Johann-Friedrich Luy
Prof. Dr.-Ing.

Grundlagen der Statistik

Einführung

Das gesamte Thema der Statistik fußt auf der Vorstellung eines großen Datensatzes, der analysiert werden soll in Bezug auf die Beziehungen der einzelnen Punkte des Datensatzes untereinander. Zunächst werden daher einige Maßzahlen betrachtet, die aus einem Datensatz abgeleitet werden können und es wird erörtert, was diese Maßzahlen für eine Aussage enthalten.

Die Standardabweichung

Für das Verständnis der Standardabweichung wird ein Datensatz benötigt. Statistik beschäftigt sich üblicherweise mit einer Stichprobe aus einer Population. Beispielsweise ist die Gesamtzahl der Menschen eines Landes bei Wahlumfragen die Population, während eine Stichprobe eine Untermenge der Population darstellt, die durch Statistik gemessen wird. Die Bedeutung von Statistik liegt darin, dass nur durch die Messung einer Stichprobe (in dem genannten Fall beispielsweise durch eine Telefonumfrage) ermittelt werden kann, was wahrscheinlich das Messergebnis wäre, wenn die gesamte Population vermessen würde. In diesem Abschnitt wird angenommen, dass die vorliegenden Datensätze Stichproben aus einer größeren Population sind.

Gegeben sei folgender Beispieldatensatz:

$$X = [1\ 2\ 4\ 6\ 12\ 15\ 25\ 45\ 68\ 67\ 65\ 98]$$

Das Symbol X wird einfach genutzt um auf den gesamten Datensatz zu verweisen. Wenn auf ein bestimmtes Element dieses Datensatzes verwiesen werden soll, werden Indizes mit dem Symbol X verwendet. Beispielsweise bezieht sich X_3 auf das dritte Element in X, also in diesem Fall die Zahl 4. Damit ist auch klar, dass mit X_1 das erste Element in X adressiert ist und nicht mit X_0, wie das teilweise in der Literatur vorkommt. Ferner bezieht sich das Symbol n auf die Anzahl der Elemente in dem Datensatz X.

Aus dem Datensatz können einige Größen berechnet werden. Beispielsweise der arithmetische Mittelwert, der hier als bekannt vorausgesetzt wird und der sich formelmäßig darstellt durch

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

Das Symbol $\bar{X}$ (gesprochen „X quer“) bezeichnet den Mittelwert des Datensatzes X. Diese Formel sagt lediglich aus: “addiere alle Zahlen auf und Teile sie dann durch ihre Anzahl“.

Leider ist dieser Mittelwert bezüglich des Datensatzes nicht besonders aussagekräftig. Beispielsweise können zwei Datensätze exakt den gleichen Mittelwert haben (10) obwohl sie offensichtlich sehr verschieden sind:

$$[0\ 8\ 12\ 20]\ und\ [8\ 9\ 11\ 12]$$

Was ist dann der Unterschied zwischen den beiden Datensätzen? Unterschiedlich ist die *Streuung* der Daten. Die Standardabweichung eines Datensatzes ist ein Maß für die Streuung der Daten. Wie wird die Standardabweichung berechnet? Die Definition besagt: „Die durchschnittliche Entfernung eines Punktes vom Mittelpunkt des Datensatzes". Die Rechenvorschrift besagt, dass die Quadrate der Entfernung jedes Datenpunktes vom Mittelpunkt zu berechnen sind und alle aufaddiert werden, durch $n-1$ dividiert werden und anschließend die Quadratwurzel gezogen wird. Die Formel lautet:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{(n-1)}}$$

Hier ist σ das verwendete Symbol für die Standardabweichung. Warum wird hier $(n-1)$ und nicht n verwendet? Die Erklärung beruht auf der Erwartungstreue der Schätzer. Wenn es sich bei dem Datensatz um eine Stichprobe handelt, also um eine Untermenge der realen Welt (im Wahlumfragebeispiel 500 Personen), dann muss $(n-1)$ verwendet werden, da dann das Ergebnis für die Standardabweichung näher an dem Ergebnis ist, das erzielt würde, wenn die *gesamte* Population verwendet würde, als wenn nur durch n dividiert würde.

Für die beiden oben angegebenen Datensätze sind die Berechnungen der Standardabweichung in der Tabelle wiedergegeben:

$Datensatz\ 1$

$\boldsymbol{X}$	$(\boldsymbol{X}-\bar{\boldsymbol{X}})$	$(\boldsymbol{X}-\bar{\boldsymbol{X}})^{\boldsymbol{2}}$
0	-10	100
8	-2	4
12	2	4
20	10	100
Summe		208
Division durch $(n-1)$		69,333
Quadratwurzel		8,3266

$Datensatz\ 2$

$\boldsymbol{X}$	$(\boldsymbol{X} - \bar{\boldsymbol{X}})$	$(\boldsymbol{X} - \bar{\boldsymbol{X}})^{\boldsymbol{2}}$
8	-2	4
9	-1	1
11	1	1
12	2	4
Summe		10
Division durch $(n-1)$		3,333
Quadratwurzel		1,8257

Beispielhafte Berechnung der Standardabweichung für zwei Datensätze

Erwartungsgemäß hat der erste Datensatz eine deutlich größere Standardabweichung als der zweite, da die Daten deutlich stärker um den Mittelwert streuen. Ein anderes Beispiel ist der folgende Datensatz:

$$[10\ 10\ 10\ 10]$$

Dieser Datensatz hat auch den Mittelwert 10, aber seine Standardabweichung ist 0, da alle Elemente dieselben sind – kein Element weicht vom Mittelwert ab.

Die Varianz

Ein anderes Maß für die Streuung der Daten in einem Datensatz ist die Varianz. Die Varianz ist das Quadrat der Standardabweichung:

$$\sigma^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{(n-1)}$$

σ^2 wird hier als Symbol für die Varianz verwendet. Beide Messgrößen (Varianz und Standardabweichung) sind *Streumaße.* Die Varianz ist eine wichtige Voraussetzung für die Einführung der Kovarianz.

Die Kovarianz

Die beiden behandelten Streumaße sind ausschließlich 1-dimensional. Beispiele für solche Datensätze könnten beispielsweise die Körpergrößen aller Menschen in einem Raum oder die Notenverteilung bei einer Prüfung sein. Viele Datensätze sind allerdings mehrdimensional, und das Ziel von statistischen Analysen kann dann sein, zu untersuchen, ob es Beziehungen zwischen den Dimensionen gibt. Beispielsweise könnten wir einen Datensatz vorliegen haben, der die Körpergrößen einer Gruppe von Menschen enthält

und ihre Noten bei einer Prüfung. Dann könnte das Ergebnis einer statistischen Analyse sein, zu erkennen, ob es eine Beziehung zwischen der Körpergröße und der erzielten Note gibt.

Die Standardabweichung und die Varianz werden nur auf eine Dimension angewandt, so dass die Standardabweichung nur für jede Dimension des Datensatzes unabhängig von den anderen Dimensionen berechnet werden kann. Insofern erscheint es nützlich ein ähnliches Maß zu finden, mit dem ausgedrückt werden kann, wie stark die Dimensionen vom Mittelwert abweichen *in Bezug zu einander.*

Die Kovarianz ist dieses Streumaß. Die Kovarianz wird immer zwischen 2 Dimensionen gemessen. Wenn die Kovarianz zwischen einer Dimension und *sich selbst* gemessen wird, ergibt sich die Varianz. In einem 3-dimensionalen Datensatz (x, y, z) kann die Kovarianz zwischen den x und y Dimensionen, den x und z Dimensionen, und den y und z Dimensionen gemessen werden. Wird die Kovarianz zwischen x und x, oder y und y, oder z und z gemessen, so ergibt sich die Varianz der x, y beziehungsweise z Dimensionen.

Die Formel für die Varianz kann auch mit ausgeschriebenem quadratischem Term dargestellt werden:

$$var(X) = \sigma^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(X_i - \bar{X})}{(n-1)}$$

Für die Kovarianz wird geschrieben:

$$cov(X,Y) = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{(n-1)}$$

Die Kovarianz unterscheidet sich also von der Varianz dadurch, dass im zweiten Klammerausdruck im Zähler die $X_i´s$ durch $Y_i´s$ und $\bar{X}$ durch $\bar{Y}$ ersetzt werden. In Worten: „Multipliziere für jeden Datenpunkt die Differenz zwischen dem x-Wert und dem Mittelwert der x-Werte mit der Differenz zwischen dem y-Wert und dem Mittelwert der y-Werte. Addiere alle diese Produkte auf und dividiere die Summe durch $(n-1)$."

Was sagt die Kovarianz aus? Dazu soll ein zweidimensionaler Datensatz betrachtet werden. Der Beispieldatensatz enthält die Vorbereitungszeit von 12 Kandidaten für eine Prüfung in Stunden (h) und die in der Prüfung jeweils erreichte Punktezahl. Die zwei Dimensionen bestehen also aus der H_i- Dimension, der Vorbereitungszeit und der M_i-Dimension, der erreichten Punktezahl. Die folgende Tabelle zeigt die fiktiven Daten.

	H_i=Vorbereitungszeit (h)	M_i=Erreichte Punkte
	9	39
	15	56
	25	93
	14	61
	10	50
	18	75
	0	32
	16	85
	5	42
	19	70
	16	66
	20	80
Summe	167	749
Durchschnitte	13,92	62,42

H_i	M_i	$(H_i - \bar{H})$	$(M_i - \bar{M})$	$(H_i - \bar{H})(M_i - \bar{M})$
9	39	-4,92	-23,42	115,23
15	56	1,08	-6,42	-6,93
25	93	11,08	30,58	338,83
14	61	0,08	-1,42	-0,11
10	50	-3,92	-12,42	48,69
18	75	4,08	12,58	51,33
0	32	-13,92	-30,42	423,45
16	85	2,08	22,58	46,97
5	42	-8,92	-20,42	182,15
19	70	5,08	7,58	38,51
16	66	2,08	3,58	7,45
20	80	6,08	17,58	106,89
Summe				1149,89
Durchschnitt				104,54

Ein zweidimensionaler Datensatz und Kovarianz-Berechnung

Der exakte Wert der Kovarianz ist nicht so bedeutend wie das Vorzeichen. Wenn der Wert positiv ist, so bedeutet das, dass beide Dimensionen gleichsinnig verlaufen, wenn also die eine Dimension zunimmt, die andere auch anwächst.

Ist der Wert negativ, dann nimmt die eine Dimension zu während die andere abnimmt.

Wenn die Kovarianz Null ist, dann sind die beiden Dimensionen unabhängig voneinander.

Im zweidimensionalen Fall kann ein solcher Zusammenhang einfach in einem Streudiagramm visualisiert werden. Auch in drei Dimensionen ist die Visualisierung noch möglich. Da die Kovarianz-Werte zwischen allen beliebigen Dimensionen in einem Datensatz berechnet werden können, wird diese Technik oft verwendet um Korrelationen zwischen den Dimensionen in hochdimensionalen Datensätzen zu finden, die einer Visualisierung sonst nicht mehr zugeführt werden können.

Aus der Definition für die Kovarianz ist auch ersichtlich, dass $cov(X, Y)$ gleich ist zu $cov(Y, X)$. Der einzige Unterschied zwischen $cov(X, Y)$ und $cov(Y, X)$ besteht darin, dass $(X_i - \bar{X})(Y_i - \bar{Y})$ ersetzt wird durch

$(Y_i - \bar{Y})(X_i - \bar{X})$ und da die Multiplikation kommutativ ist, spielt es keine Rolle auf welche Weise die beiden Klammerausdrücke multipliziert werden.

Die Kovarianz-Matrix

Die Kovarianz wird also immer zwischen zwei Dimensionen gemessen. Wenn ein Datensatz mit mehr als zwei Dimensionen vorliegt, dann können mehrere Kovarianzen berechnet werden. Für einen dreidimensionalen Datensatz mit den Dimensionen x, y, z kann daher $cov(x, y); cov(x, z)$ und $cov(y, z)$ berechnet werden. Für einen n −dimensionalen Datensatz können dann $\frac{n!}{(n-2)!*2}$ verschiedene Kovarianz-Werte berechnet werden.

Werden alle möglichen Kovarianzen zwischen verschiedenen Dimensionen berechnet, so können diese in einer Matrix angeordnet werden. Die Definition für die Kovarianz-Matrix eines n-dimensionalen Datensatzes lautet:

$$C^{n\times n} = (c_{i,j}; c_{i,j} = cov(Dim_i, Dim_j))$$

Hier ist $C^{n\times n}$ eine Matrix mit n Zeilen und n Spalten und Dim_x ist die x-te Dimension. Die Aussage dieser Formel ist, dass im Falle eines n-dimensionalen Datensatzes die Matrix n Zeilen und n Spalten hat, also quadratisch ist, und dass jeder Eintrag in der Matrix das Ergebnis der Berechnung der Kovarianz zwischen zwei verschiedenen Dimensionen ist. Zum Beispiel enthält der Eintrag in Zeile 2, Spalte 3 den berechneten Kovarianz-Wert zwischen der zweiten und der dritten Dimension.

Für den Fall eines beliebigen dreidimensionalen Datensatzes mit den Dimensionen x, y, z hat die Kovarianz-Matrix 3 Zeilen und 3 Spalten und die Einträge lauten:

$$C = \begin{pmatrix} cov(x,x) & cov(x,y) & cov(x,z) \\ cov(y,x) & cov(y,y) & cov(y,z) \\ cov(z,x) & cov(z,y) & cov(z,z) \end{pmatrix}$$

Es ist offensichtlich, dass in der Hauptdiagonale die Kovarianzen jeder Dimension mit sich selber stehen. Dies sind die Varianzen für die jeweilige Dimension. Da gilt $cov(a, b) = cov(b, a)$ ist die Matrix symmetrisch zur Hauptdiagonale.

Matrizen-Rechnung

In diesem Abschnitt werden einige Besonderheiten der Matrizenrechnung behandelt, die für das Datenhandling von Bedeutung sind. Insbesondere geht es um Eigenvektoren und Eigenwerte einer gegebenen Matrix. Grundkenntnisse der Matrizenrechnung sind erforderlich.

Eigenvektoren

Zwei Matrizen können miteinander multipliziert werden, wenn ihre Größen kompatibel sind. Ein Spezialfall davon sind die Eigenvektoren. Im folgenden Beispiel ist die Multiplikation zwischen einer Matrix und einem Vektor dargestellt.

$$\begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix} * \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 11 \\ 5 \end{pmatrix}$$

$$\begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix} * \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 12 \\ 8 \end{pmatrix} = 4 * \begin{pmatrix} 3 \\ 2 \end{pmatrix}$$

Beispiel ohne Eigenvektor und mit einem Eigenvektor

Im ersten Fall ist der Ergebnis-Vektor kein ganzzahliges Vielfaches des Originalvektors. Im zweiten Fall ist der Ergebnisvektor genau das 4-fache des Ausgangs-Vektors. Der Vektor ist ein Vektor im 2-dimensionalen Raum.

Der Vektor $\begin{pmatrix} 3 \\ 2 \end{pmatrix}$ aus der zweiten Beispiel-Multiplikation stellt einen Pfeil dar, der vom Ursprung $(0{,}0)$, zum Punkt $(3{,}2)$ zeigt. Die quadratische Matrix kann wie eine Transformations-Matrix betrachtet werden. Wird diese links von einem Vektor stehende Matrix multipliziert, so entsteht ein, von seiner ursprünglichen Position aus, transformierter Vektor.

Aus dieser Transformation entstehen Eigenvektoren. Eine linksstehende Transformations-Matrix liefere nach der Multiplikation Vektoren auf der Linie $y = x$. Dann ist ein Vektor, der *auf* der Linie $y = x$ liegt eine Darstellung von sich selbst. Dieser Vektor (sowie alle Vielfachen von ihm, da die Länge keine Rolle spielt) ist ein Eigenvektor der Transformations-Matrix.

Zu den wichtigen Eigenschaften der Eigenvektoren gehört, dass Eigenvektoren nur für *quadratische* Matrizen entstehen. Aber nicht jede quadratische Matrix hat Eigenvektoren. Hat eine $n \times n$ Matrix überhaupt Eigenvektoren, so hat sie n davon. Eine 3×3 Matrix kann 3 Eigenvektoren haben.

Eine weitere Eigenschaft von Eigenvektoren ist nachstehend dargestellt: wird der Vektor vor der Matrizen-Multiplikation skaliert, so erhält man dasselbe Vielfache als Ergebnis.

$$2 * \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 6 \\ 4 \end{pmatrix}$$

$$\begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix} * \begin{pmatrix} 6 \\ 4 \end{pmatrix} = \begin{pmatrix} 24 \\ 16 \end{pmatrix} = 4 * \begin{pmatrix} 6 \\ 4 \end{pmatrix}$$

Beispiel für einen skalierten Eigenvektor

Die Skalierung eines Vektors um einen Betrag verändert nur seine Länge, nicht seine Richtung.

Und schließlich stehen alle Eigenvektoren einer Matrix *senkrecht* aufeinander, d.h., sie bilden rechte Winkel untereinander unabhängig von der Anzahl der Dimensionen. Diese Eigenschaft wird auch *Orthogonalität* genannt. Diese wichtige Eigenschaft bedeutet, dass die Daten auch in Form der senkrechten Eigenvektoren

ausgedrückt werden können anstatt in Einheiten der x und y Achsen. Das spielt eine wichtige Rolle in der Hauptkomponentenanalyse.

In vielen Fällen sind die normierten Eigenvektoren gesucht, deren Länge genau eins beträgt. Wie bereits ausgeführt, beeinträchtigt die Länge nicht die Frage ob es sich um einen Eigenvektor handelt oder nicht, anders als die Richtung des Vektors. Daher werden die Eigenvektoren standardisiert durch eine Normierung auf die Länge eins, so dass alle Eigenvektoren die gleiche Länge haben. Wenn

$$\begin{pmatrix} 3 \\ 2 \end{pmatrix}$$

einen Eigenvektor darstellt, und die Länge des Eigenvektors dann

$$\sqrt{3^2 + 2^2} = \sqrt{13}$$

beträgt, dann wird der Original-Vektor durch diesen Wert dividiert damit er die Länge eins erhält:

$$\begin{pmatrix} 3 \\ 2 \end{pmatrix} : \sqrt{13} = \begin{pmatrix} 3/\sqrt{13} \\ 2/\sqrt{13} \end{pmatrix}$$

Zur Berechnung von Eigenvektoren ist es hilfreich zuvor die Eigenwerte zu ermitteln.

Eigenwerte

Die Eigenwerte stehen in enger Beziehung zu den Eigenvektoren und waren schon im *Beispiel ohne Eigenvektor und mit einem Eigenvektor* dargestellt. Der Betrag, um den der Original-Vektor durch die Matrix-Multiplikation skaliert wurde war derselbe: Im Beispiel ist 4 der *Eigenwert* der zu dem Eigenvektor gehört. Unabhängig davon welches Vielfache des Eigenvektors vor der Multiplikation mit der quadratischen Matrix genommen wird, ergibt sich immer der um den Faktor 4 skalierte Vektor als Ergebnis (wie in dem *Beispiel für einen skalierten Eigenvektor*).

Eigenvektoren und Eigenwerte treten also immer paarweise auf. Werden mit Hilfe einer Software Eigenvektoren berechnet, so werden üblicherweise auch die Eigenwerte ausgegeben.

Deutlich wird das noch einmal mit der mathematischen Definition: Sei **A** eine $n \times n$ Matrix. Erfüllen ein Skalar λ und ein n-dimensionaler Spaltenvektor $\boldsymbol{u}$ mit $\boldsymbol{u} \neq \mathbf{0}$ das Gleichungssystem

$$\boldsymbol{A}\boldsymbol{u} = \lambda \boldsymbol{u}$$

so heißt λ Eigenwert von **A** und $\boldsymbol{u}$ zugehöriger Eigenvektor von **A**.

Erfüllt ein Vektor $\boldsymbol{u}$ diese Gleichung, so erfüllt auch jedes Vielfache von $\boldsymbol{u}$ die Gleichung. Um eine Lösung zu erhalten, wird umgeformt in

$$(\boldsymbol{A} - \lambda \boldsymbol{I_n})\boldsymbol{u} = 0$$

Hierbei ist $\boldsymbol{I_n}$ die Einheitsmatrix. Für ein festes λ ist ein lineares homogenes Gleichungssystem entstanden. Dies besitzt genau dann Lösungen, die ungleich dem Nullvektor sind, wenn die Spalten von $\boldsymbol{A} - \lambda \boldsymbol{I_n}$ linear abhängig sind. Dies ist genau dann der Fall, wenn gilt

$$|\boldsymbol{A} - \lambda \boldsymbol{I_n}| = 0$$

Diese Gleichung ist ein Polynom n-ten Grades in λ. Dieses besitzt genau n Nullstellen. Die Nullstellen $\lambda_1, \dots, \lambda_n$ des Polynoms sind also die Eigenwerte der Matrix **A.** Diese Eigenwerte müssen nicht notwendigerweise verschieden sein. Meist werden die Eigenwerte der Größe nach durchnummeriert, wobei der erste Eigenwert der größte ist.

Beispiel

Die Eigenwerte und normierten Eigenvektoren folgender Matrix sollen bestimmt werden:

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

Es gilt

$$\boldsymbol{A} - \lambda \boldsymbol{I_2} = \begin{pmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{pmatrix}$$

Somit gilt

$$|\boldsymbol{A} - \lambda \boldsymbol{I_2}| = (2-\lambda)^2 - 1$$

Ein Eigenwert λ erfüllt also die Gleichung

$$(2-\lambda)^2 - 1 = 0$$

Die Nullstellen und somit die Eigenwerte von $\boldsymbol{A}$ sind $\lambda_1 = 3$ und $\lambda_2 = 1$.

Die Eigenvektoren zum Eigenwert $\lambda_i, i = 1, \dots, n$ werden dadurch erhalten, dass λ_i in die Definitionsgleichung eingesetzt wird und die Lösungsmenge des dadurch entstandenen linearen homogenen Gleichungssystems bestimmt wird.

Der zu $\lambda_1 = 3$ gehörende Eigenvektor

$$\boldsymbol{u_1} = \begin{pmatrix} u_{11} \\ u_{21} \end{pmatrix}$$

erfüllt also das Gleichungssystem

$$(\boldsymbol{A} - 3\boldsymbol{I_2})\boldsymbol{u} = 0$$

Wegen

$$(\boldsymbol{A} - 3\boldsymbol{I_2}) = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}$$

ergibt sich

$$-u_{11} + u_{21} = 0$$

$$u_{11} - u_{21} = 0$$

Für die Komponenten des Eigenvektors $\boldsymbol{u_1}$ zum Eigenwert $\lambda_1 = 3$ muss also gelten

$$u_{11} = u_{21}$$

Der Vektor

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

und alle Vielfachen dieses Vektors sind Eigenvektoren zum Eigenwert $\lambda_1 = 3$.

Analoge Berechnungen zum Eigenwert $\lambda_2 = 1$ ergeben, dass für die Komponenten $u_{12}\text{und}\ u_{22}$ des zu $\lambda_2 = 1$ gehörenden Eigenvektors $\boldsymbol{u_2}$ die Beziehung

$$u_{12} = -u_{22}$$

gelten muss. Der Vektor

$$\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

und alle Vielfachen dieses Vektors sind Eigenvektoren zum Eigenwert $\lambda_2 = 1$.

Für die normierten Eigenvektoren ergibt sich

$$\boldsymbol{u_1} = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$$

und

$$\boldsymbol{u_2} = \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}$$