

1 Singulärwertzerlegung und Pseudoinverse

Singulärwertzerlegung

A sei eine Matrix mit n Spalten und m Zeilen. Zunächst sei $n \leq m$.

- ① Bilde $B = A^T A$. Dies ist eine $n \times n$ -Matrix.
- ② Berechne die Eigenwerte von B . Diese sind nicht-negativ und werden in der Reihenfolge $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k > \lambda_{k+1} = \dots = \lambda_n = 0$ nummeriert. Als Kontrolle kann möglicherweise benutzt werden, dass k der Rang der Matrix A (und auch der Rang von $A^T A$) ist.
- ③ Bilde eine ON-Basis $\vec{v}_1, \dots, \vec{v}_n$ des \mathbb{R}^n . Dabei ist \vec{v}_i Eigenvektor zum Eigenwert λ_i . $V := [\vec{v}_1, \dots, \vec{v}_n]$ ist dann eine orthogonale Matrix ($V^T = V^{-1}$).
- ④ Die singulären Werte von A sind als $s_i = \sqrt{\lambda_i}$ definiert. Die Matrix $S = (s_{ij})$ ist eine Matrix vom Diagonaltyp, d.h. für $i \neq j$ ist $s_{ij} = 0$. S hat dieselbe Form wie A , also n Spalten und m Zeilen. Die Diagonalelemente sind durch die singulären Werte gegeben: $s_{ii} = s_i$.
- ⑤ Für $i \leq k$ definiere die Vektoren $\vec{u}_i = \frac{1}{\sqrt{\lambda_i}} A \vec{v}_i$. Diese Vektoren bilden ein Othonormalsystem. Ergänze diese Vektoren zu einer ON-Basis $\vec{u}_1, \dots, \vec{u}_m$ des \mathbb{R}^m und fasse sie in der Matrix $U = [\vec{u}_1, \dots, \vec{u}_m]$ zusammen.
- ⑥ Die Singulärwertzerlegung von A ist dann

$$A = USV^T.$$

singulären Werte
Diagonaltyp

Singulärwertzerlegung

Eigenschaften der Singulärwertzerlegung:

$$A = USV^T \Leftrightarrow A^T = VS^T U^T$$

Ist A invertierbar, so ist $A^{-1} = US^{-1}V^T$ (beachte: $U = U^T$ und $V^{-1} = V^T$).

Die erste Regel nutzt man aus, wenn $n > m$ ist (die Matrix ist breiter als hoch). Dann geht man so vor:

Setze $A_1 = A^T$ und zerlege $A_1 = U_1 S_1 V_1^T$. Die Zerlegung von A ist dann $A = USV^T$ mit $U = V_1$, $S = S_1^T$ und $V = U_1$.

Zerlegung von
 A^T

Für die meisten Rechnungen braucht man weder die Eigenvektoren von $A^T A$ zum Eigenwert $\lambda = 0$ noch die in ⑤ ergänzten Vektoren \vec{u}_{k+1} bis \vec{u}_m explizit.

Man erhält die Sparversion der Singulärwertzerlegung, wenn die Einträge in diesen Vektoren einfach durch $*$ ersetzt werden.

Sparversion

Beispiel 1: Singulärwertzerlegung von $A = \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 0 & 1 \end{bmatrix}$

- ① Da $n = 2 < m = 3$ ist, kann man direkt losrechnen: Es ist $B = A^\top A$.
- $$\begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 5 & 2 \\ 2 & 2 \end{bmatrix} = B$$

- ② Mit $p(\lambda) = \lambda^2 - 7\lambda + 6 = (\lambda - 6)(\lambda - 1)$ ist $\lambda_1 = 6$ und $\lambda_2 = 1$.

- ③ Eigenvektor zu $\lambda_1 = 6$: Es ist $B - 6E = \begin{bmatrix} -1 & 2 \\ 2 & -4 \end{bmatrix}$, also ist $\vec{v}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ Eigenvektor.

Entweder berechnet man analog $\vec{v}_2 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$ als Eigenvektor zu $\lambda_2 = 1$ oder man nutzt die Tatsache aus, dass die Eigenvektoren der symmetrischen Matrix B zu verschiedenen eigenwerten orthogonal sind.

V ist nun die Matrix der normierten Eigenvektoren: $V = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 & 1 \\ 1 & -2 \end{bmatrix}$

- ④ S enthält auf der Diagonalen die Wurzeln der Eigenwerte von B : $S = \begin{bmatrix} \sqrt{6} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$.

- ⑤ Es ist $\vec{u}_1 = \frac{1}{\sqrt{6}} A\vec{v}_1 = \frac{1}{\sqrt{30}} \begin{bmatrix} 2 \\ 5 \\ 1 \end{bmatrix}$ und $\vec{u}_2 = \frac{1}{\sqrt{1}} A\vec{v}_2 = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}$.

Um \vec{u}_1 und \vec{u}_2 zu einer ON-Basis des \mathbb{R}^3 zu ergänzen, kann man

- entweder das Gram-Schmidt-Verfahren verwenden:

\vec{u}_1 und \vec{u}_2 sind ja schon zueinander senkrecht und normiert, man lässt also das Orthonormalisierungsverfahren auf einen beliebigen davon linear unabhängigen Vektor des \mathbb{R}^3 los, z.B. auf \vec{e}_2 , weil dann das Skalarprodukt mit \vec{u}_2 schon Null ist:

$$\vec{w}_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} - \left\langle \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \frac{1}{\sqrt{30}} \begin{bmatrix} 2 \\ 5 \\ 1 \end{bmatrix} \right\rangle \frac{1}{\sqrt{30}} \begin{bmatrix} 2 \\ 5 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} - \frac{5}{30} \begin{bmatrix} 2 \\ 5 \\ 1 \end{bmatrix} = \frac{1}{6} \begin{bmatrix} -2 \\ 1 \\ -1 \end{bmatrix}$$

Man erhält $\vec{u}_3 = \frac{1}{\|\vec{w}_3\|} \vec{w}_3 = \frac{1}{\sqrt{6}} \begin{bmatrix} -2 \\ 1 \\ -1 \end{bmatrix}$.

- oder man nutzt die Tatsache aus, dass die Rechnung im \mathbb{R}^3 stattfindet und berechnet $\vec{u}_3 = \vec{u}_1 \times \vec{u}_2$.

Damit wird $U = \begin{bmatrix} \frac{2}{\sqrt{30}} & \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{6}} \\ \frac{5}{\sqrt{30}} & 0 & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{30}} & -\frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{6}} \end{bmatrix}$.

⑥ Die Singulärwertzerlegung von A ist

$$A = USV^T = \begin{bmatrix} \frac{2}{\sqrt{30}} & \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{6}} \\ \frac{5}{\sqrt{30}} & 0 & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{30}} & -\frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{6}} \end{bmatrix} \begin{bmatrix} \sqrt{6} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \end{bmatrix}.$$

Bei der Sparversion wird der rechenaufwändige Basisergänzungsschritt in ⑤ weggelassen und man hat

$$A = USV^T = \begin{bmatrix} \frac{2}{\sqrt{30}} & \frac{1}{\sqrt{5}} & * \\ \frac{5}{\sqrt{30}} & 0 & * \\ \frac{1}{\sqrt{30}} & -\frac{2}{\sqrt{5}} & * \end{bmatrix} \begin{bmatrix} \sqrt{6} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \end{bmatrix}.$$

Man sieht dass beim Ausmultiplizieren die *-Werte gerade mit den Nullen der S_1 -Matrix multipliziert werden.

Beispiel 2: Singulärwertzerlegung von $A = \begin{bmatrix} 2 & 2 & 1 \end{bmatrix}$

Da die Matrix A breiter als hoch ist, wird eine Zerlegung von $A_1 = A^T$ vorgenommen.

- ① Man sieht $B = A_1^T A_1 = A A^T = [9]$.
- ② $\lambda_1 = 9$ ist der einzige Eigenwert dieser 1×1 -Matrix.
- ③ $\vec{v}_1 = [1]$ ist eine ONB des \mathbb{R}^1 , $V_1 = [1]$.
- ④ $S_1 = \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}$.

⑤ $\vec{u}_1 = \frac{1}{3}A_1\vec{v}_1 = \frac{1}{3} \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}$. Statt \vec{u}_1 mit dem Gram-Schmidt-Verfahren zu einer

ON-Basis des \mathbb{R}^3 zu ergänzen, benutzen wir wieder das Kreuzprodukt: sicher steht $\vec{w}_2 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$ auf \vec{u}_1 senkrecht, und $\vec{w}_3 = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix} \times \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ -4 \end{bmatrix}$ steht auf beiden senkrecht. Damit muss man \vec{w}_2 und \vec{w}_3 nur noch normieren:

$$U_1 = [\vec{u}_1 \quad \vec{u}_2 \quad \vec{u}_3] = \begin{bmatrix} \frac{2}{3} & \frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} \\ \frac{2}{3} & -\frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} \\ \frac{1}{3} & 0 & -\frac{4}{3\sqrt{2}} \end{bmatrix}$$

⑥ Damit ist $A_1 = U_1 S_1 V_1^\top = \begin{bmatrix} \frac{2}{3} & \frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} \\ \frac{2}{3} & -\frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} \\ \frac{1}{3} & 0 & -\frac{4}{3\sqrt{2}} \end{bmatrix} \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix} [1]$

$$\text{und } A = A_1^\top = V_1 S_1^\top U_1^\top = [1] \begin{bmatrix} 3 & 0 & 0 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{3\sqrt{2}} & \frac{1}{3\sqrt{2}} & -\frac{4}{3\sqrt{2}} \end{bmatrix}$$

In der Sparversion hat man

⑤ $\vec{u}_1 = \frac{1}{3} \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}$, $\vec{u}_2 = \begin{bmatrix} * \\ * \\ * \end{bmatrix}$, $\vec{u}_3 = \begin{bmatrix} * \\ * \\ * \end{bmatrix}$ und

$$U_1 = \begin{bmatrix} \frac{2}{3} & * & * \\ \frac{2}{3} & * & * \\ \frac{1}{3} & * & * \end{bmatrix} \text{ und damit } A_1 = U_1 S_1 V_1^\top = \begin{bmatrix} \frac{2}{3} & * & * \\ \frac{2}{3} & * & * \\ \frac{1}{3} & * & * \end{bmatrix} \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix} [1]$$

Man kann die Singulärwertzerlegung von A natürlich auch direkt berechnen:

① Es ist $B = A^\top A = \begin{bmatrix} 4 & 4 & 2 \\ 4 & 4 & 2 \\ 2 & 2 & 1 \end{bmatrix}$.

② Hier sieht man, wo der größere Aufwand steckt: das charakteristische Polynom ist $p(\lambda) = -\lambda^3 + 9\lambda^2$, also $\lambda_1 = 9, \lambda_2 = \lambda_3 = 0$.

③ Mit etwas Mühe und dem Gauss-Algorithmus berechnet man $\vec{v}_1 = \frac{1}{3} \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}$ als

Eigenvektor zu $\lambda_1 = 9$ und genau wie oben z.B. $\vec{v}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$ und $\vec{v}_3 = \frac{1}{3\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ -4 \end{bmatrix}$ als Eigenvektoren zu $\lambda_2 = \lambda_3 = 0$.

④ In S stehen wieder die Wurzeln der Eigenwerte: $S = \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}$.

⑤ U ist diesmal die 1×1 -Matrix, die nur aus $\vec{u}_1 = \frac{1}{3}A\vec{v}_1 = \frac{1}{3}[3] = [1]$ besteht.

⑥ $A = USV^\top$ hat damit dieselben Komponenten wie oben.

Pseudoinverse, Moore-Penrose-Inverse, Pseudonormallösung

Ist $S = (s_{ij})$ eine $n \times m$ -Matrix vom Diagonaltyp, so definiert man die $m \times n$ -Matrix

$S^+ = (s_{ji}^+)$ durch $s_{ij}^+ = \begin{cases} \frac{1}{s_{ii}} & \text{falls } s_{ii} \neq 0 \\ 0 & \text{sonst} \end{cases}$. Das heisst, dass S^+ aus S dadurch

entsteht, dass S transponiert wird und die Nicht-Null-Elemente auf der Diagonalen durch ihre Kehrwerte ersetzt werden.

Ist $A = USV^\top$, so ist die Moore-Penrose-Inverse oder Pseudoinverse A^+ von A durch $A^+ = VS^+U^\top$ definiert.

$$A = USV^\top \Rightarrow A^+ = VS^+U^\top$$

Hier reicht die Sparversion der Singulärwertzerlegung.

$$(A^+)^\top = (A^\top)^+, \quad A^{-1} \text{ existiert} \Rightarrow A^{-1} = A^+$$

Achtung: im allg. ist $(AB)^+ \neq B^+A^+$.

Ist $A\vec{x} = \vec{b}$ ein (möglicherweise mehrdeutig lösbares oder unlösbares) lineares Gleichungssystem, so ist die Pseudonormallösung \vec{x}^+ definiert als $\vec{x}^+ = A^+\vec{b}$.

\vec{x}^+ ist dadurch charakterisiert, dass gilt:

(i) Die Norm des Fehlers $A\vec{x}^+ - \vec{b}$ ist minimal und

(ii) die Norm von \vec{x}^+ ist minimal.

Insbesondere ist im Fall eine eindeutig lösbar Systems \vec{x}^+ die Lösung und in mehrdeutig lösbar Systemen die Lösung mit der kleinsten Norm.

Ein wichtiger Spezialfall ist:

Ist A injektiv (d.h. $\text{rang } A = n$), so sind A^+ und \vec{x}^+ ohne Singulärwertzerlegung bestimmbar:

$$A^+ = (A^\top A)^{-1} A^\top \quad \text{und} \quad \vec{x}^+ \text{ ist eindeutige Lösung von } A^\top A \vec{x}^+ = A^\top \vec{b}.$$

Beispiel 3: Pseudoinverse von $A = \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 0 & 1 \end{bmatrix}$ und Pseudonormallösung von $A\vec{x} =$

$$\begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}$$

Da der Rang von A zwei ist, kann man die Formel oben anwenden:

$$B = A^\top A = \begin{bmatrix} 5 & 2 \\ 2 & 2 \end{bmatrix} \Rightarrow B^{-1} = \frac{1}{6} \begin{bmatrix} 2 & -2 \\ -2 & 5 \end{bmatrix} \Rightarrow A^+ = B^{-1} A^\top = \frac{1}{6} \begin{bmatrix} 2 & 2 & -2 \\ -2 & 1 & 5 \end{bmatrix}$$

Damit berechnet man sofort die Pseudonormallösung

$$\vec{x}^+ = A^+ \vec{b} = \frac{1}{2} \begin{bmatrix} 2 & 2 & -2 \\ -2 & 1 & 5 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Wenn man die Pseudoinverse nicht benötigt, berechnet man noch schneller $A^\top \vec{b} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$ und \vec{x}^+ als Lösung von

$$\begin{bmatrix} 5 & 2 \\ 2 & 2 \end{bmatrix} \vec{x}^+ = \begin{bmatrix} 5 \\ 2 \end{bmatrix}.$$

Die Berechnung von $A^+ = VS^+U^\top$ über die Singulärwertzerlegung bzw. die Sparversion ist mühsamer (liefert aber natürlich dasselbe Ergebnis):

$$A^+ = \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ 1 & -\frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{6}} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{2}{\sqrt{30}} & \frac{5}{\sqrt{30}} & \frac{1}{\sqrt{30}} \\ 1 & 0 & -\frac{2}{\sqrt{5}} \\ * & * & * \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{1}{6} & \frac{5}{6} \end{bmatrix}$$

Beispiel 4: Pseudonormallösung von $\begin{bmatrix} 2 & 2 & 1 \end{bmatrix} \vec{x} = [4]$

Da die Matrix nicht Rang 3 hat, kann man die obige Formel nicht direkt verwenden.

Allerdings hat $A^\top = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}$ den Rang 1 und deshalb ist

$$(A^\top)^+ = \left([2 \ 2 \ 1] \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix} \right)^{-1} \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix} = \frac{1}{9} [2 \ 2 \ 1].$$

Mit Hilfe der Formel $(A^\top)^+ = (A^+)^{\top}$, also $A^+ = ((A^\top)^+)^{\top}$ folgt

$$A^+ = \frac{1}{9} \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix} \text{ und } \vec{x}^+ = A^+ \vec{b} = \frac{1}{9} \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix} [4] = \frac{4}{9} \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}.$$

Anwendung: Numerischer Rangabfall

Von numerischem Rangabfall spricht man, wenn eine Matrix A sehr nahe an einer mit kleinerem Rang liegt. Dies führt dazu, dass die Konditionszahl ($\|A\| \|A^{-1}\|$) sehr groß wird. Kleine Abweichungen in den Daten des linearen Gleichungssystems $A\vec{x} = \vec{b}$ führen zu großen Abweichungen in der Lösung \vec{x} .

<p>Beispiel 5: $A\vec{x} = \vec{b}$ mit $A = \begin{bmatrix} 2 & 3 & 4 \\ 2 & 3 & 4.001 \\ 3 & 4 & 5 \end{bmatrix}$ und $\vec{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$</p>

Man sieht leicht (oder rechnet nach), dass A invertierbar ist. Die Lösung \vec{x} ist also eindeutig bestimmt:

Die exakte Lösung ist $\vec{x} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$. Allerdings ist z.B. $\vec{y} = \begin{bmatrix} -0.5 \\ 0 \\ 0.5 \end{bmatrix}$ eine fast „genau so gute“ Lösung: Es ist $A\vec{y} = \begin{bmatrix} 1 \\ 1.0005 \\ 1 \end{bmatrix}$.

Wenn man die Lösung des leicht gestörten Systems $A\vec{x}_1 = \vec{b}_1$ mit $\vec{b}_1 = \begin{bmatrix} 1 \\ 0.9 \\ 1 \end{bmatrix}$ berech-

net, erhält man $\vec{x}_1 = \begin{bmatrix} -101.0000 \\ 201.0000 \\ -100.0000 \end{bmatrix}$

Der Schlüssel zu dieser Problematik liegt in der Singulärwertzerlegung von A : Es ist

$$A = USV^\top \text{ mit } S = \begin{bmatrix} 10.3873 & 0 & 0 \\ 0 & 0.3338 & 0 \\ 0 & 0 & 0.0003 \end{bmatrix} \text{ und unitären Matrizen } U \text{ und } V.$$

Man sieht, dass A einen sehr kleinen singulären Wert hat. Die singulären Werte von A^{-1} sind die Kehrwerte der singulären Werte von A und der größte davon ist ≈ 3467 .

In solchen Fällen kann man so vorgehen:

- ① Zerlege $A = USV^T$.
- ② Die Matrix S_1 entsteht, indem in S alle singulären Werte, die kleiner als eine vorgegebene Zahl $\varepsilon > 0$ sind, durch Null ersetzt werden.
- ③ Statt der Lösungen von $A\vec{x} = \vec{b}$ bestimme die Pseudonormallösungen von $A_1\vec{x} = \vec{b}$ durch

$$\vec{x}^+ = A_1^+ \vec{b} = VS_1^+ U^T \vec{b}$$

In diesem Beispiel ersetzt man den dritten singulären Wert durch Null und erhält

$$S_1 = \begin{bmatrix} 10.3873 & 0 & 0 \\ 0 & 0.3338 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{und} \quad S_1^+ = \begin{bmatrix} 0.0963 & 0 & 0 \\ 0 & 2.9961 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Damit wird

$$A_1^+ = \begin{bmatrix} -1.1633 & -1.1674 & 1.8314 \\ -0.1669 & -0.1676 & 0.3342 \\ 0.8316 & 0.8344 & -1.1662 \end{bmatrix}$$

$$\text{und } \vec{x}^+ = A_1^+ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -0.4992 \\ -0.0002 \\ 0.4997 \end{bmatrix} \quad \text{und} \quad \vec{x}_1^+ = A_1^+ \begin{bmatrix} 1 \\ 0.9 \\ 1 \end{bmatrix} = \begin{bmatrix} -0.3825 \\ 0.0165 \\ 0.4163 \end{bmatrix}$$

Anwendung: Ausgleichsrechnung

Andere Bezeichnungen: Gaußsche Methode der kleinsten Quadrate, Ausgleichspolynome, Ausgleichsgerade, Regressionsgerade

Ausgleichsgerade

Wichtigster Fall: Ausgleichsgerade

Gegeben sind $n > 2$ Punktepaare (x_i, y_i) , wobei mindestens zwei verschiedene x -Koordinaten auftreten sollen. Gesucht ist eine Gerade $y = ax + b$, die diese Punkte bestmöglich annähert, d.h. Zahlen a und b , so dass der quadratische Fehler

$$\sum_{i=1}^n \left((ax_i + b) - y_i \right)^2 \text{ möglichst klein wird.}$$

Die Lösung dieses Problems wird gegeben durch die Pseudonormallösung des folgenden Gleichungssystems:

$$\begin{array}{l} b + ax_1 = y_1 \\ \vdots \\ b + ax_n = y_n \end{array}, \quad \text{also } A \begin{bmatrix} b \\ a \end{bmatrix} = \vec{y} \quad \text{mit} \quad A = \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \quad \text{und} \quad \vec{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

Da die Matrix A injektiv ist, läßt sich die Pseudonormallösung mit Hilfe der Transponierten berechnen: $\begin{bmatrix} b \\ a \end{bmatrix} = (A^\top A)^{-1} A^\top \vec{y}$.

Der Korrelationskoeffizient r ist ein Maß für die Güte der Approximation durch die Ausgleichsgerade. Stets ist $|r| \leq 1$, und je näher $|r|$ bei 1 ist, desto stärker ist die (lineare) Abhängigkeit der y - von den x -Werten.

Korrelationskoeffizient

Berechnung

Die Berechnung beruht auf der Anwendung der Cramerschen Regel auf das Gleichungssystem $A^\top A \begin{bmatrix} b \\ a \end{bmatrix} = A^\top \vec{y}$.

Der Einfachheit halber sind die Summationsgrenzen weggelassen. Alle Summen laufen von $i = 1$ bis n .

① Berechne $\Delta = n \sum x_i^2 - (\sum x_i)^2$

② Die Ausgleichsgerade $y = ax + b$ hat die Koeffizienten

$$a = \frac{1}{\Delta} (n \sum x_i y_i - \sum x_i \sum y_i)$$

und $b = \frac{1}{\Delta} (\sum x_i^2 \sum y_i - \sum x_i y_i \sum x_i)$.

③ Es ist $r = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{(n \sum x_i^2 - (\sum x_i)^2)(n \sum y_i^2 - (\sum y_i)^2)}}$

Praktische Durchführung: Die Werte werden in eine Tabelle eingetragen, die von außen nach innen ausgefüllt wird.

Beispiel 6: $n = 5$ mit den Wertepaaren $(0, 0)$, $(2, 3)$, $(2, 5)$, $(4, 4)$ und $(5, 6)$.

Die Starttabelle wird ausgefüllt. In die letzte Zeile werden die Summen eingetragen. Eine Spalte mit y_i^2 braucht man nur, wenn auch der Korrelationskoeffizient bestimmt werden soll.

x_i	x_i^2	$x_i y_i$	y_i^2	y_i		x_i	x_i^2	$x_i y_i$	y_i^2	y_i
0				0		0	0	0	0	0
2				3		2	4	6	9	3
2				5	wird zu	2	4	10	25	5
4				4		4	16	16	16	4
5				6		5	25	30	36	6
						13	49	62	86	18

Damit wird

$$\Delta = 5 \cdot 49 - 13^2 = 245 - 169 = 76$$

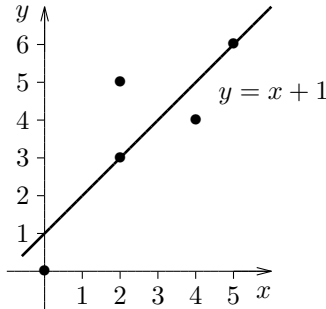
$$a = \frac{1}{76}(5 \cdot 62 - 13 \cdot 18) = \frac{1}{76}(310 - 234) = 1 \quad \text{und}$$

$$b = \frac{1}{76}(49 \cdot 18 - 62 \cdot 13) = \frac{1}{76}(882 - 806) = 1$$

Die Ausgleichsgerade ist also $y = x + 1$.

Der Korrelationskoeffizient ist

$$r = \frac{5 \cdot 62 - 13 \cdot 18}{\sqrt{(5 \cdot 49 - 13^2)(5 \cdot 86 - 18^2)}} = \frac{76}{\sqrt{76 \cdot 106}} \approx 0.847.$$



Allgemeines Grundproblem der Ausgleichsrechnung

Gegeben sind n Datenpaare (x_i, y_i) , $i = 1, \dots, n$ und $k < n$ Funktionen f_1 bis f_k .

Allgemeines
Grundproblem

Gesucht ist eine Linearkombination $f(x) = \sum_{j=1}^k \alpha_j f_j(x)$ dieser f_j , so dass die Summe der quadrierten Abweichungen von f an den Stellen x_i zu y_i minimal wird:

$$F = \sum_{i=1}^n (f(x_i) - y_i)^2 = \sum_{i=1}^n \left(\sum_{j=1}^k \alpha_j f_j(x_i) - y_i \right)^2 \stackrel{!}{=} \text{Min.}$$

Die Lösung dieses Problems wird durch die Pseudonormallösung des folgenden Gleichungssystems gegeben:

Löse $A\vec{a} = \vec{y}$. Dabei enthält $\vec{a} = (\alpha_1, \dots, \alpha_k)^\top$ die gesuchten Koeffizienten,

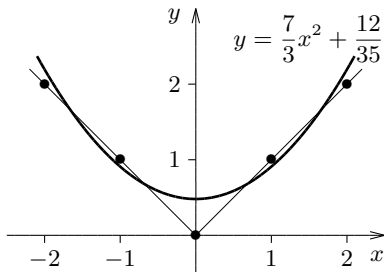
$$A = \begin{bmatrix} f_1(x_1) & f_2(x_1) & \cdots & f_k(x_1) \\ f_1(x_2) & f_2(x_2) & \cdots & f_k(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(x_n) & f_2(x_n) & \cdots & f_k(x_n) \end{bmatrix} \quad \text{und} \quad \vec{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}.$$

Beispiel 7: Approximation von $|x|$ an den Stellen $-2, -1, 0, 1$ und 2 durch eine Funktion der Form $f(x) = \alpha_1 + \alpha_2 x^2$

Es ist $f_1(x) = 1$ und $f_2(x) = x^2$.

Gesucht ist die Pseudonormallösung von

$$\begin{array}{rcl} \alpha_1 + 4\alpha_2 & = & 2 \quad (x = -2) \\ \alpha_1 + \alpha_2 & = & 1 \quad (x = -1) \\ \alpha_1 & = & 1 \quad (x = 0) \\ \alpha_1 + \alpha_2 & = & 1 \quad (x = 1) \\ \alpha_1 + 4\alpha_2 & = & 2 \quad (x = 2) \end{array}$$



Mit $A = \begin{bmatrix} 1 & 4 \\ 1 & 1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 4 \end{bmatrix}$ und $\vec{b} = \begin{bmatrix} 2 \\ 1 \\ 0 \\ 1 \\ 2 \end{bmatrix}$ berechnet man $A^\top A = \begin{bmatrix} 5 & 10 \\ 10 & 34 \end{bmatrix}$ $A^\top \vec{b} = \begin{bmatrix} 6 \\ 18 \end{bmatrix}$ und damit $\alpha_1 = \frac{24}{70} = \frac{12}{35}$ und $\alpha_2 = \frac{30}{70} = \frac{3}{7}$. Das Ausgleichspolynom ist also $y = \frac{3}{7}x^2 + \frac{12}{35}$.

Linearisierung und Gewichte

Die Ausgleichsrechnung (insbesondere Die Ausgleichsgerade) lässt sich mit den obigen Methoden nur finden, wenn die unbekannt Parameter α_1 bis α_k als Koeffizienten einer Linearkombination auftreten. Einige andere Fälle lassen sich durch Linearisierung darauf zurückführen:

Linearisierung

- Typ $y = ae^{bx}$
 Mit $z = \ln y$ erhält man $z = \ln a + bx$. Man ermittelt also eine Ausgleichsgerade $z = c + dx$ für die $(x_i, z_i) = (x_i, \ln y_i)$ -Punkte und findet so Werte für $a = e^c$ und $b = d$.
 Dieser Fall enthält auch Funktionen des Typs $y = ab^{cx}$, da sich dies wegen $b^{cx} = e^{c \ln bx}$ wie oben schreiben lässt.
- Typ $y = cx^d$
 Wegen $x^d = e^{d \ln x}$ ergibt Logarithmieren von x und y die Beziehung $\ln y = \ln c + d \ln x$. Hat also die Ausgleichsgerade für die Wertepaare $(w_i, z_i) = (\ln x_i, \ln y_i)$ die Form $z = \alpha + \beta w$, so ist die Ausgleichsfunktion der Originaldaten $y = e^\alpha x^\beta$.
- Typ $y = \frac{1}{ax + b}$
 Für $z = \frac{1}{y}$ sucht man die Ausgleichsgerade für die Wertepaare (x_i, z_i) in der Form $z = ax + b$. Die Ausgleichsfunktion ist dann $y = \frac{1}{ax + b}$.

Ein gewichtetes Ausgleichsproblem liegt vor, wenn die Quadrate von $(ax_i + b) - y_i$ nicht gleichgewichtet eingehen, sondern wenn das Minimum einer Funktion der Form

gewichtetes Ausgleichsproblem

$$F = \sum_{i=1}^n \gamma_i ((ax_i + b) - y_i)^2 \quad (\gamma_i > 0)$$

gesucht wird. Die Zahlen γ_i sind die Gewichte.

Gewichte

- ① Ist $A\vec{x} = \vec{y}$ das Gleichungssystem des ungewichteten Ausgleichsproblems, so entstehen A_1 und \vec{y}_1 dadurch, dass für $k = 1$ bis n die k -te Zeile von A und der k -te Eintrag in \vec{y} mit γ_k multipliziert werden.
- ② Die Lösung des gewichteten Systems ist die Pseudonormallösung von $(A^\top A_1)\vec{x} = A^\top \vec{y}_1$. In vielen Fällen ist $A^\top A_1$ invertierbar und das Problem hat eine eindeutige Lösung.

Beispiel 8: Die Ausgleichsfunktion vom Typ $y = \frac{1}{ax + b}$ zu den Punkten

$$(x_1, y_2) = (0, 1), (x_2, y_2) = (2, 1) \text{ und } (x_3, y_3) = (3, \frac{1}{4}).$$

- ① Rückführung auf Ausgleichsgerade: mit $z = \frac{1}{y}$ ist $z_1 = z_2 = 1, z_3 = 4$.

- ② Das Gleichungssystem $A \begin{bmatrix} b \\ a \end{bmatrix} = \vec{z}$ ist $\begin{bmatrix} 1 & 0 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix}$.

- ③ Mit der Tabelle erhält man

x_i	x_i^2	$x_i z_i$	z_i	
0	0	0	1	also $\Delta = 3 \cdot 13 - 5^4 = 14$
2	4	2	1	
3	9	12	4	
5	13	14	6	

$$\alpha_1 = \frac{13 \cdot 6 - 14 \cdot 5}{14} = \frac{8}{14} = \frac{4}{7}, \alpha_2 = \frac{3 \cdot 14 - 5 \cdot 6}{14} = \frac{12}{14} = \frac{6}{7}$$

Die Ausgleichsfunktion des Ausgangsproblems ist also $y = \frac{1}{z} = \frac{7}{6x + 4}$.

Beispiel 9: Die Ausgleichsgerade zu den Punkte $(0, 1)$, $(2, 1)$ und $(3, 4)$ mit und ohne den Gewichten $\gamma_1 = \gamma_3 = 1$ und $\gamma_2 = 3$.

Die Ausgleichsgerade für den Fall ohne Gewichte ist in Beispiel 7 schon berechnet worden: $y = \frac{6}{7}x + \frac{4}{7}$.

- ① Die Matrizen erhält man durch Abwandlung des Systems aus Beispiel 8:

$$A_1 \begin{bmatrix} b \\ a \end{bmatrix} = \vec{y}_1 \Leftrightarrow \begin{bmatrix} 1 & 0 \\ 3 & 6 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix}.$$

- ② Es ist $A^T A_1 = \begin{bmatrix} 5 & 9 \\ 9 & 21 \end{bmatrix}$ und $A^T \vec{y}_1 = \begin{bmatrix} 8 \\ 18 \end{bmatrix}$. Daraus liest man die Lösung $b = \frac{6}{24} = \frac{1}{4}$ und $a = \frac{18}{24} = \frac{3}{4}$ ab. Die Ausgleichsgerade des gewichteten Problems ist also $y = \frac{3}{4}x + \frac{1}{4}$.

Warnungen:

- (i) Die Matrix $A^T A_1$ in Beispiel 9 ist nur zufällig symmetrisch, i.allg. ist das nicht so.
- (ii) Die in Beispiel 8 berechnete Ausgleichsfunktion löst nicht das nichtlineare Problem $\sum_{i=1}^3 \gamma_i \left(\frac{1}{ax_i + b} - y_i \right)^2 \stackrel{!}{=} \text{Min}$, das mit Hilfe der mehrdimensionalen Differentialrechnung bearbeitet werden kann. Die Lösung ist ungefähr $y = \frac{1}{0.27x + 0.96}$ und ist in der Skizze dünn gestrichelt.

