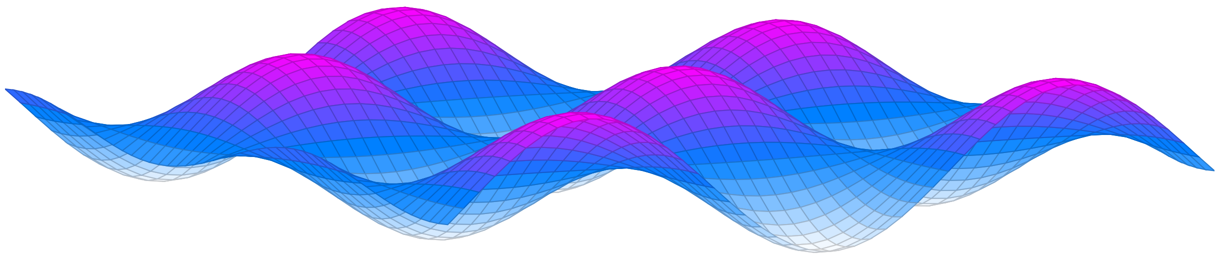


LIAL - Lineare Algebra

Matrizen, Vektoren, Gleichungssysteme und mehr

Author:
Eldar Omerovic
https://www.github.com/omeldar

Documentation Repository:
https://www.github.com/omeldar/hslu

Credits:
I wrote this summary within the LIAL class
at the Lucerne University of Applied Sciences and Arts.

June 2025

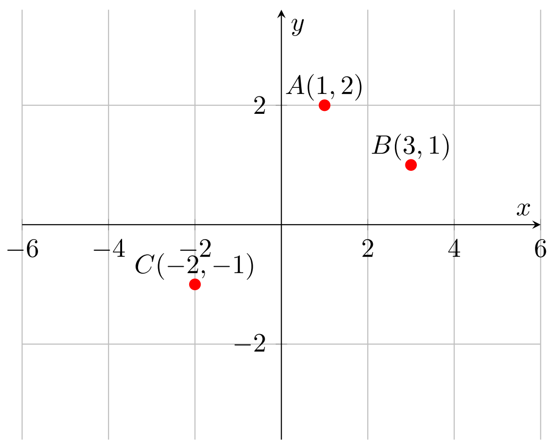
Contents

1 Vektoren, Matrizen und lineare Gleichungssysteme

Einführung in Vektoren, lineare Glecihungssysteme und Matrizen.

1.1 Vektoren

1.1.1 Punkte im zweidimensionalen kartesischen Koordinatensystem



Das zweidimensionale kartesische Koordinatensystem besteht aus:

- x-Achse (Abszisse, horizontal)
- y-Achse (Ordinate, vertikal)

In einem kartesischen Koordinatensystem stehen die Achsen senkrecht aufeinander.

Punkte werden durch Zahlen (Koordinaten) beschrieben, in unserem Beispiel:

- A (1, 2)
- B (3, 1)
- C (-2, -1)

1.1.2 Skalare und Vektoren

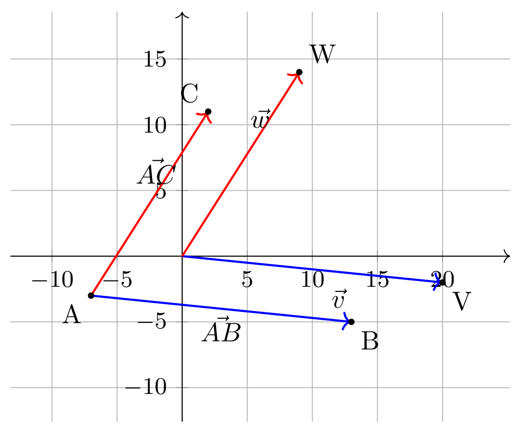
Beispiele physikalischer Grössen:

- **Skalar**: Temperatur, Druck, Luftfeuchtigkeit
- **Vektoriell**: Windgeschwindigkeit an einem Ort, Kraft auf einen Körper, Fliessgeschwindigkeit in einem Gewässer

Ein Vektor

- hat eine Länge,
- hat eine Richtung,
- lässt sich geometrisch als Pfeil darstellen,
- kann arithmetisch (in einem kartesischen Koordinatensystem) durch Zahlen beschrieben werden.

1.1.3 Vektoren im zweidimensionalen kartesischen Koordinatensystem



Ein Vektor verbindet seinen Anfangspunkt mit seinem Endpunkt:

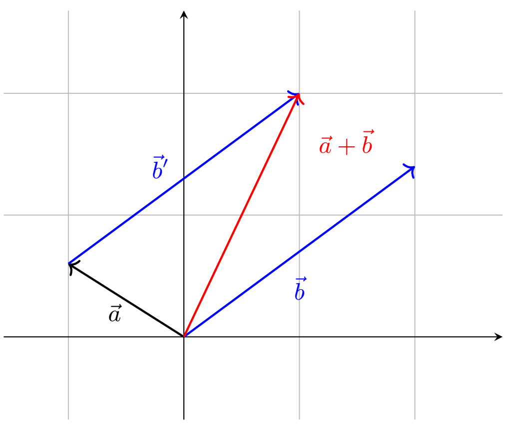
- $\vec{AB}$ verbindet $A(-7/-3)$ mit $B(2/4)$
- $\vec{AC}$ verbindet A mit $C(2/11)$

Zwei Vektoren mit gleicher Länge und gleicher Richtung sind gleich:

- $v = \vec{AB}$
- $w = \vec{AC}$
- Vektoren kann man parallel verschieben.
- Ein Ortsvektor hat den Ursprung des Koordinatensystems als Anfangspunkt und verbindet diesen mit seinem Endpunkt.
  - $\vec{v}$ ist Ortsvektor zum Punkt $V(20/-2)$
  - $\vec{w}$ ist Ortsvektor zum Punkt $W(9/14)$
- Ortsvektoren und ihre Endpunkte haben die gleichen Koordinaten.
- Schreibweise: Die Koordinaten der Vektoren werden übereinander in runden Klammern angegeben:

$$\mathbf{v} = \begin{pmatrix} 20 \\ -2 \end{pmatrix}, \quad \mathbf{w} = \begin{pmatrix} 9 \\ 14 \end{pmatrix}$$

1.1.4 Addition zweier Vektoren



- Vektor $\vec{b}$ wird in den Vektor $\vec{b}'$ verschoben, so dass der Anfangspunkt von $\vec{b}'$ am Endpunkt von $\vec{a}$ liegt.
- Die Summe $\vec{a}+\vec{b}$ ist ein Vektor vom Fusspunkt von $\vec{a}$ zur Spitze von $\vec{b}'$.

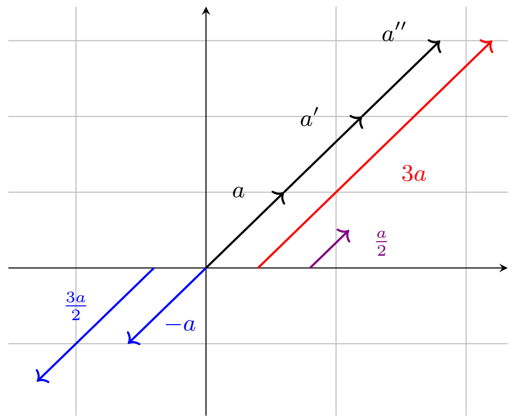
Ein Beispiel für die arithmetische Addition von Vektoren:

$$a = \begin{pmatrix} -5 \\ 3 \end{pmatrix}, \quad b = \begin{pmatrix} 11 \\ 7 \end{pmatrix} \text{ und } \quad \vec{a}+\vec{b} = \begin{pmatrix} 6 \\ 10 \end{pmatrix}$$

Allgemein werden Vektoren nach einer einfachen regel addiert: Indem man die entsprechenden Komponenten addiert:

$$\vec{a}+\vec{b} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} a_1+b_1 \\ a_2+b_2 \end{pmatrix}$$

1.1.5 Skalare Multiplikation eines Vektors



- Addiert man den Vektor $\vec{a}$ zweimal zu sich selbst (a' und a'') so erhält man den Vektor $3a$. Dieser ist dreimal so lang wie a und hat dieselbe Richtung.
- Multipliziert man den Vektor a mit -1, so hat das Ergebnis dieselbe Länge wie a, weist aber in die entgegengesetzte Richtung ($-a$).

Ein Beispiel für die arithmetische skalare Multiplikation eines Vektors:

$$a = \begin{pmatrix} 3 \\ 2 \end{pmatrix} \text{ und } \quad 3a = 3\begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 9 \\ 6 \end{pmatrix} = \begin{pmatrix} 3\cdot 3 \\ 3\cdot 2 \end{pmatrix}$$

Allgemein gilt folgende Regel:

$$\lambda a = \lambda \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} \lambda a_1 \\ \lambda a_2 \end{pmatrix}$$

1.1.6 Subtraktion zweier Vektoren

Den Vektor b vom Vektor a subtrahieren heisst, den Vektor b mit -1 zu multiplizieren und das Ergebnis zum Vektor a zu addieren:

$$\begin{aligned} a - b = a + (-1) \cdot b &= \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + (-1) \cdot \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \\ &= \begin{pmatrix} -b_1 \\ -b_2 \end{pmatrix} = \begin{pmatrix} a_1 - b_1 \\ a_2 - b_2 \end{pmatrix} \end{aligned}$$

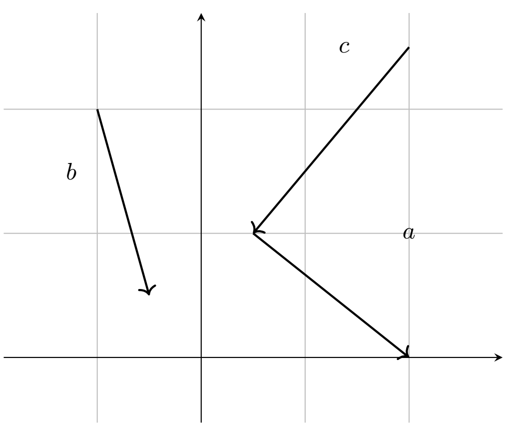
Geometrisch kann dies wie folgt interpretiert werden: Liegen die Fusspunkte von a und b am selben Ort, ist die Differenz $a - b$ ein Vektor, der von der Spitze von b zur Spitze von a verläuft.

1.1.7 Verallgemeinerung auf höhrere Dimensionen

Das bisher Gesagte gilt auch für Vektoren in höheren Dimensionen. Diese haben mehr Koordinaten, aber die Regeln für Addition, Subtraktion und skalare Multiplikation bleiben gleich.

Vektoren in n-dimensionalen Räumen bei denen $n > 3$ entziehen sich unserer Vorstellung. Diese können wir auch nicht mehr geometrisch darstellen. Dennoch können wir sie mathematisch behandeln.

1.1.8 Linearkombinationen



Multipliziert man eine Anzahl von Vektoren mit einer entsprechenden Anzahl Skalaren, und addiert man alle Ergebnisse der Skalarmultiplikationen, so spricht man von einer Linearkombination der Vektoren.

1.1.9 Regeln für das Rechnen mit Vektoren

(1)	$a + b = b + a$	Kommutativgesetz
(2)	$(a + b) + c = a + (b + c)$	Assoziativgesetz
(3)	$a + 0 = a$	Neutrales Element
(4)	$a + (-a) = 0$	Inverse
(5)	$\lambda(a + b) = \lambda a + \lambda b$	
(6)	$(\lambda + \mu)a = \lambda a + \mu a$	
(7)	$(\lambda\mu)a = \lambda(\mu a) = \mu(\lambda a)$	
(8)	$1a = a$	

Dabei sind a, b und c beliebige (auch n-dimensionale) Vektoren und λ und μ beliebige (reelle) Zahlen.

1.2 Schnittpunkte von Geraden und Ebenen

1.2.1 Schnittpunkt zweier Geraden

Zeigt die Spitze eines Vektors $\vec{v}$ zu einem Punkt auf der Geraden g, und verläuft die Gerade in Richtung des Vektors $\vec{u}$, so erzeugt folgende Gleichung alle Punkte der Geraden, wenn λ alle reellen Zahlen durchläuft:

$$g : r = \vec{v} + \lambda\vec{u}$$

Beispiel
Wir möchten den Schnittpunkt von zwei Geraden bestimmen:

$$g : r = \begin{pmatrix} -3 \\ 2 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ 2 \end{pmatrix} \text{ und } \quad h : r = \begin{pmatrix} 4 \\ 3 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ -2 \end{pmatrix}$$

Wir müssen nun beide Geraden gleichsetzen:

$$\vec{r_1} = \vec{r_2} = \begin{pmatrix} -3 \\ 2 \end{pmatrix} + \lambda_1 \begin{pmatrix} 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ -2 \end{pmatrix}$$

Das ergibt zwei Gleichungen:

$$\begin{aligned} -3 + 2\lambda &= 4 + \mu \\ 2 + 2\lambda &= 3 - 2\mu \end{aligned}$$

Wenn man das umformt und das resultierende Gleichungssystem löst, erhält man:

$$\lambda = 2.5, \quad \mu = -2$$

Dies muss man nun in eine der Geradengleichungen einsetzen, um den Schntitpunkt zu erhalten:

$$\vec{r} = \begin{pmatrix} -3 \\ 2 \end{pmatrix} + 2.5 \cdot \begin{pmatrix} 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ 7 \end{pmatrix}$$

1.2.2 Schnittpunkt einer Geraden mit einer Ebene

Zeigt die Spitze des Vektors $\vec{v}$ zu einem Punkt auf der Ebene E, und verläuft die Ebene in Richtung der Vektoren $\vec{u_1}$ und $\vec{u_2}$, so erzeugt folgende Gleichung alle Punkte der Ebene, wenn λ und μ alle reellen Zahlen durchlaufen:

$$E : r = \vec{v} + \lambda\vec{u_1} + \mu\vec{u_2}$$

Beispiel
Wir möchten den Schnittpunkt der Gerade g mit der Ebene E bestimmen:

$$g : r = \begin{pmatrix}1\\0\\2\end{pmatrix} + \lambda\begin{pmatrix}1\\2\\1\end{pmatrix}, E : r = \begin{pmatrix}2\\3\\1\end{pmatrix} + \lambda\begin{pmatrix}1\\2\\1\end{pmatrix} + \mu\begin{pmatrix}2\\-1\\3\end{pmatrix}$$

Wir setzen die Geradengleichung gleich der Ebenengleichung, um den Schnittpunkt zu finden.Dies ergibt wieder ein lineares Gleichungssystem mit 3 Gleichungen für die 3 Unbekannten t, λ, μ:

$$\begin{pmatrix}1\\0\\2\end{pmatrix} + t\begin{pmatrix}1\\2\\1\end{pmatrix} = \begin{pmatrix}2\\3\\1\end{pmatrix} + \lambda\begin{pmatrix}1\\2\\1\end{pmatrix} + \mu\begin{pmatrix}2\\-1\\3\end{pmatrix}$$

Wenn wir das umformen und gleichsetzen, erhalten wir:

$$t = 2 + 2\lambda - 2\mu$$
$$2 + 2t = 3 + \lambda + \mu$$
$$t = 1 + \lambda + 3\mu$$

Hier erhält man für $t = 2.00$, $\lambda = 1.00$ und $\mu = -1.00$. Wenn man dies wieder einsetzt, erhält man den Schnittpunkt:

$$\begin{pmatrix}3\\4\\4\end{pmatrix}$$

1.3 Matrizen

Eine Matrix ist ein rechteckiges Schema von Zahlen. Eine (2 x 3)-Matrix hat 2 Zeilen und 3 Spalten. $a_{a,2} = 3$ ist das Element in der 1. Zeile und 2. Spalte.

$$A = \begin{pmatrix}2 & 3 & -1\\1 & -2 & 1\end{pmatrix}$$

Eine $(m \times n)$-Matrix besteht aus m Zeilen und n Spalten. Das Element $a_{i,j} \in \Re$ steht in der i-ten Zeile und j-ten Spalte.

$$\mathbf{A} = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,j} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,j} & \cdots & a_{2,n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{i,1} & a_{i,2} & \cdots & a_{i,j} & \cdots & a_{i,n} \\ \vdots & \vdots & & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,j} & \cdots & a_{m,n} \end{pmatrix}$$

1.3.1 Eine Matrix mit einem Vektor multiplizieren

Will man eine Matrix A mit einem Vektor x multiplizieren, so muss die Anzahl Spalten von A gleich der Anzahl Zeilen von x sein. Das Ergebnis dieser Multiplikation ist ein Vektor:

$$\vec{y} = Ax = \begin{pmatrix}1 & 2 & 4\\3 & 1 & -1\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}$$

$$= \begin{pmatrix}1\cdot x + 2\cdot y + 4\cdot z\\3\cdot x + 1\cdot y + (-1)\cdot z\end{pmatrix} = \begin{pmatrix}1x + 2y + 4z\\3x + 1y - 1z\end{pmatrix}$$

Mit der rechten Seite dieser Gleichung (RHS) b können wir dies kurz schreiben als:

$$Ax = b \qquad , b = \begin{pmatrix}1\\-1\end{pmatrix}$$

Wenn uns b bekannt ist, können wir mithilfe der Koeffizientenmatrix A, dem Lösungsvektor x und der rechten Seite (RHS) auf das Gleichungssystem schliessen um die unbekanten Variablen x, y und z zu bestimmen:

$$x + 2y + 4z = 1$$
$$3x + y - z = -1$$

1.4 Lineare Gleichungen und Gleichungssysteme

1.4.1 Lineare Gleichungen

Lineare Gleichungen ergeben sich aus vielen praktischen Aufgabenstellungen.
Beispiel
Ein Fahrzeug fährt mit konstanster Geschwindigkeit v. Nach 45 Minuten hat es die Strecke 120km zurückgelegt. Bestimmen Sie die Geschwindigkeit v des Fahrzeugs.
Lösung:

$$v = \frac{s}{t} = \frac{120}{0.75} = 160\text{km/h}$$

1.4.2 Lineare Gleichungssysteme (LGS)

Das Fahrzeug A fährt mit konstanter Geschwindigkeit $v_A = 160\text{km/h}$ entlang einer Gerade. Das Fahrzeug B hat $d_B = 100\text{km}$ Vorsprung und fährt mit konstanter Geschwindigkeit $v_B = 120\text{km/h}$. Wann und wo treffen sich die beiden Fahrzeuge?

$$\begin{aligned} FzA: \quad & s_A - v_A \cdot t_A = 0 \\ FzB: \quad & s_B - v_B \cdot t_B = d_B \end{aligned}$$

Dabei stehen s_A und s_B für den Ort der Fahrzeuge vom Startpunkt von A aus gemessen, nachdem die Fahrzeuge die Zeiten t_A und t_B gefahren sind. Für den Treffpunkt gilt $s = s_A \equiv s_B$ und $t = t_A \equiv t_B$ und (t, s) ergibt sich als Lösung des LGS:

$$\begin{aligned} s - v_A \cdot t = 0 \\ s - v_B \cdot t = d_B \end{aligned}$$

Wenn wir uns das als Matrix anschauen, sieht die Koeffizientenmatrix des LGS wie folgt aus:

$$\begin{bmatrix} 1 & -160 \\ 1 & -120 \end{bmatrix} \begin{bmatrix} s \\ t \end{bmatrix} = \begin{bmatrix} 0 \\ 100 \end{bmatrix}$$

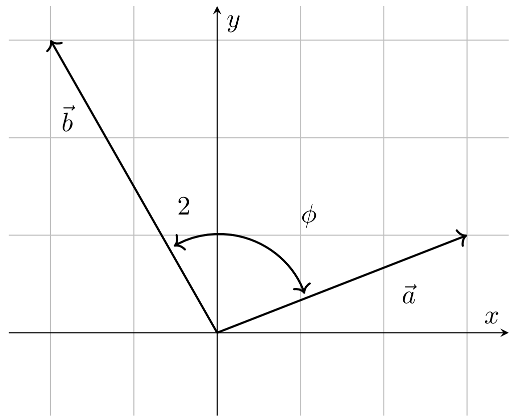
Wenn wir die erste Gleichung von der zweiten subtrahieren, erhalten wir:

$$\begin{bmatrix} 1 & -160 \\ 0 & 40 \end{bmatrix} \begin{bmatrix} s \\ t \end{bmatrix} = \begin{bmatrix} 0 \\ 100 \end{bmatrix}$$

Wir erhalten also für $t = \frac{100}{40} = 2.5$ Stunden und $s = 160 \cdot 2.5 = 400$ km den Treffpunkt der beiden Fahrzeuge. Dies sieht man etwas einfacher, wenn man dies als normales Gleichungssystem schreibt:

1.5 Das Skalarprodukt

Das Skalarprodukt zweier Vektoren a und b ist definiert durch $a \bullet b = |a| \cdot |b| \cdot cos\phi$. Dabei ist ϕ der Winkel zwischen den Vektoren a und b.



In kartesischen Koordinaten ist das Skalarprodukt einfach:

$$a \bullet b = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \bullet \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = a_1 b_1 + a_2 b_2$$

1.5.1 Berechnung von Winkel und Länge (Betrag) eines Vektors

Mit Hilfe des Skalarprodukts lässt sich der Betrag (Länge) a wie folgt definieren:

$$a \equiv |a| = \sqrt{a \bullet a} = |a|^2 = a \bullet a$$

Dann gilt im 2D-Fall (Satz von Pythagoras):

$$a \equiv |a| = \sqrt{a \bullet a} = \sqrt{a_1^2 + a_2^2}$$

Beispiel Gegeben seien zwei Vektoren:

$$\vec{a} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} 4 \\ -1 \end{pmatrix}$$

Wenn wir dies in die Formel einsetzen erhalten wir das Skalarprodukt wie folgt:

$$\vec{a} \bullet \vec{b} = 2 \cdot 4 + 3 \cdot (-1) = 8 - 3 = 5$$

Nun berechnen wir den Betrag der Vektoren:

$$\text{Für } \vec{a}: \quad |\vec{a}| = \sqrt{2^2 + 3^2} = \sqrt{4+9} = \sqrt{13} \approx 3.61$$
$$\text{Für } \vec{b}: \quad |\vec{b}| = \sqrt{4^2 + (-1)^2} = \sqrt{16+1} = \sqrt{17} \approx 4.12$$

Und in einem letzten Schritt noch den Winkel zwischen den Vektoren mit der Formel:

$$cos\phi = \frac{a \bullet b}{|a| \cdot |b|} = \frac{5}{\sqrt{13} \cdot \sqrt{17}} \approx 0.35$$

Und so dann schliesslich den Winkel:

$$cos^{-1}(0.35) \approx 70°$$

1.5.2 Rechenregeln für das Skalarprodukt und Einheitsvektor

- $a \bullet b = b \bullet a$ (Kommutativgesetz)
- $a \bullet (b + c) = a \bullet b + a \bullet c$ (Distributivgesetz)
- $\lambda(a \bullet b) = (\lambda a) \bullet b = a \bullet (\lambda b)$

Beispiel
Um einen Vektor $\vec{v}$ zu normieren, berechnen wir auch wieder zuerst den Betrag des Vektors. Wir können ihn hier als gegeben betrachten, da wir dies ja schon im Beispiel oben gemacht haben:

$$\vec{v} = \begin{pmatrix} 6 \\ 8 \end{pmatrix}, \quad |\vec{v}| = 10$$

Um den normierten Vektoren zu erhalten, teilen wir den Vektor durch seinen Betrag:

$$\vec{v}_{norm} = \frac{\vec{v}}{|\vec{v}|} = \frac{1}{10} \cdot \begin{pmatrix} 6 \\ 8 \end{pmatrix} = \begin{pmatrix} 0.6 \\ 0.8 \end{pmatrix}$$

Dies ist nun der Einheitsvektor in die Richtuung von $\vec{v}$. Ein Einheitsvektor hat immer die Länge 1.

1.5.3 Paarweise senkrechte Vektoren

Zwei Vektoren $\vec{a}$ und $\vec{b}$ stehen genau dann senkrecht aufeinander, sind also orthogonal, falls ihr Skalarprodukt verschwindet (= 0 ist).

$$a \bullet b = 0 \longleftrightarrow \vec{a} \perp \vec{b}$$

Hier haben wir angenommen, dass der Nullvektor zu jedem Vektor senkrecht steht.

Beispiel
Wenn wir einen senkrechten Vektor $\vec{v}$ zu einem gegebenen Vektor $\vec{a}$ suchen, so können wir dies ganz einfach auf diese zwei Arten tun:

$$\vec{a} = \begin{pmatrix} -a_y \\ a_x \end{pmatrix} \quad \text{oder} \quad \vec{a} = \begin{pmatrix} a_y \\ -a_x \end{pmatrix}$$

Dies würde in einem konkreten Beispiel so aussehen:

$$\vec{a} = \begin{pmatrix} 1 \\ 4 \end{pmatrix} \text{ und somit } \quad \vec{v} = \begin{pmatrix} -4 \\ 1 \end{pmatrix} \text{ oder } \begin{pmatrix} 4 \\ -1 \end{pmatrix}$$

Generell gibt es unendlich viele senkrechte Vektoren. Sie unterscheiden sich aber in ihrer Länge. Jedes Vielfache einer der Optionen für $\vec{v}$ ist ebenfalls ein senkrechter Vektor zu $\vec{a}$.

2 Lineare Gleichungssysteme : Gauss-Algorithmus

2.1 Lineare Gleichungssysteme

2.1.1 Einfache Gleichungen

Eine einfache Gleichung ist $x - 5 = 10$. Wenn man diese Glecihung nach x auflöst, ergibt sich $x = 15$. Setzt man diesen Wert ein, ist die Gleichung erfüllt: $15 - 5 = 10$.

Eine andere einfache Gleichung ist $x^2 - 9 = 16$. Das Lösen dieser Gelichung liefert zwei Werte, $x_1 = 5$ und $x_2 = -5$. Beide Werte erfüllen die Gleichung.

Je nach Typ der Gleichung gibt es unterschiedliche Lösungsverfahren und unterschiedlich viele Lösungen.

2.1.2 Gleichungssystem

Ein Gleichungssystem besteht aus mehreren Gleichungen, die alle gemeinsam erfüllt sein sollen.

Einige Beispiele
Ein lineares Gleichungssystem mit zwei Gleichungen und zwei Unbekannten:

$$\begin{aligned} 3x + 5y &= -10 \\ 0.3x + 6y &= 15.78 \end{aligned}$$

Ein lineares Gleichungssystem mit zwei Gleichungen und drei Unbekannten:

$$\begin{aligned} 3x + 5y - 4z - 28 &= 0 \\ 0.3z + 7x &= 15.78 \end{aligned}$$

Das vorangehende Gleichungssystem kann in **systematischer Notation** wie folgt geschrieben werden:

$$\begin{aligned} 3x + 5y - 4z &= 28 \\ 7x + 9y + 0.3z &= 15.78 \end{aligned}$$

Definition
Ein System, das nur aus Gleichungen mit folgenden Eigenschaften besteht, ist ein lineares Gleichungssystem:

1. Auf einer Seite kommen keine Variablen, sondern nur Konstanten vor.
2. Auf der anderen Seite stehen Summen, die aus Produkten jeweils einer Variablen mit einem konstanten Faktor, auch *Koeffizient* genannt, bestehen.
3. Die Variablen kommen nur in erster Potenz vor.

- *Tipp: Nichtlineare Funktionen wie sin, cos, ln, exp, etc. sind nicht erlaubt.*

2.1.3 Matrixschreibweise

Ein Gleichungssystem kann auch in Matrixschreibweise dargestellt werden. Dabei wird die Koeffizientenmatrix, die Variablenmatrix und die Ergebnismatrix verwendet.

So kann beispielweise die folgende Gleichung:

$$\begin{aligned} 6x - y - z &= 4 \\ x + y + 10z &= -6 \end{aligned}$$

in Matrixschreibweise dargestellt werden:

$$\begin{pmatrix} 6 & -1 & -1 \\ 1 & 1 & 10 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 4 \\ -6 \end{pmatrix}$$

2.2 Lösung linearer Gleichungssysteme

Das Gauss-Eliminationsverfahren ist ein Standardlösungsverfahren zur Lösung linearer Gleichungssysteme und beruht darauf, dass spezielle Gleichungssysteme einfach zu lösen sind.

$$\begin{aligned} 2x + 5y - 2z &= -3 \\ 3y + 2z &= 7 \\ 5z &= 10 \end{aligned}$$

- Aus der letzten Gleichung folgt direkt $z = 2$.
- Setzt man das in die zweite Gleichung ein, löst nach y auf, erhält man $y = 1$.
- Setzt man anschliessend die Werte x und y in die erste Gleichung ein, und erhält $x = -2$. Die Lösung des Gleichungssystems ist also $(x, y, z) = (-2, 1, 2)$.

2.2.1 Leitkoeffizienten und Pivotelemente

Den ersten Eintrag einer Zeile, der nicht verschwindet, nennt man Leitkoeffizient (hier rot):

$$\left(\begin{array}{ccc|c} \mathbf{2} & 5 & -2 & -3 \\ \mathbf{3} & 0 & 2 & 7 \\ 0 & 0 & \mathbf{5} & 10 \end{array}\right)$$

Steht in einer Matrix jeder Leitkoeffizient weiter rechts als der Leitkoeffizient in der Zeile darüber, und stehen alle Zeilen ohne Leitkoeffizient (diejenige mit lauter Nullen) ganz unten, hat die Matrix die *Zeilenstufenformn*. In diesem Fall nennt man die Leitkoeffizienten auch *Pivotelemente* (hier rot).

$$\left(\begin{array}{cccc|c} \mathbf{2} & 5 & -8 & -1 & -3 \\ 0 & 0 & \mathbf{2} & 4 & 7 \\ 0 & 0 & 0 & \mathbf{9} & 10 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

Die Leitkoeffizienten müssen nicht zwangsläufig alle auf der Hauptdiagonalen stehen.

2.2.2 Gauss-Eliminationsverfahren

Beim Gauss-Verfahren geht es darum, lineare Gleichungssysteme so umzuformen, dass sich eine Zeilenstufenform ergibt.

Es gibt drei Typen elementarer Zeilenumformungen, die man auf die Matrix anwenden kann:

1. Man darf eine Zeile der erweitwerten Koeffizientenmatrix mit einem Faktor, der nicht Null ist, multiplizieren. Beispiel:

$$\begin{aligned} 2x - 6y &= 1 \qquad (\times 2) \\ 4x - 12y &= 2 \end{aligned}$$

2. Man darf zwei Zeilen der erweiterten Koeffizientenmatrix vertauschen.

3. Man darf das Vielfache einer Zeile zu einer anderen addieren. Beispiel:

$$\begin{aligned} 2x - 6y &= 3 \qquad (a) \\ -4x - 12y &= 2 \qquad (b) \end{aligned}$$

Dabei entsteht dann folgendes:

$$\begin{aligned} 2x - 6y &= 3 \qquad (a) \\ -8x &= -4 \qquad (b') \end{aligned}$$

Beispiel
Wir wollen folgendes Gleichungssystem in die Zeilenstufenform bringen:

$$\begin{aligned} 6x - y - z &= 4 \qquad (\mathrm{I}) \\ x + y + 10z &= -6 \qquad (\mathrm{II}) \\ 2x - y + z &= -2 \qquad (\mathrm{III}) \end{aligned}$$

Zuerst schreiben wir das in die Matrixschreibweise:

$$\left(\begin{array}{ccc|c} 6 & -1 & -1 & 4 \\ 1 & 1 & 10 & -6 \\ 2 & -1 & 1 & -2 \end{array}\right)$$

Das Pivotelement (6) ist eigentlich an der richtigen Stelle. Da das Rechnen mit 1 oder -1 leichter ist, vertauschen wir die erste Zeile mit der zweiten Zeile.

Um den zweiten Leitkoeffizienten zu erzeugen, subtrahieren wir das 6-fache der ersten Zeile von der zweiten Zeile.

$$\begin{array}{c} \mathrm{I} \\ \mathrm{II} \\ \mathrm{III} \end{array} \left(\begin{array}{ccc|c} \mathbf{1} & 1 & 10 & -6 \\ 6 & -1 & -1 & 4 \\ 2 & -1 & 1 & -2 \end{array}\right) \quad \begin{array}{c} \\ \mathrm{II} - 6 \cdot \mathrm{I} \\ \mathrm{III} - 2 \cdot \mathrm{I} \end{array}$$

Dies gibt dann folgendes:

$$\begin{array}{c} \mathrm{I} \\ \mathrm{II} \\ \mathrm{III} \end{array} \left(\begin{array}{ccc|c} 1 & 1 & 10 & -6 \\ \mathbf{0} & \mathbf{-7} & \mathbf{-61} & \mathbf{40} \\ 0 & -3 & -19 & 10 \end{array}\right)$$

Für die dritte Zeile machen wir das selbe: $\mathrm{III} - 2 \cdot \mathrm{I}$. Auf Kosten grösserer Zahlen vermeiden wir unangenehme Brüche. Wir wenden hier die Zeilenumforumg vom Typ III an: Erst die dritte Zeile mit 7 multiplizieren:

$$\begin{array}{c} \mathrm{I} \\ \mathrm{II} \\ \mathrm{III} \end{array} \left(\begin{array}{ccc|c} 1 & 1 & 10 & -6 \\ 0 & -7 & -61 & 40 \\ \mathbf{0} & \mathbf{-21} & \mathbf{-133} & \mathbf{70} \end{array}\right) \quad \begin{array}{c} \\ \\ \mathrm{III} - 3 \cdot \mathrm{II} \end{array}$$

Und dann das dreifache der zweiten Zeile von der dritten Zeile subtrahieren:

$$\begin{array}{c} \mathrm{I} \\ \mathrm{II} \\ \mathrm{III} \end{array} \left(\begin{array}{ccc|c} 1 & 1 & 10 & -6 \\ 0 & -7 & -61 & 40 \\ \mathbf{0} & \mathbf{0} & \mathbf{50} & \mathbf{-50} \end{array}\right)$$

In dem wir rückwärts einsetzen (wie im Beispiel oben), erhalten wir die Lösung des Gleichungssystems: $(1, 3, -1)$.

2.2.3 Spezialfälle

Steht in einer Zeile in einer Gleichung folgendes:

$$0 \quad 0 \quad 0 \mid 5$$

Ist die Gleichung nicht lösbar. Denn dies ist ein Widerspruch. Eine Zeile mit nur Nullen in den Matrix und dem Resultat 5 kann nicht lösbar sein.

Anders sieht es aus, wenn wir eine Zeile mit nur Nullen haben:

$$0 \quad 0 \quad 0 \mid 0$$

Dann gibt es unendlich viele Lösungen. Während dem Gauss Verfahren kann man diese Zeilen einfach mit den unteren vertauschen um sie an das Ende zu bringen. Im Gauss-Verfahren selbst gibt es dort dann nichts mehr zu rechnen.

Jedoch muss man beim rückwärts einsetzen diese wieder mit einbeziehen. Wenn wir ein Resultat wie das folgende haben:

$$\left[\begin{array}{cccc|c} 1 & 4 & -1 & 1 & 2 \\ 0 & 10 & -4 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Wir haben hier 3 Zeilen ohne Pivotelement (Nullzeilen). Das bedeutet, dass wir 3 Parameter brauchen, um die Werte zu beschreiben.

Hier aber haben wir nicht noch 3 zusätzliche Spalten. Also haben wir nicht 3 weitere Unbekannte sondern nur 2. Das heisst wir brauchen nur 2 Parameter. Dies könnte z.B. so aussehen:

$$r = x_3, s = x_4$$

Dies können wir nun in die zwei verbleibenden Gleichungen einsetzen und dies dann nach x_1 und x_2 auflösen.

Als Lösung erhalten wir dann:

$$\begin{aligned} x_1 &= 2 - 4 \cdot \frac{1 + 4r - s}{10} + r - s \\ x_2 &= \frac{1 + 4r - s}{10} \\ x_3 &= r \\ x_4 &= s \end{aligned}$$

3 Lineare Gleichungssysteme : LU-Zerlegung, Pivotisierung

3.1 Matrixmultiplikation

Die Matrixmultiplikation wird wie folgt definiert:

$\mathbf{A} \in \Re^{l\times m}$ ist eine Matrix mit l Zeilen, m Spalten und den Elementen a_{ij}.
$\mathbf{B} \in \Re^{m\times n}$ ist eine Matrix mit m Zeilen, n Spalten und den Elementen b_{jk}.
Das Produkt $\mathbf{AB} = \mathbf{C} \in \Re^{l\times n}$ hat l Zeilen, n Spalten und die Elemente:

$$c_{ik} = \sum_{j=1}^{m} a_{ij}b_{jk} \ , \quad i = 1, 2, ...l, \quad k = 1, 2, ...n.$$

Beispiel
Wenn wir Matrizen multiplizieren wollen, müssen die Anzahl der Spalten der ersten Matrix gleich der Anzahl der Zeilen der zweiten Matrix sein:

$$\begin{pmatrix} \mathbf{4} & \mathbf{-2} & \mathbf{-2} & \mathbf{0} \\ \mathbf{-2} & \mathbf{-7} & \mathbf{3} & \mathbf{8} \\ \mathbf{0} & \mathbf{1} & \mathbf{-2} & \mathbf{-1} \end{pmatrix} \cdot \begin{pmatrix} \mathbf{4} & \mathbf{-5} \\ \mathbf{3} & \mathbf{-1} \\ \mathbf{6} & \mathbf{4} \\ \mathbf{0} & \mathbf{-3} \end{pmatrix}$$

Dies ist hier gegeben, also können wir diese zwei Matrizen multiplizieren. Die Lösung ist eine Matrix mit gleich vielen Zeilen wie in der ersten Matrix und gleich vielen Spalten wie in der zweiten Matrix. Hier also eine Matrix mit 3 Zeilen und 2 Spalten.

Wir gehen jetzt Schritt für Schritt durch, um die einzelnen Werte der Lösungsmatrix zu berechnen:

1. Um den ersten Wert zu finden, nehmen wir uns die erste Zeile der ersten Matrix und die erste Spalte der zweiten Matrix und berechnen daraus die Summe der Produkte.

   Dabei multiplizieren wir das n-te Element aus Matrix A in Zeile 1 mit dem n-ten Element aus Matrix B in Spalte 1:

   $$\mathbf{4}\cdot\mathbf{4}+\mathbf{-2}\cdot\mathbf{3}+\mathbf{-2}\cdot\mathbf{6}+\mathbf{0}\cdot\mathbf{0} = 16 - 6 - 12 + 0 = -2$$

   Dies resultiert in folgender Lösung:

   $$\begin{pmatrix} \mathbf{-2} & - \\ - & - \\ - & - \end{pmatrix}$$

2. Um den zweiten Wert zu finden, gehen wir in der ersten Matrix eine Zeile weiter und machen dasselbe nochmals mit der auch wieder ersten Spalte der zweiten Matrix:

   $$\mathbf{-2}\cdot\mathbf{4}+\mathbf{-7}\cdot\mathbf{3}+\mathbf{3}\cdot\mathbf{6}+\mathbf{8}\cdot\mathbf{0} = -8 - 21 + 18 + 0 = -11$$

   Wir können unsere Lösungsmatrix also nun erweitern mit dem zweiten Wert:

   $$\begin{pmatrix} \mathbf{-2} & - \\ \mathbf{-11} & - \\ - & - \end{pmatrix}$$

3. Nun machen wir das für die dritte Zeile nochmals gleich. Wir bewegen uns eine Zeile weiter in der ersten Matrix und nehmen wieder die erste Spalte der zweiten Matrix.

4. Um nun den nächsten Wert in der 2. Spalte der Lösungsmatrix zu finden, beginnen wir in der ersten Matrix wieder von oben bei der ersten Zeile. In der zweiten Matrix aber, gehen wir nun eine Spalte weiter:

   $$\mathbf{4}\cdot\mathbf{-5}+\mathbf{-2}\cdot\mathbf{-1}+\mathbf{-2}\cdot\mathbf{4}+\mathbf{0}\cdot\mathbf{-3}$$

   $$= -20 + 2 - 8 + 0 = -26$$

   Dies resultiert in folgender Lösungsmatrix:

   $$\begin{pmatrix} \mathbf{-2} & \mathbf{-26} \\ \mathbf{-11} & - \\ \mathbf{-9} & - \end{pmatrix}$$

5. Nun machen wir das noch für die letzten 2 Einträge genau gleich. Wir bewegen uns wieder in der ersten Matrix durch die Zeilen und bleiben in der zweiten Matrix in der zweiten Spalte. Das ergibt dann folgende Lösung:

   $$\begin{pmatrix} \mathbf{-2} & \mathbf{-26} \\ \mathbf{-11} & \mathbf{5} \\ \mathbf{-9} & \mathbf{-6} \end{pmatrix}$$

3.1.1 Eigenschaften der Matrixmultiplikation

Eine Matrixmultiplikation hat die folgenden Eigenschaften:

- Die Anzahl der Spalten der ersten Matrix A muss gleich der Anzahl Zeilen der zweiten Matrix B sein.
- Das Produkt $C = AB$ hat soviele Zeilen wie A und soviele Spalten wie B.
- Das Element in der i-ten Zeile und der k-ten Spalte von C ist das Skalarprodukt der i-ten Zeile von A und der k-ten Spalte von B.
- Bei der Matrixmultiplikation spielt die Reihenfolge der Faktoren eine Rolle. Im Allgemeinen ist $AB \neq BA$.
- Die Matrixmultiplikation ist assoziativ,

  d.h. $(AB)C = A(BC) = ABC$.

3.2 Wichtige Matrizen

3.2.1 Einheitsmatrix

Eine quadratische Matrix, deren Hauptdiagonale nur aus Einsen besteht und deren andere Einträge verschwinden, d.h. Null sind, nennt man *Einheitsmatrix*. Sie wird mit $\mathbf{E}$ oder $\mathbf{I}$ bezeichnet. Will man besonders auf die Dimensionen n der Matrix hinweisen (also ihre Zeilen und Spalten), so schreibt man $\mathbf{E}_n$ oder $\mathbf{I}_n$.

$$\mathbf{E}_4 = \mathbf{I}_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Multipliziert man eine beliebige Matrix A mit der Einheitsmatrix E, so stimmt das Produkt A überein: $EA = AE = A$. Einheitsmatrizen sind also neutrale Elemente bezüglich der Matrixmultiplikation. Sie verhalten sich wie die Zahl 1 bei der Multiplikation reeller Zahlen.

3.2.2 Transponierte Matrix

Eine Matrix A wird transponiert, indem man sie an ihrer Hauptdiagonalen spiegelt. Dabei werden die Zeilen und Spalten der Matrix vertauscht. Die transponierte Matrix A bezeichnet man mit A^T.

$$\begin{pmatrix} 4 & 5 & 0 \\ -2 & -2 & 8 \\ -1 & 4 & 3 \\ 7 & 0 & -6 \end{pmatrix}^{\mathrm{T}} = \begin{pmatrix} 4 & -2 & -1 & 7 \\ 5 & -2 & 4 & 0 \\ 0 & 8 & 3 & -6 \end{pmatrix}$$

Ein Produkt von zwei Matrizen kann man transponieren, indem man die einzelnen Faktoren transponiert und ihre Reihenfolge vertauscht:

$$(A \cdot B)^T = B^T \cdot A^T$$

3.2.3 Invertierbare oder reguläre Matrix

Ist A eine quadratische Matrix, dann heisst sie *invertierbar, regulär* oder *nichtsingulär*, wenn eine Matrix B existiert, so dass $A \cdot B = B \cdot A = E$. Man nennt B die zu A inverse Matrix und bezeichnet sie mit $B = A^{-1}$.

Nehmen wir an wir reden von einer 2×2 Matrix. Dann finden wir die Inverse wie folgt:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \rightarrow \quad \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Eine Matrix ist nur dann invertierbar, wenn $ad - bc \neq 0$ ist. Das heisst, dass die Determinante der Matrix ungleich Null sein muss.
Allgemein, für eine $n \times n$ Matrix A, ist die Inverse definiert als:

$$a_{ij}^{-1} = \frac{1}{\det(A)} \cdot C_{ji}$$

Wobei C hier die Kofaktormatrix von A ist. Somit können wir die Inverse einer Matrix auch definieren als:

$$A^{-1} = \frac{1}{\det(A)} \cdot \mathrm{Kof}(A)^T$$

Beispiel
In einem konkreten Beispiel für eine 2×2 Matrix A:

$$A = \begin{pmatrix} 1 & 3 \\ 2 & 7 \end{pmatrix}, \quad A^{-1} = \frac{1}{1} \begin{pmatrix} 7 & -3 \\ -2 & 1 \end{pmatrix}$$

3.2.4 Kofaktormatrix

Die Kofaktormatrix kann wie folgt gebildet werden:

$$\text{Kof}(A) = \begin{pmatrix} C_{11} & C_{12} & \cdots & C_{1n} \\ C_{21} & C_{22} & \cdots & C_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ C_{n1} & C_{n2} & \cdots & C_{nn} \end{pmatrix},$$

$$C_{ij} = (-1)^{i+j} \cdot \det(M_{ij})$$

An allen Stellen, an denen $i + j$ ungerade werden, wird der Wert in der Kofaktormatrix negiert.

3.2.5 Permutationsmatrix

Eine Permutationsmatrix ist eine quadratische $n \times n$-Matrix, bei der jede Zeile und jede Spalte genau ein Element 1 enthält, wobei alle anderen Elemente verschwinden (Null sind).

Multipliziert man eine Matrix A von links mit einer Permutationsmatrix (PA), führt dies zum Vertauschen der Zeilen von A. Entsprechend führt die Multiplikation von rechts (AP) zum Vertauschen der Spalten.

Mit anderen Worten: Eine Permutationsmatrix P entsteht aus einer Einheitsmatrix, indem die Zeilen (bzw. Spalten) umgeordnet werden.

Beispiel
Für eine 2×2 Matrix, kann dies so aussehen:

$$A = \begin{pmatrix} 2 & 3 \\ 4 & -5 \end{pmatrix}, A' = \begin{pmatrix} 4 & -5 \\ 2 & 3 \end{pmatrix}, A'' = \begin{pmatrix} 3 & 2 \\ -5 & 4 \end{pmatrix}$$

Die Permutationsmatrix dafür ist folgende:

$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Denn $PA = A'$:

$$A' = P \cdot A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 2 & 3 \\ 4 & -5 \end{pmatrix} = \begin{pmatrix} 4 & -5 \\ 2 & 3 \end{pmatrix}$$

Um die Spalten zu tauschen (anstelle der Zeilen), rechnen wir anstelle $P \cdot A$ ganz einfach $A \cdot P$:

$$A'' = A \cdot P = \begin{pmatrix} 2 & 3 \\ 4 & -5 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 2 \\ -5 & 4 \end{pmatrix}$$

3.3 LR-Zerlegung (LU-Decomposition)

Die LR-Zerlegung ist eine Methode, um eine Matrix in das Produkt zweier einfacherer Matrizen zu zerlegen.

$$A = L \cdot R$$

wobei:

- L: eine untere Dreiecksmatrix (links unten voll, oben Nullen)
- R: eine obere Dreiecksmatrix (rechts oben voll, unten Nullen)

Diese Zerlegung ist nützlich zum Lösen von linearen Gleichungssystemen.

Beispiel
Gegeben sei die Matrix:

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 2 & 3 \\ 3 & 5 & 4 \end{pmatrix}$$

Wir schreiben nun zuerst eine Einheitsmatrix daneben, denn durch das Anwenden der gleichen Operationen wie im Gauss-Verfahren auf die Einheitsmatrix, erhalten wir die Matrix L (Lower):

$$E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 2 & 3 \\ 3 & 5 & 4 \end{pmatrix} \begin{array}{l} \\ \text{II } \mathbf{-2} \cdot \text{I} \\ \text{III } \mathbf{-3} \cdot \text{I} \end{array}$$

Die untenstehende Matrix ist das Ergebnis dieses Schrittes.

$$\begin{pmatrix} 1 & 0 & 0 \\ \mathbf{-2} & 1 & 0 \\ \mathbf{-3} & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 1 \\ 2 & 2 & 3 \\ 3 & 5 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 1 \\ 0 & -2 & 1 \\ 0 & -1 & 1 \end{pmatrix}$$

Wir müssen noch das letzte Feld (dritte Zeile, zweite Spalte) auf 0 bringen. Dazu subtrahieren wir die Hälfte der zweiten Zeile von der dritten Zeile. Hier wichtig: Wir erstellen hier eine zweite L-Matrix, die aktuell eine neue Einheitsmatrix ist. Die Werte aus dem letzten Schritt sind hier nicht vorhanden.

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 1 \\ 2 & 2 & 3 \\ 3 & 5 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 1 \\ 0 & -2 & 1 \\ 0 & -1 & 1 \end{pmatrix} \begin{array}{l} \\ \\ \text{III } \mathbf{-0.5} \cdot \text{II} \end{array}$$

Das Ergebnis sieht dann wie folgt aus:

$$\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & \mathbf{-0.5} & 1\end{pmatrix} \cdot \begin{pmatrix}1 & 2 & 1\\ 2 & 2 & 3\\ 3 & 5 & 4\end{pmatrix} = \begin{pmatrix}1 & 2 & 1\\ 0 & -2 & 1\\ 0 & 0 & 0.5\end{pmatrix}$$

R kennen wir nun. Jetzt müssen wir nur noch L fertig bestimmen.

$$R = \begin{pmatrix}1 & 2 & 1\\ 0 & -2 & 1\\ 0 & 0 & 0.5\end{pmatrix}$$

Wir haben jetzt zwei Matrizen L_1 und L_2, die wir zusammenführen müssen, um L zu erhalten. Wir wissen, dass $L_2 \cdot L_1 \cdot A = R$. Dies können wir nach A umstellen:

$$A = \underbrace{L_1^{-1} \cdot L_2^{-1}}_{L} \cdot R$$

Wir müssen also die Inversen bilden. Üblicherweise ist dies recht aufwändig, hier aber, da es sich um eine unitäre untere Dreiecksmatrix handelt, ist es einfach:

Wir können von jedem Wert unterhalb der Hauptdiagonalen in L den Kehrwert nehmen, um die Inverse zu bilden. Die Werte auf der Hauptdiagonalen bleiben gleich.

$$L_1^{-1} = \begin{pmatrix}1 & 0 & 0\\ -2 & 1 & 0\\ -3 & 0 & 1\end{pmatrix}^{-1} = \begin{pmatrix}1 & 0 & 0\\ 2 & 1 & 0\\ 3 & 0 & 1\end{pmatrix}$$

$$L_2^{-1} = \begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & -0.5 & 1\end{pmatrix}^{-1} = \begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0.5 & 1\end{pmatrix}$$

Somit kommen wir auf ein L, das wie folgt aussieht:

$$L = \begin{pmatrix}1 & 0 & 0\\ 2 & 1 & 0\\ 3 & 0.5 & 1\end{pmatrix}$$

3.3.1 LR-Zerlegung mit Zeilentausch

Wenn man ein Gleichungssystem mit der LR-Zerlegung lösen will, benutzt man das Gauss-Verfahren, um eine Matrix schrittweise in eine obere Dreiecksmatrix (R) zu bringen. Dabei arbeitet man sich Zeile für Zeile vor und benutzt Zahlen auf der Diagonale, um darunterliegende Einträge zu nullen.

Wenn jetzt aber ein Wert auf der Diagonale schon 0 ist, kann man nicht teilen, was das Verfahren abbricht.

$$\begin{pmatrix}0 & 2\\ 1 & 3\end{pmatrix}$$

Beim ersten Schritt wollen wir die 0 links oben (die Pivot-Position) benutzen, aber:

- Man darf nich durch 0 teilen!
- Die Rechnung wäre undefiniert.

Die einfache Lösung: Man tauscht die Zeilen.

Wie wir vorher schon gesehen haben, kann man dies beispielweise mit einer Permutationsmatrix machen:

$$A = \begin{pmatrix}0 & 1 & 1\\ 1 & 2 & 1\\ 2 & 7 & 9\end{pmatrix}$$

Dann können wir mit folgender Permutationsmatrix die Zeilen tauschen:

$$P = \begin{pmatrix}0 & 1 & 0\\ 1 & 0 & 0\\ 0 & 0 & 1\end{pmatrix}$$

Was uns folgendes Resultat gibt:

$$P \cdot A = \begin{pmatrix}0 & 1 & 0\\ 1 & 0 & 0\\ 0 & 0 & 1\end{pmatrix} \begin{pmatrix}0 & 1 & 1\\ 1 & 2 & 1\\ 2 & 7 & 9\end{pmatrix} = \begin{pmatrix}1 & 2 & 1\\ 0 & 1 & 1\\ 2 & 7 & 9\end{pmatrix}$$

Auf die resultierende Matrix können wir wieder ganz normal das Gauss-Verfahren anwenden.

3.4 Lösen eines LGS mittels LR-Zerlegung

Wenn wir ein Gleichungssystem mit der LR-Zerlegung lösen wollen, stellen wir das System wie folgt auf:

$$\mathbf{Ax} = \mathbf{b}$$
$$\mathbf{A} = \mathbf{L} \cdot \mathbf{R}$$
$$\mathbf{L} \cdot \mathbf{R} \cdot \mathbf{x} = \mathbf{b}$$

Nun können wir folgendes definieren:

$$\mathbf{Ly} = \mathbf{b}$$
$$\mathbf{Rx} = \mathbf{y}$$

Die allgemeine Formel für das Vorwärtseinsetzen ist die folgende:

$$y_i = \frac{1}{l_{ii}} \left(b_i - \sum_{k=1}^{i-1} l_{ik} y_k \right)$$

Die allgemeine Formel für das Rückwärtseinsetzen ist die folgende:

$$x_i = \frac{1}{r_{ii}} \left(y_i - \sum_{k=i+1}^{n} r_{ik} x_k \right)$$

Beispiel
Gegeben ist das folgende Gleichungssystem:

$$u + 2v = 5$$
$$4u + 9v = 21$$

Das Gleichungssystem lässt sich in Matrixform schreiben als:

$$A \cdot x = b, \quad A = \begin{pmatrix} 1 & 2 \\ 4 & 9 \end{pmatrix}, \; x = \begin{pmatrix} u \\ v \end{pmatrix}, \; b = \begin{pmatrix} 5 \\ 21 \end{pmatrix}$$

Nun suchen wir wie im letzten Kapitel erklärt die Matrizen L und R (engl. U), sodass:

$$A = L \cdot R$$

Wenn wir dies tun, erhalten wir:

$$L = \begin{pmatrix} 1 & 0 \\ 4 & 1 \end{pmatrix}, \; R = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$$

Und somit die Gleichung:

$$\begin{pmatrix} 1 & 2 \\ 4 & 9 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 4 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$$

Jetzt beginnen wir mit dem Vorgehen zur Lösung der Gleichungen:

1. Wir ersetzen in $Ax = b$ die Matrix A durch $L \cdot R$:
$$L \cdot R \cdot x = b$$
2. Wir können nun $y : R \cdot x$ definieren. Daraus ergibt sich:
   (a) Lösen von: $L \cdot y = b \rightarrow$ Vorwärtseinsetzen
   (b) Lösen von: $R \cdot x = y \rightarrow$ Rückwärtseinsetzen
3. Vorwärtseinsetzen: Wir wollen hier das Gleichungssystem $Ly = b$ lösen. Dazu setzen wir die Matrix L und den Vektor b ein:
$$L = \begin{pmatrix} 1 & 0 \\ 4 & 1 \end{pmatrix}, \; b = \begin{pmatrix} 5 \\ 1 \end{pmatrix}$$
$$\begin{array}{c} y_1 = 5 \\ 4y_1 + y_2 = 21 \Rightarrow y_2 = 21 - 4 \cdot 5 = 1 \end{array}, \; y = \begin{pmatrix} 5 \\ 1 \end{pmatrix}$$
4. Rückwärtseinsetzen: Wir lösen jetzt das Gleichungssystem $Rx = y$ mit der Matrix R und dem zuvor berechneten Vektor y:
$$R = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, \; y = \begin{pmatrix} 5 \\ 1 \end{pmatrix}$$
$$\begin{array}{c} x_2 = 1 \\ x_1 + 2x_2 = 5 \Rightarrow x_1 = 5 - 2 \cdot 1 = 3 \end{array}, \; x = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$$
5. Da wir nun x kennen, kennen wir also nun auch u und v, denn:
$$x = \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$$
   Also ist das Ergebnis: $u = 3$ und $v = 1$.

3.4.1 Lösung eines LGS mit Zeilentausch

Wenn wir in einer Matrix auf einer der Pivot-Positionen eine 0 haben, können wir die Zeilen tauschen, um das Problem zu lösen. Dies geschieht mit einer Permutationsmatrix, wie wir es bereits gesehen haben. Danach ist das Verfahren dasselbe wie oben beschrieben.

Wichtig: Auch der Vektor b muss mit der Permutationsmatrix multipliziert werden, um die Zeilen zu tauschen.

4 Lineare Gleichungssysteme : Gauss-Jordan-Algorithmus und Determinante

4.1 Gauss-Jordan-Algorithmus zur Lösung linearer Gleichungssysteme

Das Gauss-Jordan-Verfahren ist eine Erweiterung des Gauss-Verfahrens, das es ermöglicht, ein lineares Gleichungssystem in eine reduzierte Zeilenstufenform zu bringen. Dies erleichtert die Bestimmung der Lösungen des Systems.

Es eliminiert weitere Elemente der Koeffizientenmatrix, so dass die Koeffizientenmatrix schliesslich eine Diagonalmatrix ist.

Eine weitere Anwendung vom Gauss-Jordan-Algorithmus ist die Berechnung der Inversen einer quadratischen Matrix.

4.1.1 Der Algorithmus

Der Gauss-Jordan-Algorithmus kann bei gegebener Koeffizientenmatrix A simultan für mehrere rechte Seiten $(a, b, c, ..)$ durchgeführt werden. Das heisst, dass die Lösungen $L_a, L_b, L_c, ..$ gleichzeitig zu folgenden Gleichungen führen:

$$A \cdot x = a \qquad A \cdot x = b \qquad A \cdot x = c$$

Ausgangssituation
Nachdem man ein lineares Gleichungssystem auf Zeilenstufenform gebracht hat, kann man es noch weiter mit HIlfe elementarer Zeilenumformungen vereinfachen. Aus der vereinfachten Form lässt sich dann die Lösung direkt, d. h. ohne Rückwärtsrechnen ablesen.

Verfahren
Ist die erweiterte Koeffizientenmatrix des linearen Gleichungssystems bereits in Zeilenstufenform, ist unser Ziel, in den Spalten oberhalb der Diagonalen Nullen zu erzeugen:

- Während wir beim Gauss-Verfahren von oben nach unten und von links nach rechts arbeiten, arbeiten wir beim Gauss-Jordan-Verfahren von rechts nach links und von unten nach oben.
- Es werden keine Zeilen mehr vertauscht. Der aktuelle Leitkoeffizient ist das Pivotelement.
- Alle Pivotelemente durch Multiplikation der entsprechenden Zeile mit dem Kehrwert auf den Wert eins gebracht.

Beispiel
Nehmen wir an, wir wollen folgendes Gleichungssystem mittels des Gauss-Jordan-Verfahren lösen:

$$\begin{aligned} a + b + c &= 0 \\ 4a + 2b + c &= 1 \\ 9a + 3b + c &= 3 \end{aligned}$$

Schreiben wir es zuerst als erweiterte Koeffizientenmatrix:

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 4 & 2 & 1 & 1 \\ 9 & 3 & 1 & 3 \end{array}\right)$$

Wir überspringen hier das Gauss-Verfahren, dass wir zuerst auf diese Matrix anwenden müssen. Nach dem Gauss-Verfahren, sieht die Matrix wie folgt aus:

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & -2 & -3 & 1 \\ 0 & 0 & 1 & 0 \end{array}\right)$$

Nun gehen wir genau in umgekehrter Reihenfolge vor, um das Gauss-Jordan-Verfahren abzuschliessen. Wir gehen von rechts nach links und von unten nach oben.

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & -2 & -3 & 1 \\ 0 & 0 & 1 & 0 \end{array}\right) \begin{array}{l} \mathrm{I} - (1) \cdot \mathrm{III} \\ \mathrm{II} - (-3) \cdot \mathrm{III} \\ \end{array}$$

Nun haben wir die zwei Nullen in der rechten Spalte in der oberen Hälfte erzeugt. Jetzt müssen wir in der zweiten Spalte auf der Hauptdiagonale das -2 zu einer 1 machen. Dies machen wir wie folgt:

$$\left(\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 0 & -2 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{array}\right) \quad \mathrm{II} \,/ -2$$

Dies resultiert dann in die folgende Matrix. Bei dieser müssen wir jetzt nur noch den mittleren Wert in der ersten Spalte auf 0 bringen. Dazu wenden wir wieder das gleiche an, wie vorher und im Gauss-Verfahren auch schon (nur hier verkehrt):

$$\left(\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & -0.5 \\ 0 & 0 & 1 & 0 \end{array}\right) \quad \mathrm{I} - \mathrm{II}$$

Dies gibt uns dann folgendes Resultat:

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & 0.5 \\ 0 & 1 & 0 & -0.5 \\ 0 & 0 & 1 & 0 \end{array}\right)$$

Aus dieser können wir die Werte für die Variabeln ablesen.

$$a = 0.5, \quad b = -0.5, \quad c = 0$$

Somit haben wir das Gleichungssystem gelöst.

4.2 Inverse Matrix und Gauss-Jordan-Algorithmus

Um die Inverse mittels des Gauss-Jordan-Verfahrens zu finden, nutzen wir folgende Regel:

$$(A|E) = (E|A^{-1})$$

Hier Beispiele einer 2×2 und einer 3×3 Matrix, welche beide bereit sind um das Gauss-Jordan-Verfahren anzuwenden:

$$\left(\begin{array}{cc|cc} 1 & 3 & 1 & 0 \\ 2 & 7 & 0 & 1 \end{array}\right), \left(\begin{array}{ccc|ccc} 2 & 1 & -1 & 1 & 0 & 0 \\ -3 & -1 & 2 & 0 & 1 & 0 \\ -2 & 1 & 2 & 0 & 0 & 1 \end{array}\right)$$

Wenn wir auf diese erweiterten Koeffizientmatrizen das Gauss-Jordan-Verfahren anwenden würden, kehren wir die Matrizen um.

Wichtig: Eine Matrix mit $\det(A) = 0$ hat keine Inverse. Auch Matrizen mit Nullzeilen haben keine Inverse. Eine Matrix muss quadratisch sein um invertierbar zu sein.

Das Ziel ist es, die Matzrizen so umzuformen, dass die Einheitsmatrix links steht. Dadurch entsteht dann auf der rechten Seite die Inverse. Bei den obigen Beispielen, wären dies die Resultate:

Die 2×2 Inverse:

$$\left(\begin{array}{cc|cc} 1 & 0 & 7 & -3 \\ 0 & 1 & -2 & 1 \end{array}\right), \quad \Rightarrow \begin{pmatrix} 7 & -3 \\ -2 & 1 \end{pmatrix}$$

Die 3×3 Inverse:

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 4 & 3 & -1 \\ 0 & 1 & 0 & -2 & -2 & 1 \\ 0 & 0 & 1 & 5 & 4 & -1 \end{array}\right), \quad \Rightarrow \begin{pmatrix} 4 & 3 & -1 \\ -2 & -2 & 1 \\ 5 & 4 & -1 \end{pmatrix}$$

4.3 Cramersche Regel

Die Cramersche Regel ist ein Verfahren zur Lösung von linearen Gleichungssystemen mit genau einer Lösung. Sie gilt für quadratische Systeme der Form $A \cdot x = b$, wenn $\det(A) \neq 0$

4.3.1 Cramersche Regel für 2x2-Systeme

Gegeb en ist ein Gleichungssystem:

$$\begin{aligned} a_1x + b_1y = c_1 \\ a_2x + b_2y = c_2 \end{aligned}$$

Die Lösung ergibt sich über die Determinanten:

$$D = \det(A) = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = a_1b_2 - a_2b_1$$

$$D_x = \begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix}, \quad D_y = \begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}$$

Lösungen:

$$x = \frac{\det(c, b)}{\det(a, b)} = \frac{D_x}{D}, \quad y = \frac{\det(a, c)}{\det(a, b)} = \frac{D_y}{D}$$

Beispiel für 2x2 System
Gegeben sei ein 2×2 Gleichungssystem:

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \quad b = \begin{pmatrix} 5 \\ 6 \end{pmatrix}$$

Nun berechnen wir die Determinante von A:

$$\det(A) = 1 \cdot 4 - 3 \cdot 2 = 4 - 6 = -2$$

Dies machen wir, weil diese Determinante der Nenner für beide Variablen x und y ist. Nur wenn $\det(A) \neq 0$, gibt es eine eindeutige Lösung.
Jetzt ersetzen wir die 1. Spalte von A durch den Vektor b:

$$A_x = \begin{pmatrix} 5 & 2 \\ 6 & 4 \end{pmatrix} \;\Rightarrow\; D_x = \det(A_x) = 5 \cdot 4 - 6 \cdot 2 = 8$$

Jetzt ersetzen wir die 2. Spalte von A durch b:

$$A_y = \begin{pmatrix} 1 & 5 \\ 3 & 6 \end{pmatrix} \;\Rightarrow\; D_y = \det(A_y) = 1 \cdot 6 - 3 \cdot 5 = -9$$

Und jetzt können wir die Lösung für x und y berechnen:

$$x = \frac{D_x}{\det(A)} = \frac{8}{-2} = -4$$

$$y = \frac{D_y}{\det(A)} = \frac{-9}{-2} = 4.5$$

4.3.2 Cramersche Regel für 3x3-Systeme

Gegeben ist:

$$a_1 x + b_1 y + c_1 z = d_1$$
$$a_2 x + b_2 y + c_2 z = d_2$$
$$a_3 x + b_3 y + c_3 z = d_3$$

Die Hauptdeterminante ist definiert als:

$$D = \det(A) = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

$$= a_1 b_2 c_3 + b_1 c_2 a_3 + c_1 a_2 b_3 - a_3 b_2 c_1 - b_3 c_2 a_1 - c_3 a_2 b_1$$

Bei den Ersetzungsdeterminanten wird, wie auch bei dem 2×2-System, Spaltenweise eine Spalte mit dem Lösungsvektor ersetzt.

Die Lösungen werden wie folgt ermittelt:

$$x = \frac{D_x}{D}, \quad y = \frac{D_y}{D}, \quad z = \frac{D_z}{D}$$

Beispiel für 3x3 System
Gegeben sei:

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 3 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

Berechnen wir die Determinante:

$$\det(A) = 2 \cdot 2 \cdot 2 + 1 \cdot 1 \cdot 1 + 1 \cdot 3 \cdot 1 - 1 \cdot 2 \cdot 1 - 1 \cdot 1 \cdot 2 - 2 \cdot 3 \cdot 1$$

$$= 8 + 1 + 3 - 2 - 2 - 6 = 12 - 10 = 2$$

Jetzt ersetzen wir die erste Spalte durch b und berechnen die Determinante dieser Matrix:

$$A_x = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 1 \\ 3 & 1 & 2 \end{pmatrix}, \quad \det(A_x) = -2.0$$

Nun machen wir dasselbe für Spalte 2 und berechnen die Determinante dieser Matrix:

$$A_y = \begin{pmatrix} 2 & 1 & 1 \\ 3 & 2 & 1 \\ 1 & 3 & 2 \end{pmatrix}, \quad \det(A_y) = 0.0$$

Und noch für die letzte Spalte und berechnen die Determinante dieser Matrix:

$$A_z = \begin{pmatrix} 2 & 1 & 1 \\ 3 & 2 & 2 \\ 1 & 1 & 3 \end{pmatrix}, \quad \det(A_z) = 0.0$$

Jetzt kennen wir alle Ersetzungsdeterminanten und auch die Hauptdeterminante und können die Ergebnisse für die Variablen berechnen:

$$D = \det(A) = 1.0,, \; D_x = -2.0, \; D_y = 0.0, \; D_z = 0.0$$

Laut Cramerscher Regel gilt für das 3×3 System:

$$x = \frac{D_x}{D}, \quad y = \frac{D_y}{D}, \quad z = \frac{D_z}{D}$$

Wenn wir hier einsetzen erhalten wir folgende Werte für x, y und z:

$$x = \frac{-2.0}{1.0} = -2.0, \quad y = \frac{0}{1.0} = 0, \quad z = \frac{0}{1.0} = 0$$

4.4 Determinante

Jede quadratische Matrix ist entweder regulär (invertierbar) oder singulär (nicht invertierbar). Ob eine Matrix regulär ist, erkennt man an ihrer Determinante:

- Ist die Determinante ungleich null, ist die Matrix regulär.
- Ist sie null, ist die Matrix singulär.

Die Determinante spielt eine wichtige Rolle bei der Frage, ob ein lineares Gleichungssystem eine eindeutige Lösung besitzt.

Definition: In der linearen Algebra ist die Determinante eine Zahl, die zu jeder quadratischen Matrix gehört. Man berechnet sie aus den Einträgen der Matrix, und sie hilft z.B. dabei zu entscheiden, ob ein lineares Gleichungssystem lösbar ist oder nicht.
Gängige Schreibweisen für die Determinante der Matrix A sind:

$$\det(A), \quad \det A, \quad |A|$$

Wie wir auch schon in den Kapiteln weiter oben gesehen habe, kann die Determinante genutzt werden, um die Inverse einer Matrix zu berechnen:

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Dies gilt hier jetzt aber nur für 2×2-Matrizen. Im allgemeinen gilt:

$$A^{-1} = \frac{1}{ad - bc} \cdot \mathrm{Kof}(A)^T = \frac{1}{ad - bc} \cdot \mathrm{adj}(A)$$

4.4.1 Determinante einer 2x2-Matrix

Gegeben ist:

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

Die Zahl $ad - bc$ nennt man die Determinante der Matrix und schreibt $\det(A)$. Die Zahl ist die Fläche, die durch die Vektoren a und b aufgespannt ist.

4.4.2 Determinante einer 3x3-Matrix

Drei Vektoren spannen in einem Raum ein Parallelepiped auf. Das Volumen dieses Körpers lässt sich mit der Determinante der Matrix berechnen, deren Spalten diese drei Vektoren sind. In diesem Zusammenhang nennt man die Determinante auch das Spatprodukt.

Die Determinante in einer 3×3-Matrix wird wie folgt berechnet:

$$A = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$$

$$D_A = a_1 b_2 c_3 + b_1 c_2 a_3 + c_1 a_2 b_3 - c_1 b_2 a_3 - a_1 c_2 b_3 - b_1 a_2 c_3$$

4.4.3 Eigenschaften von Determinanten

1. Determinante von Matrizen in Dreiecksgestalt sind ein Produklt der Diagonalelementen.
2. Vertauschen von zwei Zeilen ändert das Vorzeichen der Determinante.
3. Multiplikation einer Zeile mit einem Faktor multipliziert die Determinante mit diesem Faktor.
4. Durch die Addition des Vielfachen einer Zeile zu einer anderen Zeile ändert sich die Determinante nicht.
5. Beim Transponieren der Matrizen ändert sich die Determinante nicht.

Zusätzliche Information: Eine Einheitsmatrix hat die Determinante = 1.

4.4.4 Äquivalente Aussagen für Determinanten

Sei A eine quadratische $n \times n$-Matrix.

- A ist regulär.
- A^T ist regulär.
- Die Determinante von A ist nicht null.
- Der Nullfaktor ist die einzige Lösung von $Ax = 0$.
- Der Rang von A ist gleich n.

5 Vektorprodukt, Lineare Abbildungen

5.1 Vektorprodukt

Das Vektorprodukt (auch Kreuzprodukt genannt) ist eine Rechenoperation, die auf zwei Vektoren im dreidimensionalen Raum angewendet wird. Das Ergebnis ist ein neuer Vektor, der:

- senkrecht auf beiden Eingangsvektoren steht,
- und eine Länge hat, die dem Flächeninhalt des aufgespannten Parallelograms entspricht.

Wichtige Eigenschaften sind folgende:

- Antikommutativität:

$$\vec{a} \times \vec{b} = -(\vec{b} \times \vec{a})$$

$$a \times a = -a \times a \Rightarrow a \times a = 0$$

- Bilinearität: Distributiv und kompatibel mit Skalarmultiplikation:

$$(a + b) \times c = a \times c + b \times c$$

$$(\lambda \cdot a) \times b = \lambda \cdot (a \times b)$$

$$a \times (b + c) = a \times b + a \times c$$

$$a \times (\lambda \cdot b) = \lambda \cdot (a \times b)$$

- Beziehung zur Determinante:

$$(\vec{a} \times \vec{b} \cdot \vec{c} = \det(\vec{a}, \vec{b}, \vec{c}))$$

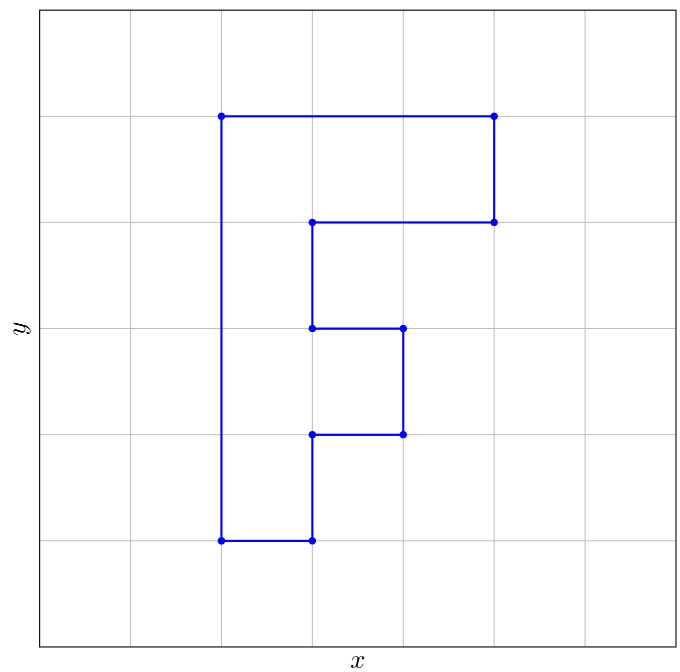
- Rechenformel:

$$\vec{a} \times \vec{b} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix}$$

5.2 Lineare Abbildungen

Wir können den Buchstaben F betrachten. Den Buchstaben F können wir mit den Koordinaten seiner Eckpunkte der Umrisslinie zeichnen lassen. In einer Matrix, sieht dies wie folgt aus:

$$P = \begin{pmatrix} 1 & 2 & 2 & 3 & 3 & 2 & 2 & 4 & 4 & 1 \\ 1 & 1 & 2 & 2 & 3 & 3 & 4 & 4 & 5 & 5 \end{pmatrix}$$



Im n-dimensionalem Raum R^n lassen sich n verschiedene Einheitsvektoren e_k mit folgenden Eigenschaften definieren:

- Alle Koordinaten eines Einheitsvektors mit Ausnahme einer Koordinate sind 0. Diese Koordinate hat den Wert 1.
- Alle Vektoren e_k haben die Länge 1.
- Die Vektoren e_k stehen paarweise aufeinander senkrecht.
- Fasst man die Einheitsvektoren $e_1, e_2, ..., e_n$ spaltenweise zusammen, bilden sie die Einheitsmatrix E_n bzw. I_n.
- Die Einheitsvektoren bilden eine Basis des R^n, d. h. jeder Vektor aus dem R^n lässt sich als Linearbkombination der Einheitsvektoren darstellen.

Beispiel

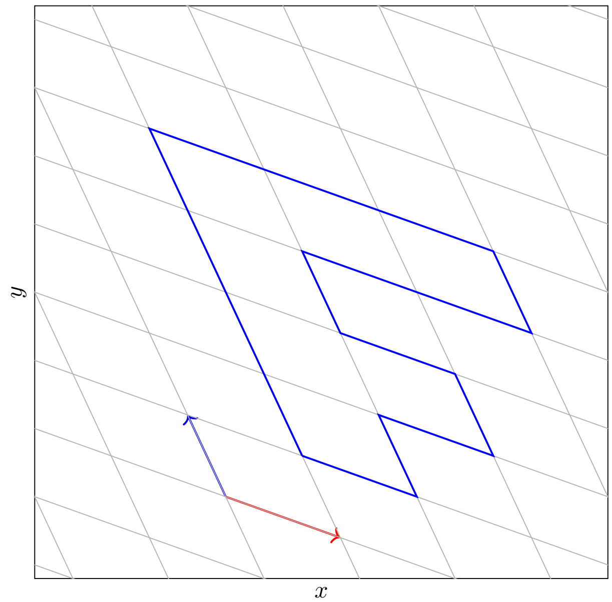
Der vierte Vektor in der Umrisslinie des Buchstabens F lässt sich auf zwei unterschiedliche Arten schreiben:

$$v = \begin{pmatrix} 3 \\ 2 \end{pmatrix} = 3 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 3 \cdot e_1 + 2 \cdot e_2$$

Wählt man an Stelle von e_1 (rot) und e_2 (blau) im linken Bild andere Basisvektoren $f_1 = (3, 1)^T$ (rot) und $f_2 = (-1, 2)^T$ (blau) im rechten Bild, so erhält man:

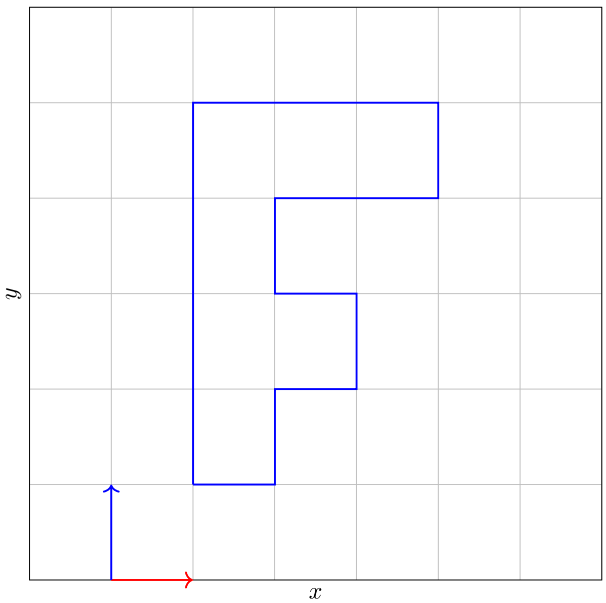
$$w = 3 \cdot f_1 + 2 \cdot f_2 = 3 \cdot \begin{pmatrix} 3 \\ -1 \end{pmatrix} + 2 \cdot \begin{pmatrix} -1 \\ 2 \end{pmatrix}$$

$$= \begin{pmatrix} 9 \\ -3 \end{pmatrix} + \begin{pmatrix} -2 \\ 4 \end{pmatrix} = \begin{pmatrix} 7 \\ 1 \end{pmatrix}$$



Dasselbe erhält man, wenn mann eine Matrix, welche die neuen Basisvektoren als Spalten hat, mit dem Vektor v multipliziert:

$$A \cdot v = \begin{pmatrix} 3 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 7 \\ 1 \end{pmatrix}$$



Die entsprechende Rechnung kann für alle Punkte der Umrisslinie des Buchstabens F wiederholt werden. Dabei nutzt man mit Vorteil aus, dass die Punkte bereits in einer Matrix zusammengestellt sind:

$$\begin{aligned} Q &= A \cdot P \\ &= \begin{pmatrix} 3 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 2 & 3 & 3 & 2 & 2 & 4 & 4 & 1 \\ 1 & 1 & 2 & 2 & 3 & 3 & 4 & 4 & 5 & 5 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 5 & 4 & 7 & 6 & 3 & 2 & 8 & 7 & -2 \\ 1 & 0 & 2 & 1 & 3 & 4 & 6 & 4 & 6 & 9 \end{pmatrix} \end{aligned}$$

- Die Vektoren der Basis (e_1, e_2) spannen ein Gitter auf, in welchem die Umrisslinien des Buchstabens F dargestellt werden.
- Die Vektoren der Basis (f_1, f_2) spannen ein verformtes Gitter auf. Stellt man die Umrisslinien in diesem neuen Gitter dar, erscheint der Buchstabe F ebenfalls verformt.
- Wenn man weiss, wie die Basis (e_1, e_2) in die Basis (f_1, f_2) transofrmiert wird, kennt man auch die Transformation des gesamten Buchstabens F.
- Abbildungen der eben beschriebenen Art haben spezielle Eigenschaften und heissen lineare Abbildungen.

Die Definition für Lineare Abbildungen sieht wie folgt aus:

Eine Abbildung A eines Vektorraums D in einen anderen Vektorraum W heisst linear, wenn sie mit der Vektoraddition und der skalaren Multiplikation verträglich ist, d. h.

$$A(u+v) = A(u) + A(v) \qquad \forall u, v \in D$$
$$A(s \cdot u) = s \cdot A(u) \qquad \forall s\Re, \forall v \in D$$

Die Multiplikation einer Matrix mit einem Vektor erfüllt diese Eigenschaften und stellt daher eine lineare Abbildung dar.

5.3 Abbildungen transformieren

Um eine Abbildung zu transformieren, können wir den Vektor mit einer Matrix mulziplizieren, also:

Jede dieser Transformationen wirkt durch Matrixmultiplikation auf einen Vektor $\vec{v}$:

$$\vec{v} = A \cdot \vec{v}$$

Kurz gesagt: Die Matrix bestimmt, wie der Raum verzerrt wird.

5.3.1 Abbildung skalieren

Um eine lineare Abbildung zu skalieren, können wir sie mit der folgender Matrix multiplizieren:

$$\begin{pmatrix} s & 0 \\ 0 & s \end{pmatrix}$$

(z.B. $s = 2$ verdoppelt die Länge in beide Richtungen). Wenn man für das untere s (unten-rechts) 2 nehmen würde, und für das obere 1, würde man die Abbildung nur nach oben strecken.

5.3.2 Abbildung drehen

Um eine Abbildung zu drehen, können wir sie mit einer Rotationsmatrix multiplizieren:

$$\begin{pmatrix} cos\phi & -sin\phi \\ sin\phi & cos\phi \end{pmatrix}$$

Dabei beschreibt ϕ, den Winkel der Drehung.

5.3.3 Abbildung an x-Achse spiegeln

Um eine Abbildung an der x-Ache zu spiegeln, können wir sie mit folgender Matrix multiplizieren:

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

5.3.4 Abbildung an y-Achse spiegeln

Um eine Abbildung an der y-Achse zu spiegeln, können wir sie mit folgender Matrix multiplizieren:

$$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

5.3.5 Abbildung an beliebiger Achse spiegeln

Um eine Abbildung auf einer beliebigen Achse zu spiegeln, drehen wir sie auf der x-Achse, spiegeln sie, und drehen sie zurück:

$$\begin{pmatrix} \cos(2\phi) & \sin(2\phi) \\ \sin(2\phi) & -\cos(2\phi) \end{pmatrix}$$

Dabei beschreibt ϕ, den Winkel der Achse.

5.3.6 Scherung von Abbildungen

Die Scherung (kippen) von Abbildungen, können wir mit folgender Matrix machen:
Um eine horizontale Scherung vorzunehmen:

$$\begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$$

Dabei beschreibt k, wie stark die Scherung ist.

- $k \geq 0$: kippt nach rechts oben
- $k \leq 0$: kippt nach links oben
- $k = 0$: keine Scherung

Um eine vertikale Scherung vorzunehmen:

$$\begin{pmatrix} 1 & 0 \\ k & 1 \end{pmatrix}$$

- $k \geq 0$: kibbt nach oben rechts
- $k \leq 0$: kippt nach unten rechts
- $k = 0$: keine Scherung

5.4 Zusammensetzungen von Abbildungen

Wenn mehrere Transformationen hintereinander ausgeführt werden sollten, z.B.:

- zuerst Matrix A: ergibt $y = A \cdot x$
- dann Matrix B: ergibt $z = B \cdot y$

Dann kannst du beide Schritte in einer Matrix zusammenfassen:

$$z = B(Ax) = (BA)x = Cx$$

Wichtig! Die Reihenfolge der Matrizen ist umgekehrt zur Reihenfolge der Anwendung:

zuerst kommt A, dann B, also: $C = B \cdot A$

Die Matrixmultiplikation ergibt automatisch die neue zusammengesetzte Transformation.

Spiegelung an einer beliebigen Achse
Vorgehen:

1. Drehung um $-\phi$: Dies bringt die Spiegelachse auf die x-Achse.
2. Spiegelung an der x-Achse mit:

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

3. Drehung zurück um $+\phi$: Dies bringt die Achse zurück in ihre ursprüngliche Lage.

Die Gesamtmatrix ergibt sich durch Multiplikation:

$$R(\phi) \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cdot R(-\phi) = \begin{pmatrix} \cos(2\phi) & \sin(2\phi) \\ \sin(2\phi) & -\cos(2\phi) \end{pmatrix}$$

Warum 2ϕ? Weil sich durch die Drehung hin und zurück eine Verdopplung des Winkels in der Spiegelmatrix ergibt - mathematisch durch die Matrixmultiplikation.

Also kurz und einfach:

- Mehrere Abbildungen können durch eine Matrixmultiplikation kombiniert werden (in umgekehrter Reihenfolge).
- Spiegelungen an geneigter Achse: Drehung, Spiegelung an x-Achse, Rückdrehung

5.4.1 Schritt-für-Schritt-Anleitung

Schritt 1: Unbekannte Transformationsmatrix definieren
Wir setzen die gesuchte Matrix T mit den unbekannten Variablen a, b, c, d an.

$$T = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Schritt 2: Zwei linear unabhängige Vektor-Paare auswählen
Wähle zwei Start-Vektoren, die nicht auf derselben Linie liegen (linear unabhängig sind), sowie deren zugehörige Ziel-Vektoren.

Start-Vektor 1: $v_1 = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$ $\rightarrow$ Ziel-Vektor 1: $w_1 = \begin{pmatrix} x'_1 \\ y'_1 \end{pmatrix}$

Start-Vektor 2: $v_2 = \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$ $\rightarrow$ Ziel-Vektor 2: $w_2 = \begin{pmatrix} x'_2 \\ y'_2 \end{pmatrix}$

Schritt 3: Zwei Matrixgleichungen aufstellen
Für jedes Vektor-Paar gilt die Grundgleichung $T \cdot v = w$. Daraus ergeben sich zwei separate Gleichungen:

1. $T \cdot v_1 = w_1 \implies \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} x'_1 \\ y'_1 \end{pmatrix}$
2. $T \cdot v_2 = w_2 \implies \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} x'_2 \\ y'_2 \end{pmatrix}$

Schritt 4: Lineares Gleichungssystem ableiten
Multipliziere die beiden Matrixgleichungen aus. Jede Matrixgleichung liefert zwei lineare Gleichungen.

- Aus Gleichung 1:

$$\begin{aligned} a \cdot x_1 + b \cdot y_1 &= x'_1 && (I) \\ c \cdot x_1 + d \cdot y_1 &= y'_1 && (II) \end{aligned}$$

- Aus Gleichung 2:

$$\begin{aligned} a \cdot x_2 + b \cdot y_2 &= x'_2 && (III) \\ c \cdot x_2 + d \cdot y_2 &= y'_2 && (IV) \end{aligned}$$

Du hast nun ein System aus vier Gleichungen mit vier Unbekannten (a, b, c, d).

Schritt 5: Gleichungssystem lösen

Löse das System. Beachte, dass die Gleichungen für a, b von denen für c, d entkoppelt sind. Du kannst also zwei kleinere 2x2-Systeme lösen:

- Löse das System (I) und (III) für die Variablen a und b.
- Löse das System (II) und (IV) für die Variablen c und d.

Schritt 6: Transformationsmatrix T zusammensetzen

Setze die gefundenen Werte für a, b, c und d in die Matrix aus Schritt 1 ein, um die finale Transformationsmatrix T zu erhalten.

$$T = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Dieser Prozess ist universell und funktioniert für jede lineare 2D-Transformation, solange die gewählten Start-Vektoren linear unabhängig sind.

6 Orthogonale Abbildungen und homegene Koordinaten

6.1 Orthogonale Abbildungen

Orthogonale Abbildungen sind Transformationen im Raum (z. B. Drehungen oder Spiegelungen), die:

- bei denen die Transponierte Matrix multipliziert mit der ursprünglichen Matrix die Einheitsmatrix ergibt:§

$$A^T \cdot A = I \qquad \text{I = Einheitsmatrix}$$

- Längen und Winkel erhalten (also keine Verzerrung erzeugen),

- das Skalarprodukt erhalten, d.h. die "innere Beziehungen" zwischen Vektoren bleiben gleich,

- Determinante ± 1 haben:

  - $+1 \rightarrow$ Drehung,
  - $-1 \rightarrow$ Spiegelung,

- aus orthonormierten Spalten (und Zeilen) bestehen.

6.2 Homogene Koordinaten in 2D

6.2.1 Rotation eines Vektors

Um einen Vektor $(x, y)^T \in \Re^2$ um den Winkel ϕ zu rotieren:

$$\begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x\cos\phi - y\sin\phi \\ x\sin\phi + y\cos\phi \end{pmatrix}$$

Eigenschaften der Rotation:

- Die Rotationsmatrix wird mit dem Vektor multipliziert (Matrixmultiplikation).

- Der Nullvektor wird auf sich selbst abgebildet.

- Die Abbildung ist linear.

6.2.2 Translation (Verschiebung) eines Vektors um einen Vektor

Um einen Vektor $(x, y)^T$ um den Vektor $(u, v)^T$ zu verschieben, machen wir folgendes:

$$\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} x + u \\ y + v \end{pmatrix}$$

Eigenschaften der Translation:

- Zum Vektor wird der Translationsvektor addiert (Vektoraddition).

- Der Nullvektor wird nicht auf sich selbst abgebildet (ausser für $u = v = 0,$) uninteressanter Spezialfall.

- Die Abbildung ist im allgemeinen nichtlinear.

Wir möchten beide Operationen - Drehung und Verschiebung - mit einem einheitlichen System (mit Matrixmultiplikation) beschreiben. Denn:

- Rotation kann man mit einer Matrix beschreiben

- Translation nicht direkt mit einer Matrix im 2D-Raum.

Translation geht nicht direkt mit einer Matrix, da wir dazu Addition brauchen und nicht Matrixmultiplikation.

Daher arbeiten wir hier mit homogenen Koordinaten. Die Rotation in homogenen Koordinaten um einen Winkel ϕ sieht wie folgt aus:

$$\begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} x\cos\phi - y\sin\phi \\ x\sin\phi + y\cos\phi \\ 1 \end{pmatrix}$$

Das ist die ganz normale 2D-Rotation, nur erweitert auf 3 Komponenten. Jetzt müssen wir noch wissen, wie wir die Translation in homogenen Koordinaten berechnen können:

Translieren um Vektor (u, v):

$$\begin{pmatrix} 1 & 0 & u \\ 0 & 1 & v \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ s \\ 1 \end{pmatrix} = \begin{pmatrix} x + u \\ y + v \\ 1 \end{pmatrix}$$

Jetzt können wir beide Transformationen mit einer Matrixmultiplikation beschreiben.

Wenn wir zum Beispiel zuerst rotieren, dann verschieben wollen, können wir jetzt das mit Matrix × Matrix × Punkt machen:

$$\begin{aligned} M_{\text{gesamt}} &= \text{Transformation} \times \text{Rotation} \\ P_{\text{neu}} &= M_{\text{gesamt}} \times P_{\text{alt}} \end{aligned}$$

6.2.3 Zusammensetzung von Abbildungen

Hier geht es darum, mehrere Transformationen nacheinander auf einen Vektor anzuwenden. Das kann z. B. sein:

- erst eine Drehung
- dann eine Verschiebung

Hier auch wieder wichtig, die erste Transformation steht rechts, die letzte Transformation am Anfang (links). Sei A die erste Drehung und B die Verschiebung, sieht dies wie folgt aus:

$$C = (B \cdot A) \cdot \vec{x}$$

Beispiel
Wir wollen den Buchstaben F um das Zentrum $(2,3)^T$ drehen:

$$P = \begin{pmatrix} 1 & 2 & 2 & 3 & 3 & 2 & 2 & 4 & 4 & 1 \\ 1 & 1 & 2 & 2 & 3 & 3 & 4 & 4 & 5 & 5 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

Da wir Drehungen immer um den Ursprung $(0,0)$ machen müssen, nutzen wir einen Trick:

1. Translation zum Ursprung: Wir wollen das Zentrum $(2,3)$ zum Ursprung verschieben. Dazu brauchen wir die Translation: $(-2,-3)$:

$$T_1 = \begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & -3 \\ 0 & 0 & 1 \end{pmatrix}$$

2. Rotation um den Ursprung: Wir drehen um 90° im Uhrzeigersinn:

$$R = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Hierbei ist R die nötige Matrix, um eine Abbildung genau um 90° zu drehen. Sollten wir dies um einen bestimmten Winkel machen wollen, nutzen wir eine Matrix, die wir weiter oben schon verwendet haben.

3. Rück-Translation zurück an den Ort: Um die Abbildung wieder an den Ursprungsort zu verschieben, nutzen wir folgende Matrix:

$$T_2 = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix}$$

4. Die Gesamtabbildung erhalten wir, wenn wir folgende Transformation ausführen:

$$M = T_2 \cdot R \cdot T_1$$

Dann ist der neue Punkt:

$$P_{\text{neu}} = M \cdot P$$

6.3 Homogoene Koordinaten in 3D

Im 3D-Raum nutzen wir 4×4-Matrizen:

$$\begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

All diese Rotationen verändern je zwei Koordinaten, während eine Achse (die Rotationsache) unverändert bleibt.

6.3.1 Rotation um x-Achse

$$R_x(\phi) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi & 0 \\ 0 & \sin\phi & \cos\phi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

6.3.2 Rotation um y-Achse

$$R_y(\phi) = \begin{pmatrix} \cos\phi & 0 & \sin\phi & 0 \\ 0 & 1 & 0 & 0 \\ -\sin\phi & 0 & \cos\phi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

6.3.3 Rotation um z-Achse

$$R_z(\phi) = \begin{pmatrix} \cos\phi & -\sin\phi & 0 & 0 \\ \sin\phi & \cos\phi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

6.3.4 Translation im 3D-Raum

Das verschieben eines Punktes (x, y, z) um den Vektor (u, v, w):

$$T(u, v, w) = \begin{pmatrix} 1 & 0 & 0 & u \\ 0 & 1 & 0 & v \\ 0 & 0 & 1 & w \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

7 Einfache lineare Regression

7.1 Motivation

Die lineare Regression ist eine von mehreren Methoden des überwachten Lernens. Sie ermöglicht die Voraussage einern stetigen Zielgrösse, auf der Basis einer Einflussgrösse oder mehrerer Einflussgrössen.

Dabei wird eine lineare Beziehung zwischen den Prediktorvariablen und der Zielvariablen angenommen.

Bei der einfachen linearen Regression gibt es nur eine Prediktorvariable.

7.2 Die beste lineare Approximation - die Regressionsgerade

Unser Ziel ist es, aus gegebenen Punkten (z.B. Messdaten) eine Gerade zu finden, die die Daten möglichst gut beschreibt. Diese Gerade nennt man Regressionsgerade.
Es seien drei Punkte gegeben: $P_1(0,6), P_2(1,0)$ und $P_3(2,0)$. Wir möchten eine lineare Funktion (eine Gerade) finden:

$$b = C + Dt$$

mit zwei Unbekannten C und D.

Wenn man versucht, die exakte Gerade zu finden, die alle drei Punkte exakt trifft, merkt man: Dies ist nicht möglich! Die drei Punkte liegen nicht auf einer einzigen Gerade. Das Gleichungssystem dazu sieht wie folgt aus:

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} C \\ D \end{pmatrix} = \begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix}$$

Dieses System aber hat keine exakte Lösung. Das bedeutet: Es gibt kein Paar C, D, das alle drei Gleichungen gleichzeitig erfüllt.

Die Idee: Wir berechnen die beste Näherung, anstelle der perfekten Lösung.

Wenn man keine exakte Lösung finden kann, sucht man die beste mögliche Gerade, also diejenige, bei der der Fehler zwischen den Punkten und der Gerade möglichst klein ist. Man nennt das: Lineare Regression oder Methode der kleinsten Quadrate.

Man sucht ein $\hat{x} = (\hat{C}, \hat{D})^T$, sodass der Abstand (Fehler) zwischen tatsächlichen Punkten und Werten auf der Gerade im Quadrat minimiert wird.

7.2.1 Geometrische Sicht: Projektion

Was passiert geometrisch?

1. Die Matrix A definiert eine Ebene im 3D-Raum (Spaltenraum).

2. Der Vektor $\vec{b}$ (die Zielwerte) liegt meistens nicht in dieser Ebene.

3. Man sucht den Punkt $A\hat{x}$ in dieser Ebene, der am nächsten bei $\vec{b}$ liegt.

4. Das ist die orthogonale Projektion von $\vec{b}$ auf den Spaltenraum von A.

Dann ist:

- $A\hat{x}$ = projizierter Punkt

- $e = b - A\hat{x}$ = Fehlervektor

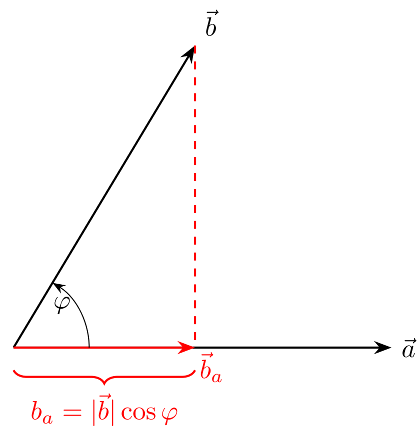
- $||e||^2$ = Fehlerquadrat, das wir minimieren

7.3 Die orthogonale Projektion eines Vektors auf einen Vektor

7.3.1 Was bedeutet Projektion hier

Wir haben zwei Vektoren a und b, die irgendwo im Raum liegen (z.B. in $\Re^2$ oder $\Re^3$).

Die orthogonale Projektion von b auf a ist der Punkt der Geraden durch a, der am nächsten bei b liegt - also der Fusspunkt des Lots von b auf die Gerade durch a.



Gesucht ist also ein Punkt auf der Gerade durch a, der am nächsten bei b liegt. Die Formel lautet:

$$\text{proj}_{\vec{a}}(\vec{b}) = \frac{\vec{a}^T\vec{b}}{\vec{a}^T\vec{a}} \cdot \vec{a}$$

Man kann dies auch als Matrix schreiben mit der Projektionsmatrix:

$$P = \frac{1}{\vec{a}^T\vec{a}} \cdot \vec{a}\vec{a}^T$$

Dann gilt:

$$\text{proj}_{\vec{a}}(\vec{b}) = P \cdot \vec{b}$$

Beispiel
Wir rechnen die orthogonale Projektion des Vektors b auf die Einheitsvektoren in $\Re^3$:

$$\vec{b} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

Die Einheitsvektoren, also Standardbasis:

$$\vec{e_1} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

$$\vec{e_2} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

$$\vec{e_3} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Um die Projektion auf $\vec{e_1}$ zu berechnen:

$$\vec{e_1}^T\vec{b} = (1 \cdot 1 + 0 \cdot 2 + 0 \cdot 3) = 1$$
$$\text{proj}_{\vec{e_1}}(\vec{b}) = 1 \cdot \vec{e_1} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

Für $\vec{e_2}$:

$$\vec{e_2}^T\vec{b} = (0 \cdot 1 + 1 \cdot 2 + 0 \cdot 3) = 2$$
$$\text{proj}_{\vec{e_2}}(\vec{b}) = 2 \cdot \vec{e_2} = \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}$$

Für $\vec{e_3}$:

$$\vec{e_3}^T\vec{b} = (0 \cdot 1 + 0 \cdot 2 + 1 \cdot 3) = 3$$
$$\text{proj}_{\vec{e_3}}(\vec{b}) = 3 \cdot \vec{e_3} = \begin{pmatrix} 0 \\ 0 \\ 3 \end{pmatrix}$$

Wir können beobachten, dass die Summe der Projektionen genau $\vec{b}$ ergibt:

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \vec{b}$$

Das ergibt Sinn, weil die Einheitsvektoren eine Orthonormalbasis bilden. Das heisst:

1. Die Basisvektoren stehen senkrecht aufeinander (also: sie sind orthogonal).
2. Jeder dieser Basisvektoren hat die Länge 1 (also: sie sind normiert).

Beispiel 2
Wir wollen herausfsinden, wie die Projektionsmatrix P aussieht, wenn wir einen beliebigen Vektor $\vec{b} = [b_1, b_2, b_3]^T$ auf den Vektor $a = [a_1, a_2, a_3]^T$ abbilden wollen?

Dazu nehmen wir wieder undere Formel:

$$P = \frac{1}{\vec{a}^T\vec{a}} \cdot \vec{a}\vec{a}^T$$

Damit ist: $\text{proj}_{\vec{a}}(\vec{b}) = P \cdot \vec{b}$. Für $\vec{a}$ bedeutet dies:

$$P = \frac{1}{a_1^2 + a_2^2 + a_3^3} \begin{bmatrix} a_1^2 & a_1a_2 & a_1a_3 \\ a_2a_1 & a_2^2 & a_2a_3 \\ a_3a_1 & a_3a_2 & a_3^2 \end{bmatrix}$$

7.4 Die orthogonale Projektion auf den Spaltenraum einer Matrix

Hier wieder das Gleiche: Wir haben ein Gleichungssystem ohne exakte Lösung. Wir wollen also eine Lösung finden, mit dem möglichst kleinen Fehler.

Der Spaltenraum $C(A)$ ist die Menge alles Vektoren, die durch Linearkombinationen der Spalten von A erzeugt werden können.
Wenn A z.B. folgende Matrix ist:

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}$$

Dann sind die Spalten:

$$\vec{a_1} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$$

Der Spaltenraum ist die Ebene, die diese beiden Spalten aufspannen.

Die Idee: Wenn b nicht in diesem Spaltenraum liegt, gibt es keine Lösung für $Ax = b$. Dann suchen wir:

- Den Punkt $p \in C(A)$, der am nächsten bei b liegt.
- Diesen Punkt nennen wir: $p = A \cdot \hat{x}$
- Das ist die orthogonale Projektion von b auf den Spaltenraum.

7.4.1 Die Lösungsmethode: Normalengleichung

Wenn der Fehler $e = b - A \cdot \hat{x}$ senkrecht zu allen Spalten von A ist, ergibt sich die sogenannte Normalengleichung:

$$A^T A \cdot \hat{x} = A^T b$$

Und daraus folgt:

$$\hat{x} = (A^T A)^{-1} A^T b$$

Die Matrix, die b direkt auf den Spaltenraum von A projiziert, ist:

$$P = A(A^T A)^{-1} A^T$$

Dann ist:

$$p = P \cdot b$$

Beispiel
Wir wollen die Projektionsmatrix für die Matrix A finden:

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}$$

Wir verwenden dazu die Formel:

$$P = A(A^T A)^{-1} A^T$$

Berechnen wir zuerst $A^T A$:

$$A^T A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 3 \\ 3 & 5 \end{bmatrix}$$

Jetzt invertieren wir $A^T A$:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Dies angewendet auf unser Resultat für $A^T A$:

$$(A^T A)^{-1} = \frac{1}{6} \begin{bmatrix} 5 & -3 \\ -3 & 3 \end{bmatrix}$$

Jetzt multiplizieren wir alles: $P = A(A^T A)^{-1} A^T$. Dazu rechnen wir:

1. $B = (A^T A)^{-1} A^T$, dies ergibt eine 2×3-Matrix.
2. Dann $P = A \cdot B$, dies ergibt eine 3×3-Matrix.

Dies sieht dann so aus:

$$B = (A^T A)^{-1} A^T = \frac{1}{6} \begin{bmatrix} 5 & -3 \\ -3 & 3 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix}$$

$$P = A \cdot B = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \cdot \frac{1}{6} \begin{bmatrix} 5 & 2 & -1 \\ -3 & 0 & 3 \end{bmatrix}$$

Am Ende bekommen wir:

$$P = \begin{bmatrix} 0.833 & 0.333 & -0.167 \\ 0.333 & 0.333 & 0.333 \\ -0.167 & 0.333 & 0.833 \end{bmatrix}$$

Diese Matrix P hat die Eigenschaft, dass sie jeden Vektor $b \in \Re^3$ orthogonal auf den Spaltenraum von A projiziert. Ausserdem ist sie symmetrisch und idempotent (d. h. $P^2 = P$).

7.4.2 Beweisen, dass Projektion richtig ist

Wenn wir zeigen wollen, dass P einen Vektor $\vec{b}$ im Spaltenraum von A auf sich abbildet, d. h. $P\vec{b} = \vec{b}$:

Wenn $\vec{b}$ im Spaltenraum von A liegt, dann gibt es ein $\vec{x}$, sodass $\vec{b} = A\vec{x}$. Die Frage ist: Was passiert, wenn man diesen Vektor durch die Projektionsmatrix schickt?

Angenommen $\vec{b} = A\vec{x}$, für irgendein $\vec{x}$. Dann folgt:

$$P\vec{b} = A(A^TA)^{-1}A^T\vec{b} = A(A^TA)^{-1}A^T\vec{x}$$

Nun ist A^TA eine $n \times n$-Matrix und damit ist:

$$\begin{aligned}&(A^TA)^{-1}A^TA = I \quad \text{, weil:}\\ &B^{-1} \cdot B = I\end{aligned}$$

Deshalb:

$$P\vec{b} = A \cdot I \cdot \vec{x} = A\vec{x} = \vec{b}$$

Wenn $\vec{b} \in C(A)$, dann gilt:

$$P\vec{b} = \vec{b}$$

7.4.3 Beweisen, dass Projektion eines orthogonalen Vektors Null ist

Wenn ein Vektor $\vec{b}$ orthogonal zum Spaltenraum von A steht, dann gilt:

$$P\vec{b} = 0 \quad \text{mit} \quad P = A(A^TA)^{-1}A^T$$

Angenommen: $\vec{b} \in C(A)^{\perp}$:

$$\forall a_i \text{ (Spalten von A): } \vec{a_i}^T\vec{b} = 0 \Rightarrow A^T\vec{b} = \vec{0}$$

Nun also:

$$P\vec{b} = A(A^TA)^{-1}A^T\vec{b} = A(A^TA)^{-1}A^T\vec{0} = A \cdot \vec{0} = \vec{0}$$

Somit haben wir gezeigt, dass:

$$\vec{b} \perp C(A) \quad \Rightarrow \quad P\vec{b} = \vec{0}$$

7.4.4 Bedeutung von p und e

Gegeben ist ein lineares Gleichungssystem $A\vec{x} = \vec{b}$, das keine exakte Lösung hat (weil $\vec{b}$ nicht im Spaltenraum von A liegt). Dann betrachten wir:

- $\vec{p} = A\hat{x}$:
  Das ist die Projektion von $\vec{b}$ auf den Spaltenraum von A - der nächstgelegene Punkt zu $\vec{b}$, der im Spaltenraum liegt.
- $\vec{e} = \vec{b} - \vec{p}$:
  Das ist der Fehlervektor, also der Abstand zwischen $\vec{b}$ und seiner Projektion $\vec{p}$. Dieser Vektor steht senkrecht auf dem Spaltenraum.

Die gesamte Formel setzt sich folgendermassen zusammen.

$$\vec{b} = \vec{p} + \vec{e} \quad \text{mit} \quad \vec{p} = A\hat{x}, \quad \vec{e} \perp C(A)$$

Folgendes können wir also daraus schliessen:

- $\vec{p} = A\hat{x}$ ist die orthogonale Projektion von $\vec{b}$ auf den Spaltenraum von A.
- $\vec{e} = \vec{b} - \vec{p}$ ist der Fehlervektor, der senkrecht zum Spaltenraum von A steht.
- Es gilt: $\vec{b} = \vec{p} + \vec{e}$
- Die Zerlegung ist die Grundlage der Methode der kleinsten Quadrate.

A^TA ist dann invertierbar, wenn A lineare unabhängige Spalten hat. Spalten sind dann linear unabhängig, wenn:

$$A\vec{x} = 0 \Rightarrow \vec{x} = 0$$

7.5 Vorgehen

Das Regressionsmodell sieht wie folgt aus:

$$b = C + Dt + \epsilon$$

- b: Zielgrösse (z.B. Messwert)
- t: Einflussgrösse (z.B. Zeit)
- C, D: Regressionskoeffizienten
- ϵ: Messfehler (Abweichung)

Das ist eine lineares Modell mit Konstante C und Steigung D.

7.5.1 Die Lösungsformel

$$\hat{x} = \begin{bmatrix} \hat{C} \\ \hat{D} = (A^T A)^{-1} A^T b \end{bmatrix}$$

Diese Formel berechnet die beste Approximation für die Regressionsgerade.

7.5.2 SQR - Summe der Quadrate der Residuen

$$SQR = ||\vec{e}||^2 = ||\vec{b} - \vec{p}||^2$$

- Mass für die Gesamtabweichung
- Je kleiner, desto besser die Anpassung der Regressionsgeraden.

7.5.3 Prädiktion (Vorhersage)

Mit dem berechneten Modell:

$$b = \hat{C} + \hat{D} \cdot t$$

können wir für neue Werte t eine Zielgrösse b vorhersagen.

8 Multiple lineare Regression

Was passiert, wenn wir mehrere Einflussgrössen haben. Mit nur einer Einflussgrösse, können wir die Methoden des Kapitel 7 nutzen. Mit mehreren Einflussgrössen brauchen wir die multiple lineare Regression.

Das Ziel ist es immernoch, auf mehreren Dimensionen (Einflussgrössen) eine Gerade zu finden, die die Datenpunkte möglichst gut beschreibt.

8.1 Die regressionshyperebene

8.2 Das Modell

Das Modell dazu sieht so aus:

$$y = b_0 + b_1 x_1 + b_2 x_2 + \cdots + b_n x_n + \epsilon$$

- y: Zielwert (z.B. gestemmtes Gewicht)
- $x_1, x_2, ...$: Einflussgrössen (z.B. Training, Proteinzufuhr)
- $b_0, b_1, ...$: Regressionskoeffizienten (die man sucht)
- ϵ: Fehler, also Abweichung vom Modell

Für jeden der Datenpunkte, bekommen wir eine Gleichung dieser Art.

8.3 Darstellung mit Matrizen

Das alles kann man kompakt schreiben als:

$$y = Xb + e$$

wobei:

- y: Spalte mit allen Zielgrössen
- X: Datenmatrix, jede Zeile ein Datenpunkt, jede Spalte ein Einflussfaktor (plus 1-er Spalte für b_0)
- b: die gesuchten Regressionskoeffizienten
- e: Fehlervektor

8.4 Möglichst kleiner Fehler

Wir wollen die summe der Fehlerquadrate minimieren:

$$||e||^2 = ||y - X\hat{b}||^2$$

Das passiert genau dann, wenn der Fehlervektor orthogonal (senkrecht) zum Spaltenraum von X steht. Daraus folgt die sogenannte Normalengleichung:

$$X^T X\hat{b} = X^T y \quad \Rightarrow \quad \hat{b} = (X^T X)^{-1} X^T y$$

8.5 Die Hutmatrix - Projektion auf Modell

Sobald wir $\hat{b}$ haben, können wir alle vorhergesagten y-Werte berechnen:

$$\hat{y} = X\hat{b} = Py$$

Die Matrix $P = X(X^T X)^{-1} X^T$ nennt man:

- Projektionsmatrix
- Prädiktionsmatrix
- oder Hutmatrix

Zusätzlich gibt es die Residualmatrix:

$$Q = I - P$$

Damit berechnet man den Fehlervektor $e = Qy$, der senkrecht zum Spaltenraum liegt.

8.6 Bildliche Vorstellung

$\hat{y} = X\hat{b}$ ist der Punkt im Spaltenraum von X, der am nächsten bei y liegt. Also ist $\hat{y}$ die orthogonale Projektion von y auf $C(X)$. Der Fehler $e = y - \hat{y}$ steht senkrecht auf der Ebene.

Die Spalten von X spannen eine Ebene oder einen Raum auf (z.b. 2D in einem 5D-Raum), und y wird orthogonal darauf projiziert.

Ein Beispiel: Mit 100 Personen (Zeilen) und drei Einflussgrössen (Spalten) (z.B.: Training, Schlaf, Ernährung):

$$X = \begin{bmatrix} 1 & 5 & 6 \\ 2 & 6 & 3 \\ 0.5 & 4 & 5 \\ \dots & \dots & \dots \\ a_{100} & b_{100} & c_{100} \end{bmatrix} \in \Re^{100\times 3}$$

8.7 Das Bestimmtheitsmass

Wir müssen auch überprüfen können, wie akkurat das Modell ist. Die Idee:

Wenn wir mit einem Modell etwas vorhersagen wollen, wollen wir wissen: Wie viel von der Variation in meinen Daten kann mein Modell erklären?

8.7.1 SQT - Gesamtabweichung (Total)

Dies beschreibt: Wie weit sind die Messwerte insgesamt vom Durchschnitt entfernt?

$$SQT = \sum_{i=1}^{m} (y_i - \overline{y})^2$$

8.7.2 SQE - Erklärte Abweichung

Wie viel kann das Modell erklären?

$$SQE = \sum_{i=1}^{m} (\hat{y}_i - \overline{y})^2$$

8.7.3 SQR - Residualfehler (Rest)

Was bleibt übrig, was das Modell nicht erklären kann?

$$SQR = \sum_{i=1}^{m} (y_i - \hat{y}_i)^2$$

8.7.4 Formel für das Bestimmtheitsmass

$$R^2 = \frac{SQE}{SQT} = 1 - \frac{SQR}{SQT}$$

wobei:

- $R^2 = 1$: Perfekte Erklärung - Modell passt perfekt
- $R^2 = 0$: Modell erklärt gar nichts
- Typisch in der Praxis: z.B. $R^2 = 0.85 \rightarrow$ Modell erklärt 85% der Streuung.

8.7.5 Mathematisch mit Matrizen

$$SQT = y^T My, \quad SQR = y^T Qy \Rightarrow R^2 = 1 - \frac{y^T Qy}{y^T My}$$

- M: Zentrierungsmatrix, entfernt den Mittelwert von y.
- $Q = I - P$: Residualmatrix (Fehlerprojektion)
- P: Projektionsmatrix (Vorhersage)

9 Eigenwerte und Eigenvektoren I

Die Motivation hinter diesem Thema ist, lineare Abbildungen besser zu verstehen - insbesondere solche, bei denen bestimmte Vektoren ihre Richtung beibehalten (auch wenn sich ihre Länge ändert).

Solche Vektoren nennt man Eigenvektoren, die zugehörigen Skalierungsfaktoren sind die Eigenwerte.

9.1 Eigenwerte und Eigenvektoren

Gegeben ist eine quadratische Matrix A. Man sucht Vektoren $y \neq 0$, für die gilt:

$$Ax = \lambda x$$

- x heisst Eigenvektor
- λ heisst Eigenwert

Das bedeutet: Die Matrix A skaliert x, ändert aber nicht seine Richtung.

9.1.1 Berechnung von Eigenvektoren und Eigenwerten

1. Eigenwerte: Löse die charakteristische Gleichung:

   $$\det(A - \lambda I) = 0$$

   Dies ergibt ein Polynom, dessen Lösungen die Eigenwerte λ sind.

2. Eigenvektoren: Für jeden gefundenen Eigenwert λ, löse:

   $$(A - \lambda I)x = 0$$

   Das ergibt den zugehörigen Eigenvektor $x \neq 0$.

Beispiel
Matrix:

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

Charakteristisches Polynom:

$$\det(A - \lambda I) = \lambda^2 - 4\lambda + 3 = 0 \Rightarrow \lambda_1 = 3, \lambda_2 = 1$$

- Zu $\lambda_1 = 3$: Eigenvektor $x_1 = (1, 1)^T$
- Zu $\lambda_2 = 1$: Eigenvektor $x_2 = (1, -1)^T$

Beispiel 2
Matrix:

$$A = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}$$

Charakteristische Gleichung:

$$\lambda^2 - 2\lambda - 8 = 0 \Rightarrow \lambda_1 = 4, \lambda_2 = -2$$

- Eigenvektor zu $\lambda_1 = 4 : x_1 = \frac{1}{\sqrt{2}}(1, 1)^T$
- Eigenvektor zu $\lambda_2 = -2 : x_2 = \frac{1}{\sqrt{2}}(1, -1)^T$

Einige Merksätze:

- Eine Matrix hat nur dann Eigenwerte und -vektoren, wenn die charakteristische Gleichung eine nicht-triviale Lösung hat.
- Der Eigenwert drückt die Skalierung eines Eigenvektors durch die Matrix aus.

9.2 Symmetrische Matrizen

Symmetrische Matrizen haben besonders gutartige Eigenwerte (reell, orthogonal normierbare Eigenvektoren).

Eine Matrix A ist symmetrisch, wenn $A^T = A$ gilt. Das bedeutet, dass die Matrix entlang ihrer Hauptdiagonalen gespiegelt ist.

Ein Beispiel für eine symmetrische Matrix:

$$A = \begin{pmatrix} 2 & 3 \\ 3 & 4 \end{pmatrix} \quad = \quad A^T = \begin{pmatrix} 2 & 3 \\ 3 & 4 \end{pmatrix}$$

Mit symmetrischen Matrizen ist auch eine Spektralzerlegung möglich: $A = V \Lambda V^T$

9.3 Eigenschaften der Eigenwerte

- Die Determinante einer Matrix ist gleich dem Produkt der Eigenwerte der Matrix:

$$\det(A) = \prod_{i=1}^{n} \lambda_i$$

- Die Spur einer Matrix ist gleich der Summe der Eigenwerte der Matrix:

$$\operatorname{spur}(A) = \sum_{i=1}^{n} \lambda_i$$

Dabei ist die Spur einer Matrix die Summe der Diagonalelemente der Matrix:

$$\operatorname{spur}(A) = \sum_{i=1}^{n} a_{i,i}$$

Beispiel
Matrix:

$$A = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}$$

Die Eigenwerte sind hier schon bekannt:
$\lambda_1 = 4, \lambda_2 = -2$.
Nun berechnen wir die Eigenschaften:

1. Determinante der Matrix A:

$$\det(A) = a \cdot d - b \cdot c$$

Für unsere Matrix:

$$\det(A) = 1 \cdot 1 - 3 \cdot 3 = 1 - 9 = -8$$

Man multipliziert also die Diagonalelemente $(1 \cdot 1)$ und subtrahiert das Produkt der Gegendiagonale $(3 \cdot 3)$. Dann zeigt man:

$$\det(A) = \lambda_1 \cdot \lambda_2 = 4 \cdot (-2) = -8$$

2. Spur der Matrix A: Die Spur (engl. trace) ist die Summe der Diagonalelemente einer Matrix:

$$\operatorname{spur}(A) = a_{11} + a_{22} = 1 + 1 = 2$$

Die Spur ist auch gleich der Summe der Eigenwerte:

$$\operatorname{spur}(A) = \lambda_1 + \lambda_2 = 4 + (-2) = 2$$

9.4 Die Spektralzerlegung einer symmetrischen Matrix

Die Spektralzerlegung (auch Spektraldarstellung) ist eine Art, eine symmetrische Matrix A in folgende FOrm zu zerlegen:

$$A = V \Lambda V^T$$

Hierbei ist:

- A: eine reelle symmetrische Matrix (also $A^T = A$)
- V: Matrix der orthonormalen Eigenvektoren (jeder als Spalte)
- Λ: Diagonalmatrix mit den zugehörigen Eigenwerten auf der Hauptdiagonale
- V^T: die Transponierte von V

Beispiel
Gegeben:

$$A = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}$$

Die Eigenwerte sind: $\lambda_1 = 4, \lambda_2 = -2$. Die zugehörigen, normierten Eigenvektoren sind:

$$x_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad x_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

Dann ist:

$$V = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

$$\Lambda = \begin{pmatrix} 4 & 0 \\ 0 & -2 \end{pmatrix}$$

Und somit: $A = V \Lambda V^T$

10 Eigenwerte und Eigenvektoren II, Diagonalisierung

10.1 Positiv definite Matrizen

Eine Matrix ist positiv definit, wenn sie immer einen positiven Wert ergibt, wenn wir den folgenden Ausdruck berechnen (für alle $x \neq 0$):

$$x^T Ax$$

Man nennt $x^T Ax$ eine quadratische Form. Die Definition sieht wie folgt aus:

$$x^T Ax > 0 \quad \forall x \neq 0$$

Sie ist positiv definit, wenn:

$$x^T Ax \geq 0$$

Beispiel
Gegeben ist die Matrix:

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$$

Wir untersuchen dies nun folgendermassen:

$$x^T Ax = (x_1, x_2) \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$
$$= x_1^2 + 4x_1x_2 + 4x_2^2 = (x_1 + 2x_2)^2$$

Dieser Ausdruck ist immer ≥ 0, aber nicht > 0, wenn $x = (-2, 1)$, denn dann: $(-2 + 2 \cdot 1)^2 = 0$.
Also A ist hier nicht positiv definit, sondern nur positiv semidefinit.

10.1.1 Alternative Prüfung - über Eigenwerte

Eine symmetrische Matrix ist genau dann positiv definit, wenn alle Eigenwerte positiv sind.

Beispiel
Gegeben ist die Matrix:

$$A = \begin{pmatrix} 5 & 3 \\ 3 & 5 \end{pmatrix}$$

Charakteristische Gleichung:

$$\begin{aligned} \det(A - \lambda I) &= (5 - \lambda)^2 - 9 \\ &= \lambda^2 - 10\lambda + 16 \end{aligned} \Rightarrow \lambda_1 = 8, \lambda_2 = 2$$

Da hier also beide Eigenwerte positiv sind, ist die Matrix positiv definit.

10.2 Diagonalisierung von beliebigen quadratischen Matrizen

Eine Matrix A ist diagonalisierbar, wenn es eine invertierbare Matrix V und eine Diagonalmatrix Λ gibt, sodass:

$$A = V \Lambda V^{-1}$$

- V ist die Matrix der Eigenvektoren von A
- Λ ist die Diagonalmatrix mit den Eigenwerten von A auf der Hauptdiagonalen.

Beispiel
Gegeben ist die Matrix:

$$A = \begin{pmatrix} 2 & 1 \\ 0 & 3 \end{pmatrix}$$

1. Eigenwerte berechnen:

$$\begin{aligned} \det(A - \lambda I) &= (2 - \lambda)(3 - \lambda) - 0 \\ &= \lambda^2 - 5\lambda + 6 = 0 \end{aligned}$$

$$\lambda_1 = 2, \lambda_2 = 3$$

2. Eigenvektoren berechnen:

   - Für $\lambda_1 = 2$:

$$(A - 2I)x = 0 \Rightarrow (1, 0)^T$$

   - Für $\lambda_2 = 3$:

$$(A - 3I)x = 0 \Rightarrow (1, 1)^T$$

3. Zusammenbauen:

$$V = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad \Lambda = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$$

   Dann ist folgendes die Diagonalisierung:

$$A = V \Lambda V^{-1}$$

   wobei Λ die Diagonalmatrix mit den entpsrechenden Eigenvektoren ist.

10.3 Eigenschaften von Determinanten und Spuren

Für beliebige Matrizen $A(n \times n)$ gilt:

- Die Spur der Matrix ist gleich der Summe der Eigenwerte.
- Die Determinante der Matrix ist gleich dem Produkt der Eigenwerte.

Für beliebige Matrizen $A(n \times n)$ und $B(n \times n)$ und ein Skalar c gilt:

$$\begin{aligned}
\mathrm{tr}(A+B) &= \mathrm{tr}(A) + \mathrm{tr}(B) \\
\mathrm{tr}(cA) &= c \cdot \mathrm{tr}(A) \\
\det(cA) &= c^n \cdot \det(A) \\
\mathrm{tr}(AB) &= \mathrm{tr}(BA) \\
\det(AB) &= \det(A) \cdot \det(B) \\
\det(AB) &= \det(BA) \\
\det(A^{-1}) &= (\det(A))^{-1}
\end{aligned}$$

10.4 Jordan-Form

Die Jordan-Form ist eine Matrix, die aus Blöcken besteht. Ein Beispiel einer Jordan-Form ist:

$$J = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 3 \end{pmatrix}$$

- Auf der Hauptdiagonale: der Eigenwert λ.
- Direkt darüber: Einsen
- Alles andere: Nullen

Wenn wir also nicht genug Eigenvektoren haben, um die Matrix zu diagonalisieren, können wir sie in Jordan-Form bringen.

Beispiel
Gegeben ist die Matrix:

$$A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$$

1. Eigenwerte berechnen:
   Berechne die charakteristische Gleichung:

   $$\det(A - \lambda I) = \det \begin{pmatrix} 2-\lambda & 1 \\ 0 & 2-\lambda \end{pmatrix} = (2-\lambda)^2 = 0$$

   Eigenwert: $\lambda = 2$ (zweifach)

2. Eigenvektoren berechnen:
   Lösen:

   $$(A - 2I)x = 0 \Rightarrow \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

   Diese Gleichung ergibt:

   $$x_2 = 0, \quad x_1 \text{ frei} \Rightarrow x = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

   Dies gibt nur einen Eigenvektor, obwohl der Eigenwert $\lambda = 2$ eine Vielfachheit von 2 hat. Das bedeutet, dass die Matrix nicht diagonalisierbar ist.

3. Wir müssen nun einen Vektor suchen, der nicht die Lösung von $(A - 2I)x = 0$ ist, sondern:

   $$(A - 2I)^2 x = 0, \text{ aber } (A - 2I)x \neq 0$$

4. Verallgemeinerten Eigenvektoren finden Wir setzen:

   $$(A - 2I) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

   Und wählen $X^{(2)} = (0,1)^T$. Dann gilt:

   $$(A - 2I)X^{(2)} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = x^{(1)} \neq 0 \Rightarrow \text{gültig}$$

5. Jordan-Form aufstellen:
   Wir haben den Eigenwert $\lambda = 2$ mit Vielfachheit 2 und nur einen Eigenvektor. Also Jordan-Block der Grösse 2. Die Jordan-Form lautet:

   $$J = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$$

   Das ist tatsächlich die gleiche Matrix wie A, das heisst: A ist bereits in Jordan-Form.

10.5 Ähnliche Matrizen

Man nennt zwei Matrizen A und B ähnlich, wenn es eine invertierbare Matrix M gibt, sodass:

$$B = M^{-1}AM$$

Sind A und B ähnliche Matrizen, dann haben sie die selben Eigenwerte.

Falls x ein Eigenvektor der Matrix B zum Eigenwert λ ist, dann gilt:

$$Bx = M^{-1}AMx = \lambda x$$

daraus folgt nach Multiplikation mit M von links:

$$AMx = \lambda Mx$$

das heisst, Mx ist ein Eigenvektor von A zum Eigenwert λ.

Änhliche Matrizen haben die selben Eigenschaften:

- gleiche Spur
- gleiche Determinante
- gleiche Eigenwerte
- gleiche Eigenvektoren (nur in anderer Basis)

10.6 Potenzen von Matrizen

In vielen Anwendungen will man wissen:

$$A^n \text{ oder } A^k x$$

Also was passiert, wenn man eine Matrix mehrfach anwendet.

Theoretisches Grundprinzip:

$$A = V\Lambda V^{-1} \quad \Rightarrow \quad A^n = V\lambda^n V^{-1}$$

Dann ist das hoch n potenzierte A ganz einfach zu berechnen, weil:

- Λ eine Diagonalmatrix ist. Potenzieren heisst: einfach jeden Diagonaleintrag hoch n rechnen.
- Das spart Zeit und Rechenaufwand

10.6.1 Merkregel: Potenzen und Eigenwerte

Wenn:

- $Ax = \lambda x$ (Eigenwertproblem), dann gilt:
- $A^n x = \lambda^n x$

Die Eigenvektoren bleiben gleich, Eigenwerte werden potenziert.

10.6.2 Entwicklung von Potenzen

Wenn wir eine Anfangsgrässe u_0 haben, und wissen wollen, wie sie sich entwickelt:

$$u_k = A^k u_0 = V\Lambda^k V^{-1} u_0$$

Wenn wir u_0 zuerst als Linearkombination der Eigenvektoren schreiben:

$$u_0 = c_1x_1 + c_2x_2 \Rightarrow u_k = c_1\lambda_1^k x_1 + c_2\lambda_2^k x_2$$

10.6.3 Beispiele

Beispiel 1 - Kleine Matrix
Gegeben ist die Matrix:

$$A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} \Rightarrow \text{hat nur einen Eigenwert } \lambda = 2$$

$$A^2 = \begin{pmatrix} 4 & 4 \\ 0 & 4 \end{pmatrix} \Rightarrow \text{Eigenwert } 2^2 = 4$$

Der Eigenvektor bleibt gleich:

$$x = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Beispiel 2 - Matrix potenzieren mit Diagonalisierung
Gegeben ist die Matrix:

$$A = \begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix} \Rightarrow \text{ Eigenwerte: } \lambda_1 = 1, \lambda_2 = 2$$

Dann:

$$A^{100} = V\Lambda^{100}V^{-1} \Rightarrow \Lambda^{100} = \begin{pmatrix} 1^{100} & 0 \\ 0 & 2^{100} \end{pmatrix}$$

So wird das rechnen von A^{100} sehr einfach.

Beispiel 3 - Fibonacci-Zahlen mit Matrixpotenz
Die Fibonacci-Folge $F_0 = 0, F_1 = 1, F_2 = 1, F_3 = 2, ...$ entsteht über:

$$\begin{pmatrix} F_{n+1} \\ F_n \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} F_n \\ F_{n-1} \end{pmatrix}$$

wobei A folgende Matrix ist:

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$$

Wir können hier mit Diagonalisierung arbeiten um von A effizient direkt F_43 zu berechnen, ohne alle vorherigen Werte zu kennen!

Konkret ein Beispiel zu F_43:

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, u_0 = \begin{pmatrix} F_1 \\ F_0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Dann gilt:

$$u_n = A^{n-1} u_0 \Rightarrow F_{43} = 1. \text{ Komponente von } A^{42} u_0$$

Für die Eigenwerte diagonalisieren wir A:

$$\det(A - \lambda I) = \lambda^2 - \lambda - 1 = 0 \Rightarrow \lambda_{1,2} = \frac{1 \pm \sqrt{5}}{2}$$

Somit ist $\lambda_1 = \phi = \frac{1+\sqrt{5}}{2} \approx 1.618$ und $\lambda_2 = \frac{1-\sqrt{5}}{2} \approx -0.618$.
Jeder verwwenden wir für die Berechnung die Binet-Formel:

$$F_n = \frac{\phi^n - \overline{\phi}^n}{\sqrt{5}}$$

Diese Formel ist eine direkte Folge davon, dass man A^n als $V \Lambda^n V^{-1}$ schreiben kann, wobei Λ^n die Diagonalmatrix mit den Eigenwerten hoch n ist.
Wir berechnen also:

$$\begin{aligned} &\phi^{43} \approx 1.618^{43} \approx 9.68 \times 10^8 \\ &\overline{\phi}^{43} \approx (-0.618)^{43} \approx -1.02 \times 10^{-9} \end{aligned}$$

Und dann ist:

$$F_{43} = \frac{9.68 \times 10^8 - (-1.02 \times 10^{-9})}{\sqrt{5}} \approx 433'494'437$$

10.6.4 Markovmatrizen

Eine Markovmatrix beschreibt ZUstandsübergänge in einem System mit Wahrscheinlichkeiten.
Folgende Eigenschaften gelten für eine Markovmatrix M:

- Alle Einträge sind ≥ 0.
- Jede Spalte summiert sich zu 1.

Wie entwickelt sich ein Zustand? Man beginnt mit einem Startzustand x_0 (z.B. mit 100% in Zustand A) und wiederholt:

$$x_1 = Ax_0, \quad x_2 = Ax_1, \quad \ldots, \quad x_k = A^k x_0$$

Ziel: Stationärer Zustand: x_∞. Was passiert, wenn man A sehr oft anwendet?

Beispiel
Gegeben ist die Matrix A und ein Startzustand x_0:

$$A = \begin{pmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{pmatrix}, \quad x_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

In einem ersten Schritt multiplizieren wir die Matrix A mit dem Startzustand x_0:

$$x_1 = A \cdot x_0 = \begin{pmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0.8 \\ 0.2 \end{pmatrix}$$

Für den nächsten Zustand x_2 multiplizieren wir erneut:

$$\begin{aligned} x_2 = A \cdot x_1 &= \begin{pmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{pmatrix} \cdot \begin{pmatrix} 0.8 \\ 0.2 \end{pmatrix} \\ &= \begin{pmatrix} 0.8 \cdot 0.8 + 0.3 \cdot 0.2 \\ 0.2 \cdot 0.8 + 0.7 \cdot 0.2 \end{pmatrix} = \begin{pmatrix} 0.70 \\ 0.30 \end{pmatrix} \end{aligned}$$

Mit jedem Schritt nähern wir uns dem stationären Zustand an. Wenn wir dies weiterführen, werden die Werte immer stabiler und ändern sich nicht mehr signifikant.

11 Singulärwertzerlegung I

11.1 Motivation und Einführung

Die Singulärwertzerlegung (engl. Singular Value Decomposition, SVD), ist eine Methode, mit der man jede beliebige Matrix A (auch nicht-quadratische) zerlegen kann in drei spezielle Matrizen:

$$A = U\Sigma V^T$$

Hierbei ist:

- U: Links-Singulärvektoren. Dies ist eine orthogonale Matrix $m \times m$, $U^T U = I$.
- Σ: Singulärwerte / Diagonalmatrix. Dies ist eine $m \times n$-Matrix mit nicht-negativen Zahlen auf der Diagonalen.
- V: Rechts-Singulärvektoren. Dies ist eine orthogonale Matrix $n \times n$, $V^T V = I$.

Warum ist das wichtig und was bringt uns das? Die SVD hilft uns zu verstehen:

- wie eine lineare Abbildung wirkt:
  Jede Abbildung $y = Ax$ ist eine Drehung/Spiegelung $\rightarrow$ Streckung *rightarrow* weitere Drehung/Spiegelung.
- wie man Daten reduziert oder komprimiert (z.B. bei Bildern)
- wie man mit nicht-invertierbaren Matrizen arbeitet: Pseudoinverse
- wie man grosse Datenmengen in Data Science oder Machine Learning vereinfacht: Principal Component Analysis (PCA)

11.1.1 Beispiel: Bildkompression

Ein Graustufenbild (z.B. 49×83 Pixel) wird als Matrix A gespeichert. Wir zerlegen es mit SVD:

$$A = U\Sigma V^T \Rightarrow A \approx \sum_{i=1}^{k} \sigma_i u_i v_i^T$$

Wenn man nur die grössten k Singulärwerte nimmt, spart man viel Speicher bei nur geringem Qualitätsverlust. Beispiel mit $k = 10$:

- statt 4101 Werte, speichert man nur 1330 Werte
- Kompressionsfaktor: 0.327 (nur $\approx 33\%$ nötig)

11.2 Was macht die Singulärwertzerlegung (SVD)

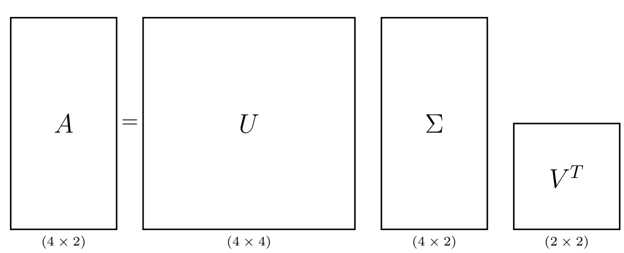
Die SVD zeigt, wie eine Matrix A als Abbildung auf einen Vektor wirkt:

$$Ax = U\Sigma V^T x$$

Schritt für Schritt:

1. V^T:
   Dreht oder spiegelt den Vektor x in eine neue Basis (Rechts-Singulärvektoren).
2. Σ:
   Skaliert die Richtungen durch Singulärwerte $\sigma_1, \ldots, \sigma_r$
3. U:
   Dreht oder spiegelt das Ergebnis in den Zielraum (Links-Singulärvektoren).

Die Eigenwert-Zerlegung funktioniert nur bei quadratischen Matrizen, die SVD funktiniert bei jeder beliebigen Matrix.



11.3 Vorgehen Singulärwertzerlegung

11.3.1 Wie berechnet man die SVD

Man verwendet zwei wichtige Produkte:

- $A^T A$: symmetrisch: liefert Rechts-Singulärvektoren V
- AA^T: symmetrisch: liefert Links-Singulärvektoren U

Die Eigenwerte dieser Matrizen sind gleich (nicht die Eigenvektoren). Die Singulärwerte von A sind:

$$\sigma_i = \sqrt{\lambda_i}$$

Wie entsteht V
V entsteht anhand der Eigenvektoren λ_i von A:

$$V = \begin{bmatrix} | & | \\ v_1 & v_2 \\ | & | \end{bmatrix}$$

Wie entsteht U
Wenn Σ und V schon berechnet sind, können wir daraus U herleiten:

$$u_i = \frac{1}{\sigma_i} \cdot Av_i$$

Das bedeutet:

- Wir nehmen den Vektor v_i aus V
- Multiplizieren ihn mit A (ergibt einen Vektor)
- Teilen diesen durch den passenden Singulärwert σ_i
- Das Ergebnis ist der zugehörige Vektor u_i

Somit:

$$U = \begin{bmatrix} | & | \\ u_1 & u_2 \\ | & | \end{bmatrix}$$

Wie entsteht Σ
Σ erstellen wir mit den berechneten Singulärwerten von A:

$$\sigma_i = \sqrt{\lambda_i}$$

Dabei ist λ_i ein Eigenwert von $A^T A$. Wir können diese dann diagonal in eine Matrix setzen, um Σ zu erhalten:

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix}$$

11.3.2 Beispiel 1

Gegeben ist die Matrix:

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

1. Berechnen wir: $A^T A$:

$$A^T A = \begin{pmatrix} 2 & 1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 2 & 2 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 5 & 5 \\ 5 & 5 \end{pmatrix}$$

2. Eigenwerte von $A^T A$:

$$\lambda_1 = 10, \lambda_2 = 0 \Rightarrow \sigma_1 = \sqrt{10}, \sigma_2 = \sqrt{0}$$

3. Eigenvektoren zu λ_1, λ_2:

$$v_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$
$$v_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$
$$V = \begin{bmatrix} | & | \\ v_1 & v_2 \\ | & | \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

Diese sind die Spalten von V.

4. Berechnen von U:

$$Av_i = \sigma_i u_i \Rightarrow u_i = \frac{1}{\sigma_i} Av_i$$

Das Ergebnis ist dann:

$$U = \tfrac{1}{\sqrt{5}} \begin{pmatrix} 2 & 1 \\ 1 & -2 \end{pmatrix}$$
$$\Sigma = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} = \begin{pmatrix} \sqrt{10} & 0 \\ 0 & 0 \end{pmatrix}$$

Sollte eines der $\sigma_i = 0$ sein, kann man die Formel für u nicht verwenden, da man eine Division durch 0 erhält. Dann sucht man einfach eines der u und sucht dazu die orthogonalen Vektoren, um die anderen u_i zu erhalten.

12 Singulärwertzerlegung II

12.1 Reduzierte Singulärwertzerlegung

Die reduzierte Singulärwertzerlegung ist eine platzsparende Version der normalen SVD. Sie wird verwendet, wenn man weiss, dass die Matrix A nur Rang r hat - also nur r Singulärwerte ungleich Null.

Dann muss man auch nur die ersten r Spalten von U und V, sowie die $r \times r$-Diagonalmatrix Σ_r betrachten.

12.1.1 Normale vs. reduzierte SVD

Die normale SVD ist:

$$A = U\Sigma V^T$$

mit:

- $U \in \Re^{m \times n}$
- $\Sigma \in \Re^{m \times n}$
- $V \in \Re^{n \times n}$

Bei reduzierter SVD kürzt man alles, was Null wäre weg:

$$A = U_r \Sigma_r V_r^T$$

mit:

- $U_r \in \Re^{m \times r}$ - nur die ersten r Spalten von U
- $\Sigma_r \in \Re^{r \times r}$ - nur die r Singulärwerte
- $V_r \in \Re^{n \times r}$ - nur die ersten r Spalten von V

12.1.2 Ein Beispiel

Gegeben sei die Matrix:

$$A = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}$$

Normale SVD:

$$A = U\Sigma V^T = \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \cdot \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

hier ist nur $\sigma_1 = 2 \neq 0$ also Rang $r = 1$

Reduzierte SVD:

Jetzt kürzen wir:

$$A = U_r \Sigma_r V_r^T = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \cdot \begin{bmatrix} 2 \end{bmatrix} \cdot \begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

- U_r ist 2×1
- Σ_r ist 1×1
- V_r^T ist 1×2

Diese kompakte Darstellung ist kleiner, effizienter - und bei grossen Matrizen besonders praktisch.

12.2 Hauptkomponentenanalyse (PCA)

Die Hauptkomponentenanalyse (Principal Component Analysis, PCA) ist eine Methode, um:

- hochdimensionale Daten zu vereinfachen
- wichtige Strukturen zu erkennen
- Daten zu komprimieren (weniger Dimensionen, wenig Informationsverlust)

PCA findet die Richtung mit der meisten Varianz - diese nennt man Hauptkomponenten. PCA sucht neue Achsen, auf denen die Daten:

- möglichst gut gestreckt sind (hohe Varianz)
- nicht redundant sind (unkorreliert)
- geordnet nach Wichtigkeit (grösste Varianz zuerst)

Die erste Hauptkomponente zeigt, wo die Daten am meisten streuen, die zweite ist orthogonal dazu, usw.

12.2.1 Mathematischer Ablauf (vereinfacht)

Normalerweise macht man die PCA mit einem Rechner. Hier wird der Ablauf vereinfacht erklärt:

Gegeben: n Datenpunkte $x_1, x_2, ..., x_n \in \Re^d$

1. Zentriere die Daten (Mittelwert abziehen)
2. Bilde die Kovarianzmatrix:

$$C = \frac{1}{n} X^T X$$

3. Berechne die Eigenvektoren und Eigenwerte der Kovarianzmatrix
4. Die Eigenvektoren sind die Hauptachsen (Hauptkomponenten)
5. Die zugehörigen Eigenwerte sagen, wie wichtig die Achse ist (Varianz)

12.2.2 Verbindung zur SVD

Die PCA basiert mathematisch auf der SVD der Datenmatrix:

$$X = U \Sigma V^T$$

- Die Spalten von V (bzw. V^T) sind die Hauptkomponenten
- Die Singulärwerte σ_i sagen, wie stark eine Komponente ist
- Die ersten k Spalten von V: neue Basis: Reduktion auf k Dimensionen

12.2.3 Dimension reduzieren

Wenn wir viele Daten in 20 Dimensionen haben, können wir mit PCA auf 2 oder 3 Hauptkomponenten reduzieren, ohne viel Informationen zu verlieren. Die Daten lassen sich dann darstellen als:

$$X_{\text{neu}} = X V_k$$

mit $V_k =$ Matrix der ersten k Hauptkomponenten.

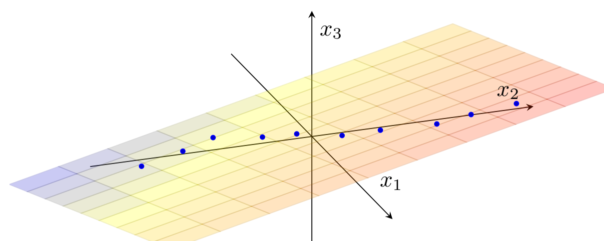
12.2.4 Beispiel

Wir haben Messdaten mit 3 Merkmalen: Grösse, Gewicht, Alter. PCA findet eine neue Achse, auf der diese Daten am besten auseinandergezogen sind - das ist die 1. Hauptkomponente. Die 2. Achse ist senkrecht dazu, fängt den nächstgrössten Unterschied ein, usw.

PCA sucht nun die Achse, entlang der:

- die meiste Streuung (Varianz) der Punkte liegt
- die Punkte am weitesten auseinander liegen
- man also am meisten Informationen gewinnt

Wir finden also eine neue Achse (Fläche) die eine Wolke von Datenpunkten am besten beschreibt. Diese neue Fläche verwenden wir als Achse. Somit haben wir zwei Dimensionen in eine zusammengefasst.



PCA-Ebene (reduzierter Raum)

12.3 Die Moore-Penrose-Pseudoinverse

Die normale Inverse A^{-1} gibt es nur, wenn A eine quadratische und reguläre Matrix ist (also vollen Rang hat). Aber in vielen Fällen ist A:

- nicht quadratisch ($m \times n, m \neq n$)
- oder nicht vollen Rangs

Trotzdem möchten wir eine art Inverse haben. Dafür gibt es die Moore-Penrose-Pseudoinverse.

12.3.1 Definition und Eigenschaften

Für eine beliebige Matrix $A \in \Re^{m \times n}$ existiert eine eindeutige Matrix $A^+ \in R^{n \times m}$ sodass:

$$\begin{aligned} AA^+A &= A \\ A^+AA^+ &= A^+ \\ (AA^+)^T &= AA^+ \\ (A^+A)^T &= A^+A \end{aligned}$$

Diese vier Bedingungen definieren die Pseudoinverse eindeutig.

12.3.2 Berechnung durch SVD

Wenn eine Matrix A gegeben ist:

$$A = U\Sigma V^T$$

Dann ist die Pseudoinverse:

$$A^+ = V\Sigma^+U^T$$

Dabei ist Σ^+ die Diagonalmatrix, in der auf der Diagonale die Kehrwerte der Singulärwerte von Σ stehen:

$$\Sigma^+ = diag(\frac{1}{\sigma_1}, \dots, \frac{1}{\sigma_r})$$

(alle weiteren Nullen bleiben Nullen).

12.3.3 Sonderfälle

Wenn A vollen Spaltenrang hat (z.B. bei linearer Regression):

$$A^+ = (A^TA)^{-1}A^T \quad \text{Links-Inverse}$$

Wenn A vollen Zeilenrang hat:

$$A^+ = A^T(AA^T)^{-1} \quad \text{Rechts-Inverses}$$

12.3.4 Beispiel

Gegeben sei:

$$A = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}$$

Das SVD ergibt:

$$A = U\Sigma V^T = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} \cdot \frac{1}{\sqrt{2}} \begin{bmatrix} -1 & -1 \\ -1 & 1 \end{bmatrix}$$

Dann ist A^+:

$$\begin{aligned} A^+ &= V\Sigma^+U^T \\ &= \frac{1}{\sqrt{2}} \begin{bmatrix} -1 & -1 \\ -1 & 1 \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 0 \end{bmatrix} \cdot \frac{1}{\sqrt{2}} \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix} \\ &= \frac{1}{4} \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix} \end{aligned}$$

12.3.5 Anwendung: Lineare Regression

In der lineares Regression:

$$X^TX\hat{b} = X^Ty \Rightarrow \hat{b} = (X^TX)^{-1}X^Ty$$

Aber wenn X^TX nicht invertierbar ist:

$$\hat{b} = (X^TX)^+X^Ty$$

Wir ersetzen die normale Inverse durch die Pseudoinverse.

13 Vektorräume

13.1 Vektorräume

Ein Vektorraum ist eine Sammlung von Vektoren. In einem Vektorraum gibt es immer den Nullvektor (einen Vektor mit nur Nullen). Beispiel:

Der Raum $\Re^2$ enthält Vektoren wie $\begin{pmatrix}3\\2\end{pmatrix}$ und $\begin{pmatrix}0\\0\end{pmatrix}$

Der Raum $\Re^3$ enthält Vektoren mit drei Werten.

13.1.1 Regeln in einem Vektorraum

Ein Vektorraum muss 8 Regeln (Axiome) einhalten. Diese Regeln sagen, wie Vektoren addiert oder mit Zahlen multipliziert werden dürfen:

V1: $\forall x, y \in V : x + y = y + x$

V2: $\forall x, y, z \in V : (x + y) + z = x + (y + z)$

V3: $\exists 0 \in V\ \forall x \in V : x + 0 = x$

V4: $\forall x \in V\ \exists - x \in V : x + (-x) = 0$

S1: $\forall \lambda, \mu \in S\ \forall x \in V : (\lambda \cdot \mu) \cdot x = \lambda \cdot (\mu \cdot x)$

S2: $\exists 1 \in S\ \forall x \in S : 1 \cdot x = x$

D1: $\forall \lambda \in S\ \forall x, y \in V : \lambda \cdot (x + y) = \lambda \cdot x + \lambda \cdot y$

D2: $\forall \lambda, \mu \in S\ \forall x \in V : (\lambda + \mu) \cdot x = \lambda \cdot x + \mu \cdot x$

13.1.2 Wichtige Arten von Vektorräumen

1. Matrizenräume:
   - Vektoren sind hier Matrizen
   - Beispiel: In $\Re^{2\times2}$ sind Vektoren z.B.:
   $$\begin{pmatrix}1 & 2\\3 & 4\end{pmatrix}$$
2. Funktionsräume:
   - Vektoren sind hier Funktionen, z.B. $f(x) = e^{-x}$
   - Man kann Funktionen addieren und mit Zahlen multiplizieren.
3. Symmetrische Matrizen:
   - Eine Matrix ist symmetrisch, wenn sie an der Diagonalen gespiegelt gleich aussieht.
   - Beispiel:
   $$\begin{pmatrix}7 & 5\\5 & -3\end{pmatrix}$$
   - Die Summe zweier symmetrischer Matrizen ist wieder symmetrisch. Sie bilden einen Untervektorraum.
4. Der kleinste Vektorraum Z:
   - Enthält nur den Nullvektor
   - Hat die Dimension 0

Errinerung: Der Nullvektor ist ein Vektor, bei dem alle Komponenten 0 sind. Jeder Vektorraum muss ihn enthalten!

13.1.3 Basis von Vektorräumen

Eine Basis ist eine kleine Sammlung von Vektoren, mit denen man alle andere Vektoren im Raum zusammensetzen kann - durch Linearkombination (z.B. $2 \cdot v_1 + 3 \cdot v_2$).

- Jeder Vektorraum hat eine Basis.
- Die Anzahl der Basis-Vektoren nennt man Dimension
- Alle Basen eines Raumes haben gleich viele Vektoren

Beispiele:

- Für $\Re^2$ ist die Standardbasis:
$$\left\{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right\}$$
- Im Matrizenraum $\Re^{2\times2}$ sind die Matrizen, bei denen nur ein Eintrag eine 1 ist, der Rest ist 0.

Die Dimensionen eines Vektorraums ist die Anzahl Vektoren in einer Basis.

13.2 Untervektorräume

Ein Untervektorraum ist eine Teilmenge eines Vektorraums, die selbst alle Bedingungen eines Vektorraums erfüllt.

Ein Untervektorraum ist ein kleinerer Vektorraum innerhalb eines grösseren.

13.2.1 Definition

Eine Menge $U \subseteq V$ ist ein Untervektorraum, wenn:

1. Der Nullvektor liegt in U
$$0 \in U$$
2. Für alle $v, w \in U$ gilt:
    - $v + w \in U$: abgeschlossen unter Addition
    - $c \cdot v \in U$ für alle Skalare $c \in \Re$: abgeschlossen unter skalrer Multiplikation

Ein Untervektorraum ist abgeschlossen unter Addition und Skalarmultiplikation und enthält den Nullvektor.

13.2.2 Beispiele

Typische Untervektorräume in $\Re^3$:

1. Eine Ebene durch den Nullpunkt $(0, 0, 0)$ ist ein Untervektorraum von $\Re^3$.
2. Eine Gerade durch den Nullpunkt ist auch ein Untervektorraum.
3. Der ganze Raum $\Re^3$ ist Untervektorraum von sich selbst.
4. Die Menge $\{(0, 0, 0)\}$ (nur der Nullvektor) ist auch ein Untervektorraum - der kleinste.

Ebenen oder Geraden, die den Nullpunkt nicht enthalten, sind keine Unterräume.

13.3 Vektorräume und lineare Gleichungssysteme

Wir wollen wissen:

- Wann ein lineares Gleichungssystem $Ax = b$ lösbar ist,
- und was das mit dem Spaltenraum zu tun hat.

Der Spaltenraum $C(A)$ ist:

- die Menge aller Linearkombinationen der Spalten einer Matrix A.
- Man schreibt ihn als $C(A)$
- Wenn $A \in \Re^{m \times n}$, dann ist $C(A) \subseteq \Re^m$

Das heisst: Der Spaltenraum ist ein Untervektorraum von $\Re^m$.

Die Gleichung $Ax = b$ hat genau dann eine Lösung, wenn b im Spaltenraum von A liegt, also:

$$b \in C(A)$$

Wenn b nicht im Spaltenraum liegt, dann hat das Gleichugnssystem keine Lösung.

13.3.1 Beispiel 1

Gegeben ist:

$$A = \begin{pmatrix} 1 & 0 \\ 4 & 3 \\ 2 & 3 \end{pmatrix}, \quad b = \begin{pmatrix} 4 \\ 25 \\ 17 \end{pmatrix}$$

Frage: Ist $Ax = b$ lösbar?

$$b = 4 \cdot \begin{pmatrix} 1 \\ 4 \\ 2 \end{pmatrix} + 3 \cdot \begin{pmatrix} 0 \\ 3 \\ 3 \end{pmatrix} \Rightarrow x = \begin{pmatrix} 4 \\ 3 \end{pmatrix}$$

13.4 Nullraum

Der Nullraum enthält alle Lösungen von $Ax = 0$. Man nennt ihn auch den Kern der Matrix A:

$$N(A) = \{x \in \Re^n | Ax = 0\}$$

Der Nullraum ist ein Untervektorraum von $\Re^n$. Wichtig zu wissen:

- Der Nullvektor ist immer im Nullraum (weil: $A0 = 0$).
- Wenn $v, w \in N(A)$, dann sind auch $v + w$ und λv im Nullraum.

13.4.1 Beispiel

Gegeben ist:

$$A = \begin{pmatrix} 1 & 1 & 2 \\ 2 & 1 & 3 \\ 3 & 1 & 4 \\ 4 & 1 & 5 \end{pmatrix}$$

Die Lösungen von $Ax = 0$:

$$x = c \cdot \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \quad c \in \Re$$

Das bedeutet: $N(A)$ ist eine Gerade in $\Re^3$, die durch $(1, 1, -1)^T$ verläuft.