Tutorium Allgemeines Lineares Modell

BSc Psychologie SoSe 2022

2. Termin: Vektoren und Regression

Belinda Fleischmann

Organisatorisches

Termine:

- jeden Donnerstag
  - außer 26.05. (Christi Himmelfahrt)
  - Gruppe 1: 11.15 - 12.45 Uhr [G22A-112]
  - Gruppe 2: 13.15 - 14.45 Uhr [G22A-209]

Follow-up letzte Woche

Selbstkontrollfragen - Vektoren

Selbstkontrollfragen - Regression (1 bis 7)

Follow-up letzte Woche

Selbstkontrollfragen - Vektoren

Selbstkontrollfragen - Regression (1 bis 7)

Definition von Vektoren in R

Die R-Funktion `c(data)` erzeugt einen Zeilenvektor

```
c(1:12)
```

```
>  [1]  1  2  3  4  5  6  7  8  9 10 11 12
```

Um einen Spaltenvektor zu erzeugen, verwenden wir die R-Funktion `matrix(data, nrow, ncol, byrow)`

```
matrix(c(1:12), nrow = 12)     # Vektor mit 12 Zeilen
```

```
>       [,1]
>  [1,]    1
>  [2,]    2
>  [3,]    3
>  [4,]    4
>  [5,]    5
>  [6,]    6
>  [7,]    7
>  [8,]    8
>  [9,]    9
> [10,]   10
> [11,]   11
> [12,]   12
```

Definition von Vektoren in R

`matrix(data, nrow, ncol, byrow)` befüllt eine Matrix mit Vektorelementen. Mit den Argumenten `nrow` und `ncol` bestimmen wir den Matrixtyp und mit `byrow` können wir bestimmen, *wie* die Matrix befüllt wird.

```
matrix(c(1:12), nrow = 3) # Vektor mit 3 Zeilen
```

```
>      [,1] [,2] [,3] [,4]
> [1,]    1    4    7   10
> [2,]    2    5    8   11
> [3,]    3    6    9   12
```

```
matrix(c(1:12), ncol = 4) # Vektor mit 4 Spalten
```

```
>      [,1] [,2] [,3] [,4]
> [1,]    1    4    7   10
> [2,]    2    5    8   11
> [3,]    3    6    9   12
```

Per default ist `byrow=FASLE`. D.h. die Matrix wird *nicht* "byrow", also Reihen-weise" befüllt, sondern Spalten-weise. Wenn wir `byrow=TRUE` setzen, wird die Matrix "byrow", also Reihenweise befüllt.

```
matrix(c(1:12), nrow = 3, byrow = TRUE) # Vektor mit 3 Zeilen, reihenweise befüllt
```

```
>      [,1] [,2] [,3] [,4]
> [1,]    1    2    3    4
> [2,]    5    6    7    8
> [3,]    9   10   11   12
```

Follow-up letzte Woche

Selbstkontrollfragen - Vektoren

Selbstkontrollfragen - Regression (1 bis 7)

Selbstkontrollfragen - Vektoren

1. Geben Sie die Definition eines Vektorraums wieder.
2. Geben Sie die Definition des reellen Vektorraums wieder.
3. Es seien
$$x := \begin{pmatrix} 2 \\ 1 \end{pmatrix}, y := \begin{pmatrix} 0 \\ 1 \end{pmatrix} \text{ und } a := 2. \tag{1}$$
Berechnen Sie
$$v = a(x+y) \text{ und } w = \frac{1}{a}(y-x) \tag{2}$$
und überprüfen Sie ihre Rechnung mit R.
4. Geben Sie die Definition des Skalarproduktes auf $\mathbb{R}^m$ wieder.
5. Für
$$x := \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}, y := \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, z := \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix} \tag{3}$$
berechnen Sie
$$\langle x, y \rangle, \langle x, z \rangle, \langle y, z \rangle \tag{4}$$
und überprüfen Sie ihre Rechnung mithilfe von R.
6. Geben Sie die Definition des Euklidischen Vektorraums wieder.
7. Definieren Sie die Länge eines Vektors im Euklidischen Vektorraum.
8. Berechnen Sie die Längen der Vektoren x, y, z aus Aufgabe 5 und überprüfen Sie ihre Rechnung mit R.

Selbstkontrollfragen - Vektoren

9. Geben Sie Definition des Abstands zweier Vektoren im Euklidischen Vektorraum wieder.
10. Berechnen Sie $d(x, y)$, $d(x, z)$ und $d(y, z)$ für x, y, z aus Aufgabe 5.
11. Geben Sie die Definition des Winkels zwischen zwei Vektoren im Euklidischen Vektorraum wieder.
12. Berechnen Sie die Winkel zwischen den Vektoren x und y, x und z, sowie y und z aus Aufgabe 5 mit R.
13. Definieren Sie die Begriffe der Orthogonalität und Orthonormalität von Vektoren.
14. Definieren Sie den Begriff der Linearkombination von Vektoren.
15. Definieren Sie den Begriff der linearen Unabhängigkeit von Vektoren.
16. Woran kann man erkennen, dass zwei Vektoren linear abhängig sind?
17. Dokumentieren Sie alle in dieser Einheit eingeführten R Befehle in einem Skript.

1. Geben Sie die Definition eines Vektorraums wieder.

Definition (Vektorraum)

Es seien V eine nichtleere Menge und S eine Menge von Skalaren. Weiterhin sei eine Abbildung

$$+ : V \times V \to V, (v_1, v_2) \mapsto +(v_1, v_2) =: v_1 + v_2, \tag{5}$$

genannt *Vektoraddition*, definiert. Schließlich sei eine Abbildung

$$\cdot : S \times V \to V, (s, v) \mapsto \cdot(s, v) =: sv, \tag{6}$$

genannt *Skalarmultiplikation* definiert. Dann wird das Tupel $(V, S, +, \cdot)$ genau dann *Vektorraum* genannt, wenn für beliebige Elemente $v, w, u \in V$ und $a, b \in S$ folgende Bedingungen gelten:

(1) Kommutativität der Vektoraddition $\quad v + w = w + v$

(2) Assoziativität der Vektoraddition $\quad (v + w) + u = v + (w + u)$

(3) Existenz eines neutralen Elements der Vektoraddition $\quad \exists\, 0 \in V$ mit $v + 0 = 0 + v = v.$

(4) Existenz inverser Elemente der Vektoraddition $\quad \forall v \in V\, \exists -v \in V$ mit $v + (-v) = 0.$

(5) Existenz eines neutralen Elements der Skalarmultiplikation $\quad \exists\, 1 \in S$ mit $1 \cdot v = v.$

(6) Assoziativität der Skalarmultiplikation $\quad a \cdot (b \cdot c) = (a \cdot b) \cdot c.$

(7) Distributivität hinsichtlich der Vektoraddition $\quad a \cdot (v + w) = a \cdot v + a \cdot w.$

(8) Distributivität hinsichtlich der Skalaraddition $\quad (a + b) \cdot v = a \cdot v + b \cdot v.$

2. Geben Sie die Definition des reellen Vektorraums wieder.

Definition (Reeller Vektorraum)

Für alle $x, y \in \mathbb{R}^m$ und $a \in \mathbb{R}$ definieren wir die *Vektoraddition* durch

$$+ : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^m, (x, y) \mapsto x + y = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} := \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_m + y_m \end{pmatrix} \tag{7}$$

und die *Skalarmultiplikation* durch

$$\cdot : \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}^m, (a, x) \mapsto ax = a \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} := \begin{pmatrix} ax_1 \\ \vdots \\ ax_m \end{pmatrix}. \tag{8}$$

Dann bildet $(\mathbb{R}^m, +, \cdot)$ mit den Rechenregeln der Addition und Multiplikation in $\mathbb{R}$ einen Vektorraum, den wir den *reellen Vektorraum* nennen.

3. Es seien $x := \begin{pmatrix} 2 \\ 1 \end{pmatrix}, y := \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ und $a := 2$.

Berechnen Sie $v = a(x+y)$ und $w = \frac{1}{a}(y-x)$ und überprüfen Sie ihre Rechnung mit R.

$$v = a(x+y) = 2\left(\begin{pmatrix} 2 \\ 1 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} 2(2+0) \\ 2(1+1) \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \end{pmatrix}$$

$$w = \frac{1}{a}(y-x) = \frac{1}{2}\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix} - \begin{pmatrix} 2 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} 0.5(0-2) \\ 0.5(1-1) \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

```r
x = matrix(c(2,1), nrow = 2)   # Vektordefinition
y = matrix(c(0,1), nrow = 2)   # Vektordefinition
a = 2                          # Skalardefinition
v = a*(x + y)                  # Vektoraddition und Skalarmultiplikation
w = 1/a * (y - x)
print(v)
```

```
>      [,1]
> [1,]    4
> [2,]    4
```

```r
print(w)
```

```
>      [,1]
> [1,]   -1
> [2,]    0
```

4. Geben Sie die Definition des Skalarproduktes auf $\mathbb{R}^m$ wieder

Definition (Skalarprodukt auf $\mathbb{R}^m$)

Das *Skalarprodukt auf* $\mathbb{R}^m$ ist definiert als die Abbildung

$$\langle\rangle : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}, (x, y) \mapsto \langle(x, y)\rangle := \langle x, y\rangle := \sum_{i=1}^{m} x_i y_i. \tag{9}$$

5. Für $x := \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}, y := \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, z := \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}$, berechnen Sie $\langle x, y\rangle, \langle x, z\rangle, \langle y, z\rangle$ und überprüfen Sie ihre Rechnung mithilfe von R.

$$\langle x, y\rangle = x_1y_1 + x_2y_2 + x_3y_3 = 2 \cdot 1 + 1 \cdot 0 + 3 \cdot 1 = 2 + 0 + 3 = 5$$

$$\langle x, z\rangle = x_1z_1 + x_2z_2 + x_3z_3 = 2 \cdot 3 + 1 \cdot 1 + 3 \cdot 0 = 6 + 1 + 0 = 7$$

$$\langle y, z\rangle = y_1z_1 + y_2z_2 + y_3z_3 = 1 \cdot 3 + 0 \cdot 1 + 1 \cdot 0 = 3 + 0 + 0 = 3$$

```r
# Vektordefinition
x = matrix(c(2,1,3), nrow = 3)
y = matrix(c(1,0,1), nrow = 3)
z = matrix(c(3,1,0), nrow = 3)

# Skalarprodukt mit R's komponentenweiser Multiplikation * und sum()
skalarprodukt_xy = sum(x*y)
skalarprodukt_xz = sum(x*z)
skalarprodukt_yz = sum(y*z)
print(c(skalarprodukt_xy, skalarprodukt_xz, skalarprodukt_yz))
```

```
> [1] 5 7 3
```

5. Für $x := \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}$, $y := \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, $z := \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}$, berechnen Sie $\langle x, y \rangle, \langle x, z \rangle, \langle y, z \rangle$ und überprüfen Sie ihre Rechnung mithilfe von R.

$$\langle x, y \rangle = x_1 y_1 + x_2 y_2 + x_3 y_3 = 2 \cdot 1 + 1 \cdot 0 + 3 \cdot 1 = 2 + 0 + 3 = 5$$

$$\langle x, z \rangle = x_1 z_1 + x_2 z_2 + x_3 z_3 = 2 \cdot 3 + 1 \cdot 1 + 3 \cdot 0 = 6 + 1 + 0 = 7$$

$$\langle y, z \rangle = y_1 z_1 + y_2 z_2 + y_3 z_3 = 1 \cdot 3 + 0 \cdot 1 + 1 \cdot 0 = 3 + 0 + 0 = 3$$

```r
# Vektordefinition
x = matrix(c(2,1,3), nrow = 3)
y = matrix(c(1,0,1), nrow = 3)
z = matrix(c(3,1,0), nrow = 3)

# Skalarprodukt mit R's Matrixtransposition t() und -multiplikation (%*%)
skalarprodukt_xy = t(x) %*% y
skalarprodukt_xz = t(x) %*% z
skalarprodukt_yz = t(y) %*% z
print(c(skalarprodukt_xy, skalarprodukt_xz, skalarprodukt_yz))
```

```
> [1] 5 7 3
```

6. Geben Sie die Definition des Euklidischen Vektorraums wieder.

Definition (Euklidischer Vektorraum)

Das Tupel $((\mathbb{R}^m, +, \cdot), \langle\rangle)$ aus dem reellen Vektorraum $(\mathbb{R}^m, +, \cdot)$ und dem Skalarprodukt $\langle\rangle$ auf $\mathbb{R}^m$ heißt *reeller kanonischer Euklidischer Vektorraum*.

7. Definieren Sie die Länge eines Vektors im Euklidischen Vektorraum.

Definition (Länge)

$((\mathbb{R}^m, +, \cdot), \langle\rangle)$ sei der Euklidische Vektorraum.

- Die *Länge* eines Vektors $x \in \mathbb{R}^m$ ist definiert als

$$\|x\| := \sqrt{\langle x, x\rangle}. \tag{10}$$

8. Berechnen Sie die Längen der Vektoren x, y, z aus Aufgabe 5 und überprüfen Sie ihre Rechnung mit R.

$$\|x\| = \left\| \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} \right\| = \sqrt{\left\langle \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} \right\rangle} = \sqrt{2^2 + 1^2 + 3^2} = \sqrt{4 + 1 + 9} = \sqrt{14} \approx 3.74$$

$$\|y\| = \left\| \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\| = \sqrt{\left\langle \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\rangle} = \sqrt{1^2 + 0^2 + 1^2} = \sqrt{1 + 0 + 1} = \sqrt{2} \approx 1.41$$

$$\|z\| = \left\| \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix} \right\| = \sqrt{\left\langle \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix} \right\rangle} = \sqrt{3^2 + 1^2 + 0^2} = \sqrt{9 + 1 + 0} = \sqrt{10} \approx 3.16$$

```r
norm_x = norm(matrix(c(2,1,3), nrow = 3), type = "2")  # Vektorlänge = l_2 Norm (type = "2")
norm_y = norm(matrix(c(1,0,1), nrow = 3), type = "2")
norm_z = norm(matrix(c(3,1,0), nrow = 3), type = "2")
print(c(norm_x, norm_y, norm_z))
```

```
> [1] 3.74 1.41 3.16
```

9. Geben Sie Definition des Abstands zweier Vektoren im Euklidischen Vektorraum wieder.

Definition (Abstand)

$((\mathbb{R}^m, +, \cdot), \langle\rangle)$ sei der Euklidische Vektorraum.

- Der *Abstand* zweier Vektoren $x, y \in \mathbb{R}^m$ ist definiert als

$$d(x, y) := \|x - y\|. \tag{11}$$

10. Berechnen Sie $d(x, y), d(x, z)$ und $d(y, z)$ für x, y, z aus Aufgabe 5.

$$d(x,y) = d\left(\begin{pmatrix}2\\1\\3\end{pmatrix}, \begin{pmatrix}1\\0\\1\end{pmatrix}\right) = \left\|\begin{pmatrix}2\\1\\3\end{pmatrix} - \begin{pmatrix}1\\0\\1\end{pmatrix}\right\| = \left\|\begin{pmatrix}1\\1\\2\end{pmatrix}\right\| = \sqrt{1^2+1^2+2^2} = \sqrt{6} \approx 2.45$$

$$d(x,z) = d\left(\begin{pmatrix}2\\1\\3\end{pmatrix}, \begin{pmatrix}3\\1\\0\end{pmatrix}\right) = \left\|\begin{pmatrix}2\\1\\3\end{pmatrix} - \begin{pmatrix}3\\1\\0\end{pmatrix}\right\| = \left\|\begin{pmatrix}-1\\0\\3\end{pmatrix}\right\| = \sqrt{(-1)^2+0^2+3^2} = \sqrt{10} \approx 3.16$$

$$d(y,z) = d\left(\begin{pmatrix}1\\0\\1\end{pmatrix}, \begin{pmatrix}3\\1\\0\end{pmatrix}\right) = \left\|\begin{pmatrix}1\\0\\1\end{pmatrix} - \begin{pmatrix}3\\1\\0\end{pmatrix}\right\| = \left\|\begin{pmatrix}-2\\-1\\1\end{pmatrix}\right\| = \sqrt{(-2)^2+(-1)^2+1^2} = \sqrt{6} \approx 2.45$$

```r
Abstand_xy = norm(matrix(c(2,1,3), nrow = 3) - matrix(c(1,0,1), nrow = 3), type = "2")
Abstand_xz = norm(matrix(c(2,1,3), nrow = 3) - matrix(c(3,1,0), nrow = 3), type = "2")
Abstand_yz = norm(matrix(c(1,0,1), nrow = 3) - matrix(c(3,1,0), nrow = 3), type = "2")
print(c(Abstand_xy, Abstand_xz, Abstand_yz))
```

```
> [1] 2.45 3.16 2.45
```

11. Geben Sie die Definition des Winkels zwischen zwei Vektoren im Euklidischen Vektorraum wieder.

Definition (Winkel)

$((\mathbb{R}^m, +, \cdot), \langle\rangle)$ sei der Euklidische Vektorraum.

- Der *Winkel* α zwischen zwei Vektoren $x, y \in \mathbb{R}^m$ mit $x, y \neq 0$ ist definiert durch

$$0 \leq \alpha \leq \pi \text{ und } \cos\alpha := \frac{\langle x, y\rangle}{\|x\|\|y\|} \tag{12}$$

12. Berechnen Sie die Winkel zwischen den Vektoren x und y, x und z, sowie y und z aus Aufgabe 5 mit R.

Winkel zwischen x und y in Radians

$$\alpha = \text{acos}\left(\frac{\left\langle \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\rangle}{\left\| \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} \right\| \left\| \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\|}\right) = \text{acos}\left(\frac{2 \cdot 1 + 1 \cdot 0 + 3 \cdot 1}{\sqrt{2^2 + 1^2 + 3^2} \cdot \sqrt{1^2 + 0^2 + 1^2}}\right) = \text{acos}\left(\frac{5}{2 \cdot \sqrt{7}}\right) \approx 0.333$$

Winkel zwischen x und y in Grad

$$\text{acos}\left(\frac{5}{2 \cdot \sqrt{7}}\right) \cdot \frac{180}{\pi} \approx 19.107$$

Berechnung in R

```
x = matrix(c(2,1,3), nrow = 3)
y = matrix(c(1,0,1), nrow = 3)
alpha = acos(sum(x*y)/(sqrt(sum(x*x))*sqrt(sum(y*y))))          # Winkel in Radians
w = acos(sum(x*y)/(sqrt(sum(x*x))*sqrt(sum(y*y)))) * 180/pi     # Winkel in Grad
print(c(alpha,w))
```

```
> [1]  0.333 19.107
```

12. Berechnen Sie die Winkel zwischen den Vektoren x und y, x und z, sowie y und z aus Aufgabe 5 mit R.

Winkel zwischen x und z in Radians

$$\alpha = \text{acos}\left(\frac{\left\langle \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix} \right\rangle}{\left\| \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} \right\| \left\| \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix} \right\|}\right) = \text{acos}\left(\frac{2 \cdot 3 + 1 \cdot 1 + 3 \cdot 0}{\sqrt{2^2 + 1^2 + 3^2} \cdot \sqrt{3^2 + 1^2 + 0^2}}\right) = \text{acos}\left(\frac{7}{2 \cdot \sqrt{35}}\right) \approx 0.938$$

Winkel zwischen x und y in Grad

$$\text{acos}\left(\frac{7}{2 \cdot \sqrt{35}}\right) \cdot \frac{180}{\pi} \approx 53.729$$

Berechnung in R

```
x = matrix(c(2,1,3), nrow = 3)
z = matrix(c(3,1,0), nrow = 3)
alpha = acos(sum(x*z)/(sqrt(sum(x*x))*sqrt(sum(z*z))))        # Winkel in Radians
w = acos(sum(x*z)/(sqrt(sum(x*x))*sqrt(sum(z*z)))) * 180/pi   # Winkel in Grad
print(c(alpha,w))
```

```
> [1]  0.938 53.729
```

12. Berechnen Sie die Winkel zwischen den Vektoren x und y, x und z, sowie y und z aus Aufgabe 5 mit R.

Winkel zwischen y und z in Radians

$$\alpha = \operatorname{acos}\left(\frac{\left\langle \begin{pmatrix}1\\0\\1\end{pmatrix}, \begin{pmatrix}3\\1\\0\end{pmatrix} \right\rangle}{\left\| \begin{pmatrix}1\\0\\1\end{pmatrix} \right\| \left\| \begin{pmatrix}3\\1\\0\end{pmatrix} \right\|}\right) = \operatorname{acos}\left(\frac{1\cdot 3 + 0\cdot 1 + 1\cdot 0}{\sqrt{1^2+0^2+1^2}\cdot\sqrt{3^2+1^2+0^2}}\right) = \operatorname{acos}\left(\frac{3}{2\cdot\sqrt{5}}\right) \approx 0.835$$

Winkel zwischen x und y in Grad

$$\operatorname{acos}\left(\frac{3}{2\cdot\sqrt{5}}\right)\cdot\frac{180}{\pi} = 47.87$$

Berechnung in R

```
y = matrix(c(1,0,1), nrow = 3)
z = matrix(c(3,1,0), nrow = 3)
alpha = acos(sum(y*z)/(sqrt(sum(y*y))*sqrt(sum(z*z))))          # Winkel in Radians
w = acos(sum(y*z)/(sqrt(sum(y*y))*sqrt(sum(z*z)))) * 180/pi      # Winkel in Grad
print(c(alpha,w))
```

```
> [1]  0.835 47.870
```

13. Definieren Sie die Begriffe der Orthogonalität und Orthonormalität von Vektoren.

Definition (Orthogonalität und Orthonormalität von Vektoren)

$((\mathbb{R}^m, +, \cdot), \langle\rangle)$ sei der Euklidische Vektorraum.

- Zwei Vektoren $x, y \in \mathbb{R}^m$ heißen *orthogonal*, wenn gilt, dass

$$\langle x, y \rangle = 0 \tag{13}$$

- Zwei Vektoren $x, y \in \mathbb{R}^m$ heißen *orthonormal*, wenn gilt, dass

$$\langle x, y \rangle = 0 \text{ und } \|x\| = \|y\| = 1. \tag{14}$$

14. Definieren Sie den Begriff der Linearkombination von Vektoren.

Definition (Linearkombination)

$\{v_1, v_2, ..., v_k\}$ sei eine Menge von k Vektoren eines Vektorraums V. Dann ist die *Linearkombination* der Vektoren in $v_1, v_2, ..., v_k$ mit den skalaren Koeffizienten $a_1, a_2, ..., a_k$ definiert als der Vektor

$$w := \sum_{i=1}^{k} a_i v_i \in V. \tag{15}$$

15. Definieren Sie den Begriff der linearen Unabhängigkeit von Vektoren.

Definition (Lineare Unabhängigkeit)

V sei ein Vektorraum. Eine Menge $W := \{w_1, w_2, ..., w_k\}$ von Vektoren in V heißt *linear unabhängig*, wenn die einzige Repräsentation des Nullelements $0 \in V$ durch eine Linearkombination der $w \in W$ die triviale Repräsentation

$$0 = a_1 w_1 + a_2 w_2 + \cdots + a_k w_k \text{ mit } a_1 = a_2 = \cdots = a_k = 0 \tag{16}$$

ist. Wenn die Menge W nicht linear unabhängig ist, dann heißt sie *linear abhängig*.

16. Woran kann man erkennen, dass zwei Vektoren linear abhängig sind?

Theorem (Lineare Abhängigkeit von zwei Vektoren)

V sei ein Vektorraum. Zwei Vektoren $v_1, v_2 \in V$ sind linear abhängig, wenn einer der Vektoren ein skalares Vielfaches des anderen Vektors ist.

Follow-up letzte Woche

Selbstkontrollfragen - Vektoren

Selbstkontrollfragen - Regression (1 bis 7)

Selbstkontrollfragen - Regression (1 bis 7)

1. Geben Sie die funktionale Form eine linear-affinen Funktion an.
2. Erläutern Sie die Bedeutung der Parameter einer linear-affinen Funktion.
3. Definieren Sie den Begriff der Ausgleichsgerade.
4. Erläutern Sie die intuitive Bedeutung der Funktion der quadrierten vertikalen Abweichungen.
5. Geben Sie das Theorem zur Ausgleichsgerade wieder.
6. Skizzieren Sie den Beweis des Theorems zur Ausgleichsgeraden.
7. Definieren Sie den Begriff des Ausgleichspolynoms.

1. Geben Sie die funktionale Form eine linear-affinen Funktion an.

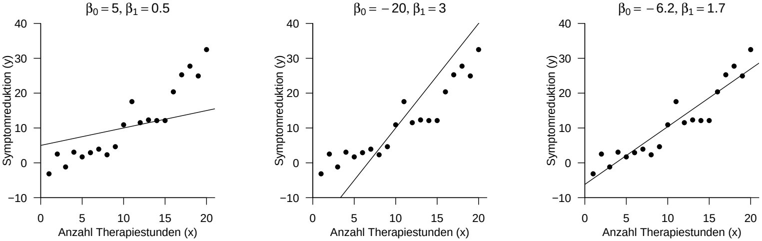
Für $\beta := (\beta_0, \beta_1)^T \in \mathbb{R}^2$ hat die linear-affine Funktion $f_\beta : \mathbb{R} \to \mathbb{R}, x \mapsto f_\beta(x)$ die funktionale Form $f_\beta(x) := \beta_0 + \beta_1 x$.

2. Erläutern Sie die Bedeutung der Parameter einer linear-affinen Funktion.

Für linear-affine Funktionen $f_\beta(x) := \beta_0 + \beta_1 x$ ist

- β_0 der Schnittpunkt von Gerade und y-Achse ("Offset Parameter") und
- β_1 die y-Differenz pro x-Einheitsdifferenz ("Steigungsparameter").

Beispiele:



3. Definieren Sie den Begriff der Ausgleichsgerade.

Definition (Ausgleichsgerade)

Für $\beta := (\beta_0, \beta_1)^T \in \mathbb{R}^2$ heißt die linear-affine Funktion

$$f_\beta : \mathbb{R} \to \mathbb{R}, x \mapsto f_\beta(x) := \beta_0 + \beta_1 x, \tag{17}$$

für die für eine Wertemenge $\{(x_1, y_1), ..., (x_n, y_n)\} \subset \mathbb{R}^2$ die Funktion

$$q : \mathbb{R}^2 \to \mathbb{R}_{\geq 0}, \beta \mapsto q(\beta) := \sum_{i=1}^{n} (y_i - f_\beta(x_i))^2 = \sum_{i=1}^{n} (y_i - (\beta_0 + \beta_1 x_i))^2 \tag{18}$$

der quadrierten vertikalen Abweichungen der y_i von den Funktionswerten $f_\beta(x_i)$ ihr Minimum annimt, die *Ausgleichsgerade* für die Wertemenge $\{(x_1, y_1), ..., (x_n, y_n)\}$.

4. Erläutern Sie die intuitive Bedeutung der Funktion der quadrierten vertikalen Abweichungen.

Die Funktion der quadrierten vertikalen Abweichungen $q(\beta)$ quantifiziert die Abweichungen der Funktionswerte $f_\beta(x_i)$ der Ausgleichsgeraden f_β von den (beobachteten) Werten y_i.

Die Abweichungen werden *quadriert*, damit sich positive und negative Abweichungen nicht zu null ausgleichen.

Intuitiv misst die Funktion $q(\beta)$ "wie gut" eine Ausgleichsgerade mit bestimmten Werten für β_0 und β_1 in das Streudiagramm der Datenpunkte "rein passt".

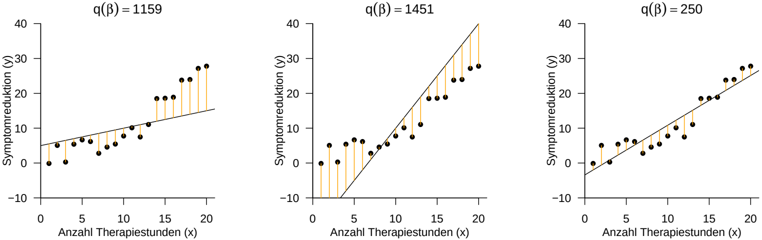
Je größer ein Funktionswert $q(\beta)$, desto größer ist die aufsummierte Abweichung und desto "schlechter" passt die Ausgleichsgerade f_β in das Streudiagramm der Datenpunkte.

Je geringer ein Funktionswert $q(\beta)$, desto geringer ist die aufsummierte Abweichung und desto "besser" passt die Ausgleichsgerade f_β in das Streudiagramm der Datenpunkte.

Zur Veranschaulichung

Funktion der quadrierten vertikalen Abweichungen

$$q(\beta) := \sum_{i=1}^{n} (y_i - (\beta_0 + \beta_1 x_i))^2$$



$y_i - (\beta_0 + \beta_1 x_i)$ für $i = 1, \ldots, n$

5. Geben Sie das Theorem zur Ausgleichsgerade wieder.

Theorem (Ausgleichsgerade)

Für eine Wertemenge $\{(x_1, y_1), ..., (x_n, y_n)\} \subset \mathbb{R}^2$ hat die Ausgleichsgerade die Form

$$f_\beta : \mathbb{R} \to \mathbb{R}, x \mapsto f_\beta(x) := \hat{\beta}_0 + \hat{\beta}_1 x, \tag{19}$$

wobei mit der Stichprobenkovarianz c_{xy} der (x_i, y_i)-Werte, der Stichprobenvarianz s_x^2 der x_i-Werte und den Stichprobenmitteln $\bar{x}$ und $\bar{y}$ der x_i- und y_i-Werte, respektive, gilt, dass

$$\hat{\beta}_1 = \frac{c_{xy}}{s_x^2} \text{ und } \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \tag{20}$$

6. Skizzieren Sie den Beweis des Theorems zur Ausgleichsgeraden.

Wir wollen zeigen, dass die Summe der quadrierten Abweichungen $q(\beta_0, \beta_1)$ für die Werte $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$ und $\hat{\beta}_1 = \frac{c_{xy}}{s_x^2}$ ihr Minimum annimmt, wobei $\bar{x}$ und $\bar{y}$ die Stichprobenmittel der x_i- und y_i-Werte, repektive, sind, c_{xy} die Stichprobenkovarianz der (x_i, y_i)-Werte und s_x^2 die Stichprobenvarianz der x_i-Werte entspricht.

Dafür bestimmen wir zunächst das Minimum der Funktion $q(\beta_0, \beta_1)$, indem wir die partiellen Ableitungen hinsichtlich β_0 und β_1 berechnen und diese gleich 0 setzen. Formal,

$$\frac{\partial}{\partial \beta_0} q(\beta_0, \beta_1) = 0 \text{ und } \frac{\partial}{\partial \beta_1} q(\beta_0, \beta_1) = 0$$

Durch die partielle Ableitungen und das anschließende Umstellen von Termen ergibt sich ein Gleichungssystem, das das *System der Normalengleichungen* genannt wird und die notwendige Bedingung für ein Minimum von q beschreibt. Auflösen dieses Gleichungssystems nach β_0 und β_1 liefert dann die Werte $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$ und $\hat{\beta}_1 = \frac{c_{xy}}{s_x^2}$ des Theorems.

7. Definieren Sie den Begriff des Ausgleichspolynoms.

Definition (Ausgleichspolynom)

Für $\beta := (\beta_0, ..., \beta_k)^T \in \mathbb{R}^{k+1}$ heißt die Polynomfunktion $k$ten Grades

$$f_\beta : \mathbb{R} \to \mathbb{R}, x \mapsto f_\beta(x) := \sum_{i=0}^{k} \beta_i x^i, \tag{21}$$

für die für eine Wertemenge $\{(x_1, y_1), ..., (x_n, y_n)\} \subset \mathbb{R}^2$ die Funktion

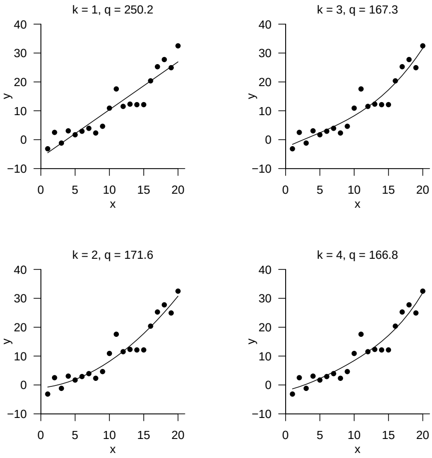
$$q : \mathbb{R}^2 \to \mathbb{R}_{\geq 0}, \beta \mapsto q(\beta) := \sum_{i=1}^{n} \left(y_i - f_\beta(x_i)\right)^2 = \sum_{i=1}^{n} \left(y_i - \sum_{i=0}^{k} \beta_i x^i\right)^2 \tag{22}$$

der quadrierten vertikalen Abweichungen der y_i von den Funktionswerten $f_\beta(x_i)$ ihr Minimum annimt, das *Ausgleichspolynom $k$ten Grades* für die Wertemenge $\{(x_1, y_1), ..., (x_n, y_n)\}$.

Methode der kleinsten Quadrate - Selbstkontrollfragen

Zur Veranschaulichung

Beispieldatensatz Ausgleichspolynome 1ten bis 4ten Grades



$$\bullet (x_i, y_i) \qquad — \; f_{\hat{\beta}}(x) = \sum_{i=0}^{k} \hat{\beta}_i x^i$$