Lineare Systeme & Least Squares

15 Min Lineare Systeme
15 Min Least Squares
Anschließend Diskussion

Lineare Systeme

$x + y - z = 9$
$y + 3z = 3$
$-x - 2z = 2$

Was bedeutet linear?

- Man verwendet bei linearen Gleichungssystemen ausschließlich lineare Operatoren (Addition, Skalar-Multiplikation)

- Dinge wie sin(x), xy sind nicht erlaubt, da sie die Linearität verletzen

Das lineare System $Ax = b$

Was ist eine lineare Gleichung im Kontext der linearen Algebra?

Gaußsches Eliminationsverfahren

- Algorithmus in der linearen Algebra zur Lösung von linearen Gleichungssystemen
- Nötige Schritte:
  - Übertrage das Gleichungssystem in eine sogenannte augmentierte Matrix
  - Verwende Reihenoperationen, um die Matrix in eine Stufenform zu bringen (Pivotisierung)
  - Verwende Rückwärtseinsetzen, um das Gleichungssystem zu lösen

Pivotisierung

Ziel: Wende Reihenoperationen so lange an, bis du bei den jeweiligen Achsen Einsen stehen hast und für alle Einträge, falls $i > j$, 0 stehen hast

Rückwärtseinsetzen

Alternative: LU Dekomposition

- Man kann jede Matrix A in zwei Dreiecksmatrizen faktorisieren
- A = LU
- P = Permutationsmatrizen, die man mit A multipliziert, um die gesuchten Faktoren LU zu bekommen
- L1xL2xA = U
- L1 xL2 = I
- L = Inverse von I

Lineare Systeme und **Least Squares**

Amon Soares de Souza, Jakob Jungmaier

Ludwig-Maximilians-Universität München
Centrum für Informations- und Sprachverarbeitung (CIS)

9. Januar 2020

Inhalt

Überdeterminierte Systeme

Beispiel:

- Messwertreihe:

a	1	2	3	4
b	1.5	2.7	3.5	4.4

- Ziel: x_1 und x_2 finden, sodass $b = x_1 a + x_2$ für alle Messwerte
  $\Rightarrow$ Gerade $b = x_1 a + x_2$ finden
- Gleichungen:

$$\begin{aligned} x_1 + x_2 &= 1.5 \\ 2x_1 + x_2 &= 2.7 \\ 3x_1 + x_2 &= 3.5 \\ 4x_1 + x_2 &= 4.4 \end{aligned}$$

- Mehr Gleichungen als Unbekannte $\Rightarrow$ überdeterminiert
  $\Rightarrow$ im Allgemeinen keine Lösung!
  $\Rightarrow$ Linie durch alle vier Punkte: unmöglich!

Überdeterminierte Systeme

Lineares Gleichungssystem in Matrixform:

- $\mathbf{A}\vec{x} = \vec{b}$
- hier: $\begin{pmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1.5 \\ 2.7 \\ 3.5 \\ 4.4 \end{pmatrix}$
- überdeterminiert: $\mathbf{A} \in \mathbb{R}^{m \times n}, m > n$
  $\Rightarrow$ keine lineare Kombination der Spaltenvektoren von **A** ergibt $\vec{b}$
  $\Rightarrow$ $\vec{b}$ nicht im Spaltenraum von **A**

Least Squares Methode

Lösung: Least Squares Methode

- Gerade $b = x_1 a + x_2$ so gut wie möglich annähern
- $\vec{x}$ so wählen, dass Residuenvektor $\vec{r}$ so klein wie möglich: $\vec{r} = \vec{b} - \mathbf{A}\vec{x}$
- Least Squares $\Rightarrow$ euklidische Länge von $\vec{r}$ minimieren:

$$\min_{\vec{x}} ||\vec{b} - \mathbf{A}\vec{x}||_2$$

$$= \min_{\vec{x}} \sqrt{(b_1 - a_1.x_1)^2 + (b_2 - a_2.x_2)^2 + ... + (b_n - a_n.x_n)^2}$$
$$= \min_{\vec{x}}(b_1 - a_1.x_1)^2 + (b_2 - a_2.x_2)^2 + ... + (b_n - a_n.x_n)^2$$

Least Squares Methode

Wie?

- *(Gaußsche) Normalgleichungen* ("Normal Equations") lösen:

$$\mathbf{A}^\top \mathbf{A}\vec{x} = \mathbf{A}^\top \vec{b}$$

- hier:

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1.5 \\ 2.7 \\ 3.5 \\ 4.4 \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} 30 & 10 \\ 10 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 35 \\ 12.1 \end{pmatrix}$$

- wieder lineares Gleichungssystem, aber mit genau einer Lösung ($\mathbf{A}^\top \mathbf{A}$ reguläre Matrix (quadratisch und invertierbar)!
(Lösung hier: $x_1 = \frac{19}{20} = 0.95$, $x_2 = \frac{13}{20} = 0.65$)

Herleitung der Normalgleichungen

Woher kommen die Normalgleichungen $\mathbf{A}^\top \mathbf{A}\vec{x} = \mathbf{A}^\top \vec{b}$?

Geometrische Herleitung:

- $\vec{b}$ nicht in Spaltenraum von **A**
- Lösung: Projektion von $\vec{b}$ auf Spaltenraum von **A**

$\Rightarrow$ Abstand von $\vec{b}$ zu Spaltenraum von **A** minimieren
$\Rightarrow \vec{r}$

$\Rightarrow$ $\vec{r}$ muss orthogonal zu Spaltenraum von **A** gemacht werden
$\Rightarrow \vec{r}^\top \mathbf{A} = 0$

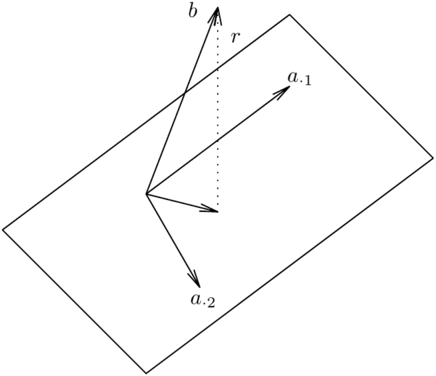

Abbildung: Siehe Eldén 2007, S. 33.

Herleitung der Normalgleichungen

- $\vec{r}^\top \mathbf{A} = 0$
- $\vec{r} = \mathbf{A}\vec{x} - \vec{b}$

$\Rightarrow$ Einsetzen: $(\mathbf{A}\vec{x} - \vec{b})^\top \mathbf{A} = 0$

$\Rightarrow \vec{x}^\top \mathbf{A}^\top \mathbf{A} - \vec{b}^\top \mathbf{A} = 0$

$\Rightarrow \vec{x}^\top \mathbf{A}^\top \mathbf{A} = \vec{b}^\top \mathbf{A}$

$\Rightarrow (\vec{x}^\top \mathbf{A}^\top \mathbf{A})^\top = (\vec{b}^\top \mathbf{A})^\top$

$\Rightarrow \mathbf{A}^\top \mathbf{A}\vec{x} = \mathbf{A}^\top \vec{b}$

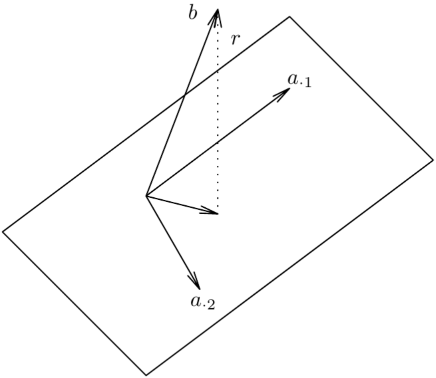

Abbildung: Siehe Eldén 2007, S. 33.

Herleitung der Normalgleichungen

Herleitung mit *Calculus* (Skizze, siehe Townsend 2015):

- Ziel: $\min_{\vec{x}} ||\vec{b} - \mathbf{A}\vec{x}||_2$
- Minimum folgender Funktion finden: $f(\vec{x}) = ||\vec{b} - \mathbf{A}\vec{x}||_2$

  $f(\vec{x}) = (\vec{b} - \mathbf{A}\vec{x})^\top(\vec{b} - \mathbf{A}\vec{x})$
- Partielle Ableitung nach jeder Variablen $x_1, ..., x_n$:
  $\Rightarrow$ Gradient: $\nabla f(\vec{x}) = 2\mathbf{A}^\top\mathbf{A}\vec{x} - 2\mathbf{A}^\top\vec{b}$
- Gradient gleich Null setzen:
  $\nabla f(\vec{x}) = \vec{0}$
  $\Rightarrow 2\mathbf{A}^\top\mathbf{A}\vec{x} - 2\mathbf{A}^\top\vec{b} = \vec{0}$
  $\Rightarrow \mathbf{A}^\top\mathbf{A}\vec{x} = \mathbf{A}^\top\vec{b}$

Zusammenfassung

- Darstellung und Lösung linearer Gleichungssysteme in Matrixform
- Überdeterminierte Gleichungssysteme:

  $\Rightarrow$ Keine Lösung für $\mathbf{A}\vec{x} = \vec{b}$

  $\Rightarrow$ Lösung (bzw. Annäherung) mit den Normalgleichungen $\mathbf{A}^\top\mathbf{A}\vec{x} = \mathbf{A}^\top\vec{b}$

Danke!

Literatur

Eldén, Lars (2007). *Matrix methods in data mining and pattern recognition*. Siam. Kap. 3.

Townsend, Alex (2015). *Least squares and the normal equations*. Available at `http://math.mit.edu/icg/resources/teaching/18.085-spring2015/LeastSquares.pdf`.