Wissenschaftliches Rechnen - Großübung 2.2

Themen: Ausgleichsrechnung, Definitheit

Ugo & Gabriel

22. November 2022

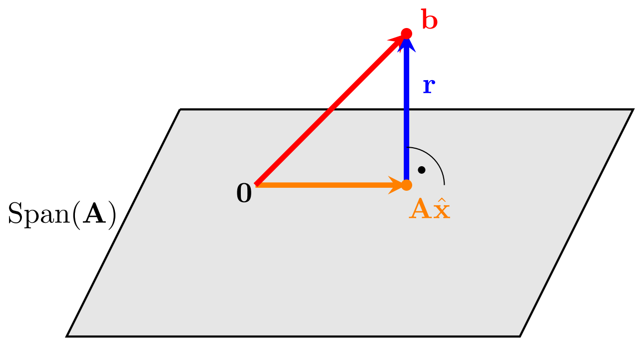
Aufgabe 1: Ausgleichsrechnung

Bei der linearen Ausgleichsrechnung versucht man ein unlösbares, überbestimmtes LGS approximativ zu lösen. Statt $\mathbf{Ax} = \mathbf{b}$ löst man $\mathbf{A}^\mathsf{T}\mathbf{Ax} = \mathbf{A}^\mathsf{T}\mathbf{b}$ und minimiert dabei die ℓ^2-Norm des Residuums $\mathbf{r} = \mathbf{Ax} - \mathbf{b}$, sodass $\mathbf{Ax}$ möglichst nah an $\mathbf{b}$ sein muss.

1. Welche geometrische Interpretation hat die Normalengleichung?

   Lösung

   Es wird die Lösung gefunden, die erreicht wird, wenn man den Lösungsvektor auf den durch die Systemmatrix aufgespannten Raum orthogonal projiziert.

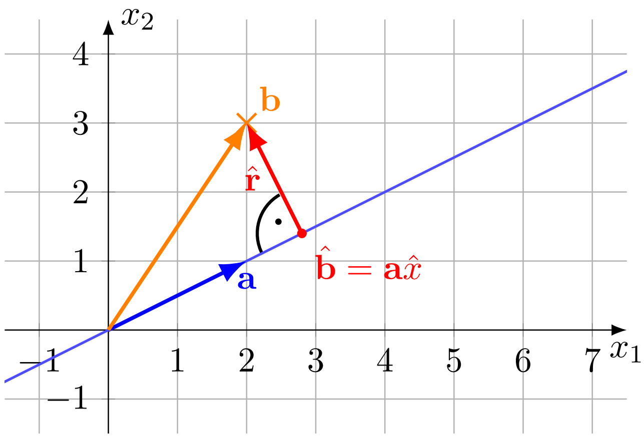
   

   Falls $\mathbf{b}$ bereits im Spaltenraum von $\mathbf{A}$ liegt, dann wird er auf sich selbst projiziert (ist also ein Fixpunkt). Fall nicht, dann wird er auf den nächstgelegenen Punkt bezüglich der ℓ^2-Metrik abgebildet.

   Lösung Ende

2. Zunächst betrachten wir den einfachen Fall, dass die Systemmatrix ein einziger Vektor $\mathbf{a}$ ist. Dann haben wir ein lineares Gleichungssystem $\mathbf{a}x = \mathbf{b}$ mit $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$. Welcher Punkt $\hat{\mathbf{b}}$ im linearen Unterraum, der von $\mathbf{a}$ aufgespannt wird, ist am nächsten zu $\mathbf{b}$?

   Hinweis: Schauen Sie sich die Übungsaufgabe zu Skalarprodukten an.



Lösung

Wir wissen, dass für das Skalarprodukt folgendes gilt: $\langle \mathbf{a}, \mathbf{b} \rangle = \|\mathbf{a}\|\|\mathbf{b}\| \cos(\varphi)$. Dabei ist $\|\mathbf{b}\| \cos(\varphi) = \frac{\langle \mathbf{a},\mathbf{b} \rangle}{\|\mathbf{a}\|}$ die orthogonale Projektion von $\mathbf{b}$ auf den Spaltenraum von $\mathbf{a}$, also die Distanz vom Ursprung bis zu dem Punkt $\hat{\mathbf{b}}$, welcher im Spann von $\mathbf{a}$ liegt und zu $\mathbf{b}$ am nächsten ist. Um diese Distanz müssen wir den auf Länge 1 normierten Vektor $\frac{\mathbf{a}}{\|\mathbf{a}\|}$ skalieren.

$$\hat{\mathbf{b}} = \frac{\mathbf{a}^\intercal \mathbf{b}}{\|\mathbf{a}\|^2}\mathbf{a}$$

Lösung Ende

3. Gegeben ein überbestimmtes LGS $\mathbf{Ax} = \mathbf{b}$. Geben Sie eine Formel für den nächsten Punkt $\hat{\mathbf{b}}$ an, welcher sich im Spann der Matrix $\mathbf{A}$ befindet.

Lösung

Da gilt $\mathbf{A}^\intercal \mathbf{A} \mathbf{x} = \mathbf{A}^\intercal \mathbf{b}$ folgt direkt $\mathbf{x} = \left(\mathbf{A}^\intercal \mathbf{A}\right)^{-1} \mathbf{A}^\intercal \mathbf{b}$. Da wir aber wissen wollen, worauf $\mathbf{x}$ abgebildet wird, lautet die Lösung $\hat{\mathbf{b}} = \mathbf{Ax} = \mathbf{A}\left(\mathbf{A}^\intercal \mathbf{A}\right)^{-1} \mathbf{A}^\intercal \mathbf{b}$. Es ergibt sich also der *Projektor* $\mathbf{P} = \mathbf{A}\left(\mathbf{A}^\intercal \mathbf{A}\right)^{-1} \mathbf{A}^\intercal$. Dieser führt die orthogonale Projektion durch.

Lösung Ende

4. Die letzte Aufgabe lässt sich als lineares Gleichungssystem $\mathbf{Pb} = \hat{\mathbf{b}}$ schreiben. Zeigen Sie, dass diese Matrix $\mathbf{P}$ eine Projektion ist, d.h. $\mathbf{P}^2 = \mathbf{PP} = \mathbf{P}$ (diese Eigenschaft nennt sich idempotent).

Lösung

Da $\mathbf{P} = \mathbf{A}\left(\mathbf{A}^\intercal \mathbf{A}\right)^{-1} \mathbf{A}^\intercal$, ergibt sich:

$$\mathbf{P}^2 = \mathbf{A}\left(\mathbf{A}^\intercal \mathbf{A}\right)^{-1} \cancel{\mathbf{A}^\intercal \mathbf{A}\left(\mathbf{A}^\intercal \mathbf{A}\right)^{-1}} \mathbf{A}^\intercal = \mathbf{A}\left(\mathbf{A}^\intercal \mathbf{A}\right)^{-1} \mathbf{A}^\intercal = \mathbf{P}$$

Lösung Ende

5. Gegeben sei die folgende Basis $\mathbf{A}$ eines linearen Unterraums des $\mathbb{R}^3$:

$$\mathbf{A} = \begin{bmatrix} 3 & 1 \\ 1 & 0 \\ 0 & 0 \end{bmatrix}$$

Berechnen Sie die Projektionsmatrix $\mathbf{P}$, welche jeden Punkt $\mathbf{x} \in \mathbb{R}^3$ auf den nächstgelegenen Punkt $\hat{\mathbf{x}} \in \mathrm{Span}(\mathbf{A})$ projiziert. Ist das Ergebnis überraschend?

Lösung

In den Rechnungen nutzen wir folgende Identität:

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

Damit gilt:

$$\begin{aligned}
\mathbf{P} &= \begin{bmatrix} 3 & 1 \\ 1 & 0 \\ 0 & 0 \end{bmatrix} \left(\begin{bmatrix} 3 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 1 & 0 \\ 0 & 0 \end{bmatrix} \right)^{-1} \begin{bmatrix} 3 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \\
&= \begin{bmatrix} 3 & 1 \\ 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 10 & 3 \\ 3 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 3 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \\
&= \begin{bmatrix} 3 & 1 \\ 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & -3 \\ -3 & 10 \end{bmatrix} \begin{bmatrix} 3 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \\
&= \begin{bmatrix} 3 & 1 \\ 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 1 & -3 & 0 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}
\end{aligned}$$

Die Lösung ist nicht überraschend, da der von $\mathbf{A}$ aufgespannte lineare Unterraum die x-y-Ebene ist und die Projektion dadurch zustande kommt, dass man die z-Komponente auf 0 setzt.

Lösung Ende

Aufgabe 2: Definitheit

Die Definitheit einer Matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ ist über die quadratische Form definiert: Wir sagen $\mathbf{A}$ ist

$$\left.\begin{array}{ll}\text{positiv definit,} & \text{falls } \mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x} > 0 \\ \text{positiv semidefinit,} & \text{falls } \mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x} \geq 0 \\ \text{negativ definit,} & \text{falls } \mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x} < 0 \\ \text{negativ semidefinit,} & \text{falls } \mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x} \leq 0 \\ \text{indefinit,} & \text{sonst}\end{array}\right\} \text{ für alle } \mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}.$$

1. Was genau ist die quadratische Form $\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x}$? Gib die quadratische Form folgender Matrizen an:

$$\mathbf{A} = \begin{bmatrix} 5 & 1 \\ 3 & 4 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 5 & 2 \\ 2 & 4 \end{bmatrix} \quad \mathbf{C} = \begin{bmatrix} 1 & 2 & 5 \\ 2 & 4 & 2 \\ 5 & 2 & 7 \end{bmatrix}$$

Lösung

Die quadratische Form ist eine quadratische Funktion, welche nur quadratische Terme enthält.

- $A(x, y) = 5x^2 + 4xy + 4y^2$
- $B(x, y) = 5x^2 + 4xy + 4y^2$
- $C(x, y, z) = x^2 + 4y^2 + 7z^2 + 2xy + 10xz + 4yz$

Lösung Ende

2. Warum spricht man bei Definitheit per Konvention über symmetrische bzw. hermitesche Matrizen? Schau dir dazu die quadratische Form der Matrizen $\mathbf{A}$ und $\mathbf{B}$ in der vorherigen Teilaufgabe an!

Lösung

Zunächst gilt:

$$\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x} = (\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x})^\mathsf{T} = \mathbf{x}^\mathsf{T}\mathbf{A}^\mathsf{T}\mathbf{x}$$

Damit hat jede Matrix $\mathbf{A}$ und ihre Transponierte $\mathbf{A}^\mathsf{T}$ dieselbe quadratische Form. Somit ist:

$$\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x} = \frac{1}{2}\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x} + \frac{1}{2}\mathbf{x}^\mathsf{T}\mathbf{A}^\mathsf{T}\mathbf{x} = \mathbf{x}^\mathsf{T}\left(\frac{1}{2}(\mathbf{A} + \mathbf{A}^\mathsf{T})\right)\mathbf{x}$$

Jede nicht symmetrische Matrix $\mathbf{A}$ hat dieselbe quadratische Form wie ihr symmetrischer Anteil $\frac{1}{2}(\mathbf{A} + \mathbf{A}^\mathsf{T})$. Um die Definitheit einer nicht symmetrischen Matrix zu untersuchen, kann man stattdessen ihren symmetrischen Anteil untersuchen. Dies ist vorteilhaft, da viele Kriterien nur für symmetrische Matrizen gelten.

Lösung Ende

3. Welche geometrische Bedeutung hat es, dass eine Matrix positiv semidefinit ist?

Lösung

Es gilt:

$$\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x} = \langle \mathbf{x}, \mathbf{A}\mathbf{x} \rangle = \underbrace{\|\mathbf{x}\|}_{\geq 0}\underbrace{\|\mathbf{A}\mathbf{x}\|}_{\geq 0}\cos(\alpha)$$

Der Winkel α zwischen $\mathbf{x}$ und $\mathbf{Ax}$ ist also maximal $\frac{\pi}{2}$. Damit liegt die Transformation $\mathbf{Ax}$ immer im Halbraum von $\mathbf{x}$.

Lösung Ende

4. Welche Kriterien gibt es, um Definitheit zu untersuchen? Untersuche die folgenden Matrizen auf Definitheit!

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 5 & 2 \\ 2 & 1 \end{bmatrix} \quad \mathbf{C} = \begin{bmatrix} 2 & -3 \\ -3 & 6 \end{bmatrix}$$

$$\mathbf{D} = \begin{bmatrix} 2 & 3 & 1 \\ 3 & -6 & 2 \\ 1 & 2 & 9 \end{bmatrix} \quad \mathbf{E} = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \quad \mathbf{F} = \begin{bmatrix} -4 & 1 & 4 \\ 1 & -5 & 2 \\ 4 & 2 & -6 \end{bmatrix}$$

Lösung

Zur Untersuchung der Definitheit symmetrischer Matrizen gibt es zwei nützliche Kriterien:

(1) Eigenwertkriterium:

Sei $\mathbf{A} \in \mathbb{R}^{n\times n}$ eine **symmetrische** Matrix. Dann ist $\mathbf{A}$

positiv definit,	gdw. alle Eigenwerte größer als 0 sind.
positiv semidefinit,	gdw. alle Eigenwerte größer gleich 0 sind.
negativ definit,	gdw. alle Eigenwerte kleiner als 0 sind.
negativ semidefinit,	gdw. alle Eigenwerte kleiner gleich 0 sind.
indefinit,	sonst.

Es ist an dieser Stelle nochmal erwähnenswert, dass symmetrische Matrizen immer ausschließlich reelle Eigenwerte haben.

(2) Hauptminorenkriterium:

Sei $\mathbf{A} \in \mathbb{R}^{n\times n}$ eine **symmetrische** Matrix. Dann ist $\mathbf{A}$

- positiv definit, genau dann wenn alle führenden Hauptminoren positiv sind.
- negativ definit, genau dann wenn alle führenden Hauptminoren abwechselnd negativ und positiv sind.

Die führenden Hauptminoren sind die Determinanten aller $k \times k$ Untermatrizen von $\mathbf{A}$ mit $1 \leq k \leq n$, welche oben links beginnen.

Es gibt im Folgenden immer mehrere Lösungsmöglichkeiten, von denen wir jeweils nur eine (clevere) vorstellen:

a) Es gilt:

$$\det \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} = -3 = \lambda_1 \cdot \lambda_2$$

Damit muss ein Eigenwert positiv und der andere negativ sein. Somit ist die Matrix indefinit.

b) Es gilt:

$$\det 5 = 5, \quad \det \begin{bmatrix} 5 & 2 \\ 2 & 1 \end{bmatrix} = 1$$

Da alle führenden Hauptminoren größer als Null sind, ist die Matrix positiv definit und damit auch positiv semidefinit.

c) Es gilt:

$$\det 2 = 2, \quad \det \begin{bmatrix} 2 & -3 \\ -3 & 6 \end{bmatrix} = 3$$

Da alle führenden Hauptminoren größer als Null sind, ist die Matrix positiv definit und damit auch positiv semidefinit.

d) Da es ein positives und ein negatives Diagonalelement gibt, ist die Matrix indefinit. Schließlich gilt:

$$\begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \mathbf{D} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = 2, \quad \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} \mathbf{D} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = -6$$

e) Es gilt:

$$\det 2 = 2, \quad \det \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} = 3, \quad \det \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} = 4$$

Da alle führenden Hauptminoren größer als Null sind, ist die Matrix positiv definit und damit auch positiv semidefinit.

Hinweis: Für die Berechnung der Determinante einer 3×3-Matrix ist die Regel von Sarrus hilfreich.

f) Es gilt:

$$\det -4 = -4, \quad \det \begin{bmatrix} -4 & 1 \\ 1 & -5 \end{bmatrix} = 19, \quad \begin{bmatrix} -4 & 1 & 4 \\ 1 & -5 & 2 \\ 4 & 2 & -6 \end{bmatrix} = -2$$

Da der erste Hauptminor negativ, der zweite positiv und der dritte negativ ist, ist die Matrix negativ definit und damit auch negativ semidefinit.

Lösung Ende

5. Beweisen Sie, dass die Summe zweier positiv definiter Matritzen $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$ wieder eine positiv definite Matrix sein muss.

Lösung

Es gilt:

$$\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x} > 0 \text{ sowie } \mathbf{x}^\mathsf{T}\mathbf{B}\mathbf{x} > 0, \text{ für alle } \mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}.$$

Daher ergibt sich

$$\mathbf{x}^\mathsf{T}(\mathbf{A}+\mathbf{B})\mathbf{x} = \mathbf{x}^\mathsf{T}(\mathbf{A}\mathbf{x}+\mathbf{B}\mathbf{x}) = \mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x} + \mathbf{x}^\mathsf{T}\mathbf{B}\mathbf{x} > 0,$$

denn Summe zweier Werte größer als 0 muss auch größer als 0 sein.

Lösung Ende

6. Die quadratische ℓ^2-Norm $\|\mathbf{x}\|^2$ eines Vektors $\mathbf{x} \in \mathbb{R}^n$ lässt sich als eine quadratische Form $\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x}$ darstellen mit $\mathbf{A} \in \mathbb{R}^{n \times n}$. Geben Sie die zugehörige Matrix $\mathbf{A}$ an.

Lösung

Die Matrix $\mathbf{A}$ muss die Identität sein.

Lösung Ende