführt und ihre Anwendung in den Wirtschaftswissenschaften aufgezeigt. Weiterhin werden einige elementare Lösungsmethoden behandelt, die einen ersten allgemeinen Einblick vermitteln.

Differentialgleichungen dienen der Beschreibung stetiger ökonomischer Prozesse. Differenzengleichungen kommen hingegen bei diskreten Prozessen zur Anwendung, da ökonomische Größen häufig nur für konstante Zeitintervalle zur Verfügung stehen. Sobald bei Prozessen von der diskreten zur stetigen Betrachtungsweise gewechselt wird, gehen Differenzengleichungen in Differentialgleichungen über. Umgekehrt ergeben sich Differenzengleichungen durch die Diskretisierung von Differentialgleichungen. Der enge Zusammenhang von Differential- und Differenzengleichungen wird bereits an dieser Stelle deutlich und zeigt sich wiederholt in den Parallelen der Aufgabenstellungen und den Lösungsmethoden.

Ähnlichkeitsdifferentialgleichungen

Eine weitere elementare Lösungsmethode existiert für Ähnlichkeitsdifferentialgleichungen. Diese DGLn lassen sich in separable DGLn überführen und schließlich mit der Methode „Trennung der Variablen“ lösen. Bezeichnet werden diese DGLn in der Literatur auch häufig als *homogene* Differentialgleichungen.

Definition 5.4.

Eine gewöhnliche DGL erster Ordnung heißt **Ähnlichkeitsdifferentialgleichung**, wenn sie sich in der Form

$$y' = g\left(\frac{y}{x}\right) \quad \text{bzw.} \quad y'(x) = g\left(\frac{y(x)}{x}\right) \quad (5.14)$$

darstellen lässt. Dabei ist g eine auf einem Intervall definierte, stetige Funktion.

Um Ähnlichkeitsdifferentialgleichungen zu lösen, erfolgt zuerst die Substitution

$$z(x) := \frac{y(x)}{x} \quad \text{bzw.} \quad y(x) = x \cdot z(x) \quad (5.15)$$

mit der entsprechenden Ableitung gemäß der Produktregel:

$$y'(x) = z(x) + x \cdot z'(x). \quad (5.16)$$

Grundlagen 5.7. (Produktregel) [G: 2.3.3]

$$\left(f(x) \cdot g(x)\right)' = f'(x) \cdot g(x) + f(x) \cdot g'(x)$$

Im Folgenden wird zur besseren Übersichtlichkeit die Schreibweise ohne Argument verwendet, d.h. $y' = z + x \cdot z'$. Einsetzen von (5.15) und (5.16) in (5.14) ergibt die DGL $z + x \cdot z' = g(z)$.

Um die Lösungsmethode „Trennung der Variablen“ anwenden zu können, ist zunächst in Schritt T1) eine Substitution von $z = y/x$ notwendig, die dann in Schritt T5) wieder rückgängig gemacht werden muss (Resubstitution). Dieses Vorgehen wird im folgenden Beispiel veranschaulicht.

Beispiel 5.13

Gesucht sei die allgemeine Lösung der DGL $y' = \frac{y}{x} + \left(\frac{y}{x}\right)^3$.

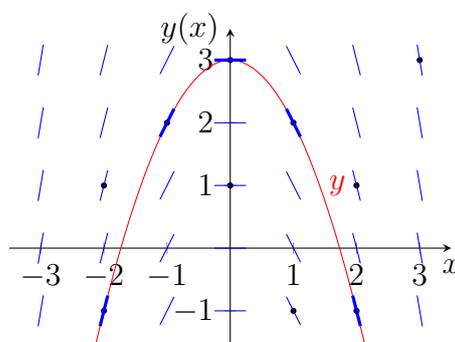
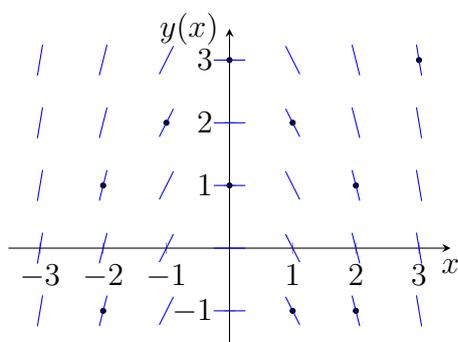
*0) Typus: Ähnlichkeitsdifferentialgleichung mit $g\left(\frac{y}{x}\right) = \frac{y}{x} + \left(\frac{y}{x}\right)^3$.

Lösungen zu den Übungen

Lösung Übung 5.1 auf S. 180:

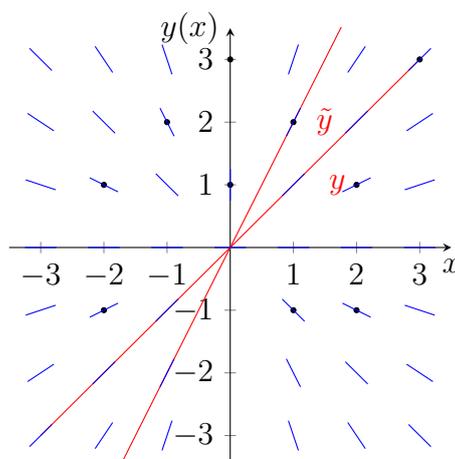
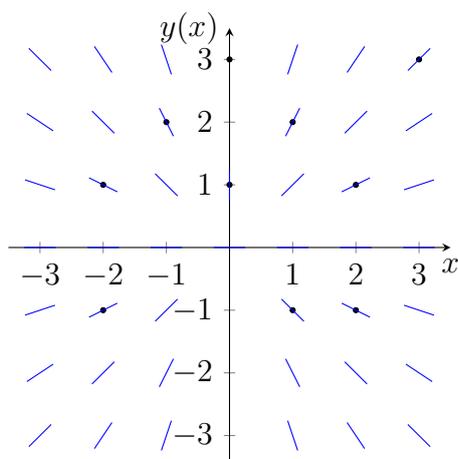
$(x, y(x))$	$(-2, -1)$	$(-2, 1)$	$(-1, 2)$	$(0, 1)$	$(0, 3)$
$y'(x) = -2 \cdot x$	4	4	2	0	0

$(x, y(x))$	$(1, -1)$	$(1, 2)$	$(2, 1)$	$(2, -1)$	$(3, 3)$
$y'(x) = -2 \cdot x$	-2	-2	-4	-4	-6



Richtungsfeld der expliziten DGL $y'(x) = -2 \cdot x$ mit Lösung $y(x) = 3 - x^2$

Lösung Übung 5.2 auf S. 180:



Richtungsfeld der expliziten DGL $y'(x) = \frac{y}{x}$ mit den Lösungen $y(x) = x$ und $\tilde{y}(x) = 2x$.

Aus dem Vergleich möglicher Lösungskurven folgt, dass die allgemeine Lösung die Form $y = c \cdot x$ mit $c \in \mathbb{R}$ besitzt.

Vertiefung der Statistik

Fakultät für
**Wirtschafts-
wissenschaft**

$$P \left\{ \frac{(N-1)S^2}{\chi^2(1-\alpha/2, N-1)} \leq \sigma^2 \leq \frac{(N-1)S^2}{\chi^2(\alpha/2, N-1)} \right\} = 1 - \alpha \quad (2.58)$$

(Konfidenzintervall für die Varianz σ^2).

Ein wichtiger Unterschied zu den Konfidenzintervallen für den Erwartungswert besteht in der Asymmetrie des Konfidenzintervalls für die Varianz. Dies ist auf die asymmetrische χ^2 -Verteilung zurückzuführen.

Beispiel 20 (Inflationserwartungen in den USA im Jahr 2022: Konfidenzintervalle).

Für das Quartal 2022Q1 finden wir:

$$s_1 = 1.893, N_1 = 34, \alpha = 0.10$$

$$\chi^2(\alpha/2, N_1 - 1) = \chi^2(0.05, 33) = 20.867$$

$$\chi^2(1 - \alpha/2, N_1 - 1) = \chi^2(0.95, 33) = 47.400$$

$$\text{Konfidenzintervall für } \sigma_1^2: \left[\frac{(34-1)1.893^2}{47.400}, \frac{(34-1)1.893^2}{20.867} \right] = [2.495, 5.667]$$

Analog ergeben sich die Konfidenzintervalle für die anderen Quartale.

■

2.4.6 Konfidenzintervall für den Korrelationskoeffizienten ρ : Normalverteilte Grundgesamtheit

Der Korrelationskoeffizient zwischen 2 Zufallsvariablen X, Y ist ein Maß für den linearen Zusammenhang, bei normalverteilten Größen folgt sogar die Unabhängigkeit aus der Unkorreliertheit der Merkmale.

Die Unkorreliertheit impliziert jedoch nicht unbedingt die Unabhängigkeit, etwa bei einem parabelförmigen Zusammenhang ($Y = X^2$; unkorreliert, jedoch perfekt abhängig).

Zunächst ist die Korrelation ganz allgemein die normierte Kovarianz der Zufallsvariablen, d. h.

$$\rho = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)}\sqrt{\text{Var}(Y)}} \quad (2.59)$$

Ein Punktschätzer für ρ ist durch die Produkt-Moment-Korrelation

$$R = \frac{S(X, Y)}{S(X)S(Y)} \quad (2.60)$$

gegeben, wobei die Stichproben(ko)varianzen durch

$$S(X, Y) = \frac{1}{N-1} \sum_{n=1}^N (X_n - \bar{X})(Y_n - \bar{Y}) \quad (2.61)$$

$$S^2(X) = \frac{1}{N-1} \sum_{n=1}^N (X_n - \bar{X})^2 = S(X, X) \quad (2.62)$$

$$S^2(Y) = \frac{1}{N-1} \sum_{n=1}^N (Y_n - \bar{Y})^2 = S(Y, Y) \quad (2.63)$$

definiert sind.

Explizit kann man auch

Produkt-Moment-Korrelation

$$R = \frac{\sum_{n=1}^N (X_n - \bar{X})(Y_n - \bar{Y})}{\sqrt{\sum_{n=1}^N (X_n - \bar{X})^2 \sum_{n=1}^N (Y_n - \bar{Y})^2}} \quad (2.64)$$

schreiben.

Fisher Z-Transformation

Wenn die Zufallsvariablen (X_n, Y_n) bivariat normalverteilt sind, so ist die transformierte Zufallsvariable

$$Z = \frac{\sqrt{N-3}}{2} \left(\ln \frac{1+R}{1-R} - \ln \frac{1+\rho}{1-\rho} \right) \quad (2.65)$$

für große Stichproben $N(0, 1)$ -verteilt (**Fisher Z-Transformation**).

Somit ergibt sich das Intervall $[z = z(1 - \alpha/2)]$

$$P \{-z \leq Z \leq z\} = 1 - \alpha, \quad (2.66)$$

das man wie immer nach dem Parameter von Interesse, hier ρ , umformen kann:

$$P \left\{ \frac{e^A - 1}{e^A + 1} \leq \rho \leq \frac{e^B - 1}{e^B + 1} \right\} = 1 - \alpha, \quad (2.67)$$

wobei

$$A = \ln \frac{1 + R}{1 - R} - \frac{2z}{\sqrt{N - 3}} \quad (2.68)$$

$$B = \ln \frac{1 + R}{1 - R} + \frac{2z}{\sqrt{N - 3}}. \quad (2.69)$$

Beispiel 21 (Neu- und Gebrauchtwagenpreise).

Wir betrachten den linearen Zusammenhang von Neu- und Gebrauchtwagenpreisänderungen (Änderung in Prozent vom Vorjahreswert) in den USA im Zeitraum von Januar 1954 bis Oktober 2024 mit $N = 850$ monatlichen Beobachtungen für beide Variablen. Das Streudiagramm ist in Abb. 2.8 gegeben. Der geschätzte Korrelationskoeffizient beträgt $r = 0.579$. Diese Berechnung kann durch die empirische Stichprobenkovarianz von 16.368 und die empirischen Standardabweichungen von 3.252 (Neuwagen) und 8.688 (Gebrauchtwagen) nachvollzogen werden:

$$r = \frac{16.368}{3.252 \times 8.688} = 0.579.$$

Gesucht ist ein 99%-Konfidenzintervall für den unbekanntem Korrelationskoeffizienten ρ .

99.5%-Quantil: $z(0.995) = 2.58$

$$a = \ln \frac{1 + 0.579}{1 - 0.579} - \frac{2 \times 2.58}{\sqrt{847}} = 1.144 \quad (2.70)$$

$$b = 1.499 \quad (2.71)$$

$$\frac{e^a - 1}{e^a + 1} = 0.517 \quad (2.72)$$

$$\frac{e^b - 1}{e^b + 1} = 0.635 \quad (2.73)$$

Daraus findet man das 99%-Konfidenzintervall $[0.517, 0.635]$.

■

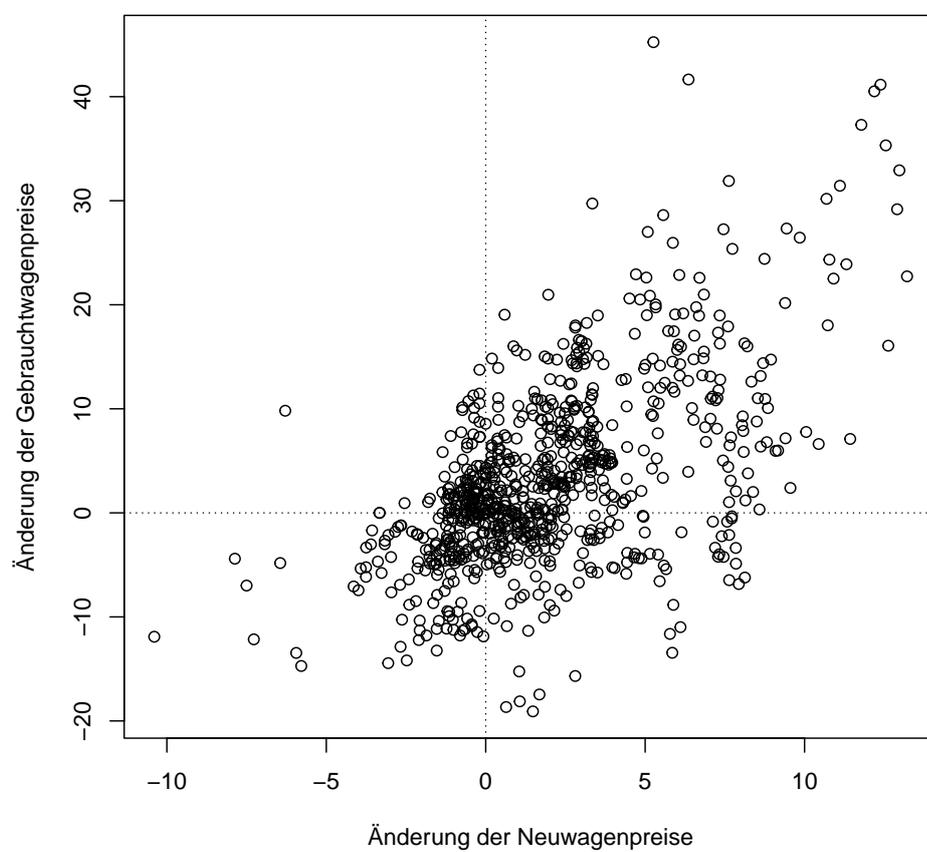


Abbildung 2.8: Änderungen von Neu- und Gebrauchtwagenpreisen.

Aufbau eines Hypothesentests

1. Festlegung einer sogenannten Nullhypothese H_0 und einer Alternativhypothese H_1 .
2. Festlegung eines Signifikanz-Niveaus α :
 $\alpha = P\{H_0 \text{ ablehnen} | H_0 \text{ richtig}\} = \alpha\text{-Fehler}$
3. Bestimmung einer Testfunktion (Teststatistik)
 $T = T(X_1, \dots, X_N)$.
 Die Verteilung der Teststatistik muss von der Hypothese abhängen und wird unter H_0 berechnet.
4. Festlegung eines Verwerfungs (Ablehnungs)-Bereichs R mit
 $P\{T \in R | H_0\} = \alpha$.
5. Entscheidungsregel:
 Wenn $T \in R$ wird H_0 abgelehnt (Realisation $T = t$ einsetzen).

Bemerkungen:

- Die Notation $P\{A|H_0\}$ bedeutet, dass die Wahrscheinlichkeit des Ereignisses A unter Gültigkeit der Hypothese H_0 berechnet wird. Es handelt sich nicht um eine bedingte Wahrscheinlichkeit, da H_0 kein Ereignis ist.
- Die Wahrscheinlichkeit, H_0 abzulehnen, wenn H_0 richtig ist, soll klein sein (z. B. $\alpha = 10\%, 5\%, 1\%$).
- Die Hypothese selbst hat keine Wahrscheinlichkeit. Man kann also nicht sagen: H_0 ist zu 10% falsch oder H_1 ist zu 95% richtig, etc.
- Der Ablehnungsbereich (rejection region) R ist eine Untermenge des Parameter-Raums (Wertebereich des Parameters, etwa $-1 \leq \rho \leq 1$, $-\infty < \mu < \infty$, $0 < \sigma^2 < \infty$).

Da der Test für den Korrelationskoeffizienten etwas komplizierter ist (es wird eine transformierte Testgröße benutzt), hier der Test für den Erwartungswert μ :

Beispiel 22 (Test des Erwartungswerts μ ; σ^2 bekannt).

Wir betrachten die Inflationserwartungen in den USA im vierten Quartal 2022. Folgende Informationen liegen vor: $\bar{x}_4 = 4.459\%$ (Mittelwert über die vorliegenden Prognosen) basierend auf $N_4 = 35$ (Anzahl der vorliegenden individuellen

Prognosen). Die Standardabweichung ist $s_4 = 1.396$. Es wird unterstellt, dass s mit der tatsächlichen Standardabweichung σ übereinstimmt (ansonsten führt man einen t -Test durch).

Ablauf des Tests:

1. Festlegung einer Nullhypothese $H_0 : \mu = \mu_0 = 4$ (Prozent) und einer Alternativhypothese $H_1 : \mu \neq \mu_0 = 4$.

2. Festlegung eines Signifikanz-Niveaus $\alpha = 0.05$:
 $\alpha = P\{H_0 \text{ ablehnen} | H_0 \text{ richtig}\} = \alpha$ -Fehler

3. Bestimmung einer Testfunktion (Teststatistik)

$$T = T(X_1, \dots, X_N) = \bar{X}$$

(Mittelwert).

Es gilt unter H_0 :

$$E[T|H_0] = E[\bar{X}|H_0] = \mu_0 = 6$$

Die Verteilung der Teststatistik hängt somit, wie gefordert, von der Nullhypothese ab.

4. Verwerfungsbereich R festlegen:

$$P\{T \in R | H_0\} = P\{\bar{X} \notin [\mu_0 - z \sigma / \sqrt{N}, \mu_0 + z \sigma / \sqrt{N}] | H_0\} = \alpha$$

bzw. für die standardisierte Teststatistik $Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{N}}$

$$P\{|Z| > z | H_0\} = \alpha$$

[Quantil $z = z(1 - \alpha/2) = 1.96$].

Ablehnungsbereich:

$$\begin{aligned} R &= (-\infty, \mu_0 - z \sigma / \sqrt{N}) \cup (\mu_0 + z \sigma / \sqrt{N}, \infty) \\ &= (-\infty, 4 - 1.96 \times 1.396 / \sqrt{35}) \cup (4 + 1.96 \times 1.396 / \sqrt{35}, \infty) \\ &= (-\infty, 3.538) \cup (4.462, \infty) \end{aligned}$$

5. Entscheidungsregel: wenn $t = \bar{x} \in R$, wird H_0 abgelehnt.

Da $4.459 \in [3.538, 4.462] = \bar{R}$ wird $H_0 : \mu = \mu_0 = 4$ beibehalten (kann nicht abgelehnt werden).

Entsprechend wäre das 95%-Konfidenzintervall $[3.997, 4.921]$. Es überdeckt also den Wert unter der Nullhypothese (keine Ablehnung).

■

Kapitel 5

Regressionsanalyse

5.1 Lineares Modell und Kleinst-Quadrate-Schätzung (KQ)

5.1.1 Das lineare Modell

Die Regressionsanalyse zählt zu den wichtigsten Werkzeugen der empirischen Wirtschaftswissenschaft. Der einfachste Ansatz ist das lineare Regressionsmodell (lineare Prognoseregeln)

$$Y_n = \alpha + \beta X_n + \varepsilon_n, n = 1, \dots, N, \quad (5.1)$$

**lineares
Regressionsmodell**

bei dem eine Gerade an die Daten angepasst wird (Abb. 5.1, 5.2). Hierbei wird X als unabhängige Variable (oder auch erklärende Variable bzw. Regressor) und Y als abhängige Variable (zu erklärende Variable bzw. Regressand) bezeichnet.

Da in der Realität nicht alle Daten-Punkte (X_n, Y_n) exakt auf der Geraden liegen, muss das (zu einfache) lineare Modell durch stochastische Gleichungsfehler (oder auch Störgröße) ε_n ergänzt werden. Diese modellieren Spezifikationsfehler des Modells, etwa Abweichungen des Zusammenhangs vom linearen Modell oder weggelassene bzw. nicht erfasste Variablen, die aber zumindest in stochastischer Weise berücksichtigt werden können.

Die Parameter des Modells sind:

- Achsenabschnitt α
- Steigung β
- Varianz des Gleichungsfehlers (Störgröße) $\sigma^2 = \text{Var}(\varepsilon_n)$

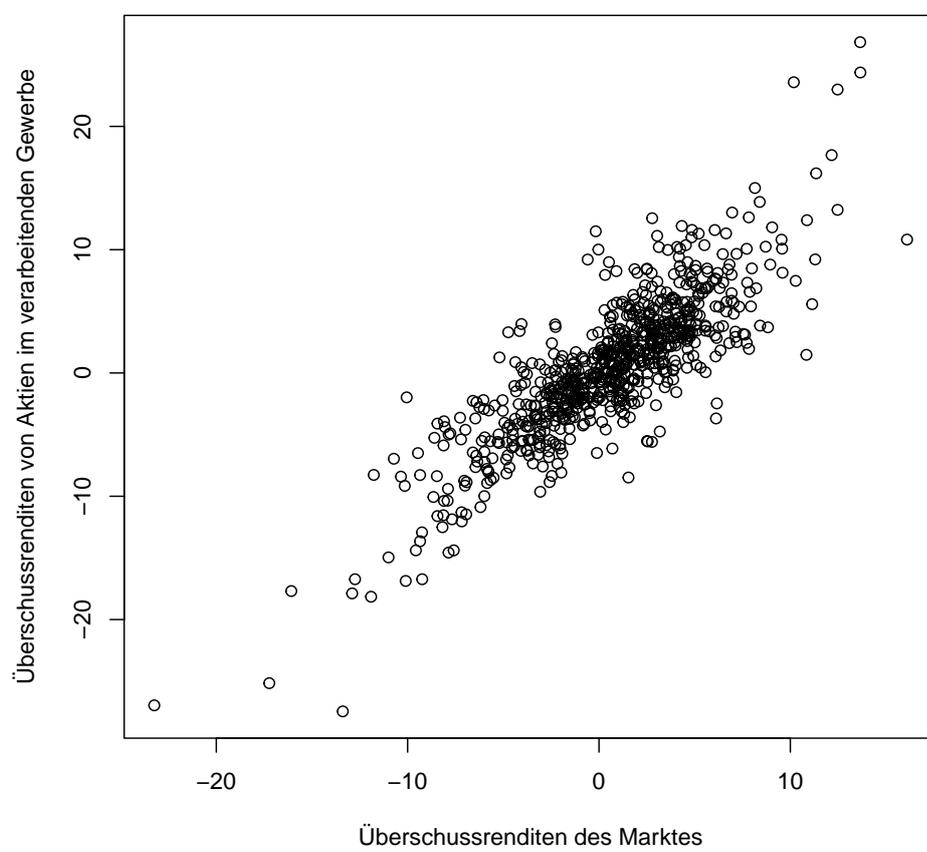


Abbildung 5.1: Streudiagramm der Überschussrenditen.

wahres Modell

Bei Glg. 5.1 handelt es sich allerdings um das wahre Modell, das durch die Populationsparameter α, β und $\sigma^2 = \text{Var}(\varepsilon_n)$ definiert ist. Diese können nur geschätzt werden. Im Bild sieht man die geschätzte Gerade (lineare Prognose)

$$\hat{Y} = \hat{\alpha} + \hat{\beta}X. \quad (5.2)$$

5.1.2 Kleinst-Quadrate-Schätzung (KQ)**Berechnung der Kleinst-Quadrate-Schätzer**

Die Schätzer für α, β und $\sigma^2 = \text{Var}(\varepsilon_n)$ können mit Hilfe des Kleinst-Quadrate (KQ) Methode gefunden werden. Hierbei werden α und β so gewählt, dass die Quadratsumme

$$S(\alpha, \beta) = \sum_n (Y_n - \alpha - \beta X_n)^2 = \sum_n \varepsilon_n^2 \quad (5.3)$$

minimal wird, also die zu schätzende Gerade optimal an die Datenpunkte (X_n, Y_n) angepasst wird.

Differenziert man $S(\alpha, \beta)$ nach α und β und setzt die Ableitungen Null, so ergeben sich die KQ-Schätzer

Kleinst-Quadrate-Schätzer (KQ, OLS)

$$\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{X} \quad (5.4)$$

$$\hat{\beta} = \frac{\sum_n (X_n - \bar{X})(Y_n - \bar{Y})}{\sum_n (X_n - \bar{X})^2} \quad (5.5)$$

$$= R_{xy} \frac{S_y}{S_x}. \quad (5.6)$$

Hierbei ist R_{xy} der Korrelationskoeffizient und $\bar{X} = \frac{1}{N} \sum X_n$ sowie $\bar{Y} = \frac{1}{N} \sum Y_n$ sind die Mittelwerte. Man kann also direkt ersehen, dass zum einen eine analytische Lösung existiert (gegeben eine positive Varianz der Daten X_n) und dass die KQ-Schätzung des Steigungsparameters β dasselbe Vorzeichen wie die Kovarianz zwischen bzw. der Korrelationskoeffizient von Y_n und X_n aufweist. Sind die Daten zentriert, also $\bar{Y} = \bar{X} = 0$, so folgt $\hat{\alpha} = 0$. Der geschätzte Achsenabschnitt ist demnach also auch passenderweise gleich 0. Für den Fall, dass $\hat{\beta} = 0$ gilt — dies tritt auf, wenn die erklärende Variable X aus dem Modell ausgelassen wird (oder die tatsächliche Korrelation zwischen Y_n und X_n exakt 0 ist) und $Y_n = \alpha + \varepsilon_n$ gilt, so ergibt sich direkt $\hat{\alpha} = \bar{Y}$. Die Schätzung für den Parameter α entspricht also dem Mittelwert von Y_n .

Für die Berechnung ist es offenbar hilfreich, zunächst $\hat{\beta}$ und auf dessen Basis dann $\hat{\alpha}$ zu berechnen.

Aus den geschätzten Parametern können die **geschätzten bzw. prognostizierten Y-Werte**

$$\hat{Y}_n = \hat{\alpha} + \hat{\beta}X_n \quad (5.7)$$

$$= \bar{Y} + \hat{\beta}(X_n - \bar{X}), \quad (5.8)$$

die Residuen (geschätzte Gleichungsfehler)

Residuen

$$\hat{\varepsilon}_n = Y_n - \hat{Y}_n \quad (5.9)$$

sowie der Varianz-Schätzer

$$\hat{\sigma}^2 = \frac{1}{N-2} \sum_n \hat{\varepsilon}_n^2 \quad (5.10)$$

gewonnen werden.

Der Mittelwert der geschätzten Y-Werte ist gleich dem Mittelwert von Y_n

$$\tilde{Y} = \hat{\alpha} + \hat{\beta}\bar{X} = \bar{Y} \quad (5.11)$$

sodass sich die Residuen exakt zu Null addieren

$$\sum_n \hat{\varepsilon}_n = \sum_n (Y_n - \hat{Y}_n) = N(\bar{Y} - \bar{Y}) = 0. \quad (5.12)$$

Daraus folgt unmittelbar, dass der Mittelwert der Residuen auch exakt Null ist. Es liegt also kein systematischer Fehler vor. Eine weitere Implikation ist, dass die geschätzte Regressionsgerade durch den Punkt (\bar{Y}, \bar{X}) verläuft. Es ist an dieser Stelle wichtig anzumerken, dass die genannte Eigenschaft durch die Berücksichtigung des Parameters α sichergestellt wird. Es empfiehlt sich also diesen stets im Regressionsmodell zu berücksichtigen und nicht $Y_n = \beta X_n + \varepsilon_n$ (ohne α) zu schätzen.

Beispiel 30 (Capital Asset Pricing Model als lineare Regression).

In Abb. 5.2 wurde eine Gerade an die Datenpunkte (x_n, y_n) angepasst. Bei den Daten handelt es sich jeweils um monatliche Überschussrenditen, d. h. prozentuale Änderungen in risikobehafteten Aktienkursen abzüglich der Rendite einer risikolosen Anlage. Bei der abhängigen Variable handelt es sich um die Portfolio-Überschussrenditen von US-amerikanischen Aktien im verarbeitenden Gewerbe, während die erklärende Variable die Überschussrenditen des gesamten Aktienmarktes sind. Der Zeitraum ist Januar 1960 bis September 2024 mit $N = 777$. Das Capital Asset Pricing Model (CAPM) unterstellt einen

linearen Zusammenhang zwischen diesen Variablen und impliziert einen Achsenabschnitt von 0. Demnach liegen keine systematisch höheren oder niedrigen Portfoliorenditen im Vergleich zum Gesamtmarkt vor. Der Steigungsparameter misst das systematische Risiko des Portfolios. Hierbei wird zwischen den Fällen $\beta < 1$ (defensiv), $\beta = 1$ (neutral) und $\beta > 1$ (aggressiv) unterschieden. Weitere Details finden sich im Zusatzmaterial in der Moodle-Umgebung.

Ausgehend vom dem Streudiagramm in Abb. 5.2 halten wir zunächst fest, dass der geschätzte Korrelationskoeffizient $r = 0.839$ (positiver Zusammenhang) ist. Für die Standardabweichungen haben wir $s_y = 5.723$ und $s_x = 4.467$. Damit ergibt sich

$$\hat{\beta} = r_{xy} \frac{s_y}{s_x} = 0.839 \times \frac{5.723}{4.467} = 1.075 .$$

Auf Basis der Mittelwerte $\bar{y} = 0.543$ und $\bar{x} = 0.577$ finden wir

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x} = 0.543 - 1.075 \times 0.577 = -0.077 .$$

Die geschätzten Werte für y_n ergeben sich aus

$$\hat{y}_n = \hat{\alpha} + \hat{\beta}x_n, \tag{5.13}$$

die Residuen (geschätzte Gleichungsfehler) durch

$$\hat{\varepsilon}_n = y_n - \hat{y}_n \tag{5.14}$$

sowie der Varianz-Schätzer

$$\hat{\sigma}^2 = \frac{1}{N-2} \sum_n \hat{\varepsilon}_n^2 = 7512.177/775 = 9.693. \tag{5.15}$$

Die geschätzte (empirische Regressionsgerade) ist also

$$\hat{y}_n = -0.077 + 1.075 \times x_n \tag{5.16}$$

mit Achsenabschnitt $\hat{\alpha} = -0.077 = \hat{y}_n(x_n = 0)$ und Steigungsparameter $\hat{\beta} = 1.075$.

■

Es ist wichtig, sich klarzumachen, dass die geschätzte Regressionsgerade durch die Mittelwerte $\bar{y} = 0.543$, $\bar{x} = 0.577$ verläuft, denn

$$\hat{y} = \hat{\alpha} + \hat{\beta}x \tag{5.17}$$

$$= \bar{y} - \hat{\beta}\bar{x} + \hat{\beta}x \tag{5.18}$$

$$= \bar{y} + \hat{\beta}(x - \bar{x}) \tag{5.19}$$

$$\hat{y}(x = \bar{x}) = \bar{y}. \tag{5.20}$$