Computergestützte Mathematik zur Linearen Algebra

Wiederholung Lineare Algebra

Heinrich-Heine-Universität

28. Nov. & 5. Dez. & 19. Dez. 2024

Vektorräume

Beispiele für $\mathbb{R}$-Vektorräume

1. $\mathbb{P}_m := \left\{\sum_{i=0}^m a_i X^i | a_i \in \mathbb{R}\right\}$

   Addition:

   $p + q := \sum_{i=0}^m (p_i + q_i) X^i$ für $p := \sum_{i=0}^m p_i X^i$ und $q := \sum_{i=0}^m q_i X^i$

   Skalarmultiplikation:

   $\alpha \cdot p := \sum_{i=0}^m (\alpha p_i) X^i$ für $p := \sum_{i=0}^m p_i X^i$ und $\alpha \in \mathbb{R}$.

2. Reelle 2×2 Matrizen

   Addition:

   $A + B := (a_{ij} + b_{ij})_{i,j=1}^2$ für $A = (a_{ij})_{i,j=1}^2$ und $B = (b_{ij})_{i,j=1}^2$
   $(a_{ij}, b_{ij} \in \mathbb{R})$

   Skalarmultiplikation:

   $\alpha \cdot A := (a_{ij})_{i,j=1}^2$ für $A = (\alpha a_{ij})_{i,j=1}^2$ und $\alpha \in \mathbb{R}$.

Lineare (Un-)Abhängigkeit

Definition

Eine Menge von n Vektoren $M = \{u_0, \ldots, u_{n-1}\} \subset V$ heißt *linear unabhängig*, falls gilt

$$\sum_{j=0}^{n-1} \alpha_j u_j = 0 \ \Rightarrow \alpha_j = 0 \text{ für alle } j = 0, \ldots, n-1.$$

Definition

Ein linear unabhängiges Erzeugendensystem heißt *Basis*.

Definition

Die Anzahl der Elemente einer Basis eines Vektorraums heißt *Dimension* des Vektorraums.

Lineare Abbildungen

Definition

Seien U, V zwei $\mathbb{R}$-Vektorräume. Eine Abbildung $\alpha : U \to V$ heißt *linear*, falls

- $\alpha(u + v) = \alpha(u) + \alpha(v)$
- $\alpha(zu) = z\alpha(u)$

für alle Vektoren $u, v \in U$ und reelle z gilt.

Welche Abbildungen sind linear?

1. $\mathbb{R}^2 \to \mathbb{R}^2 : (x, y) \mapsto (-y, x)$
2. $\mathbb{R}^2 \to \mathbb{R}^2 : (x, y) \mapsto (xy, 0)$
3. $\mathbb{R}^2 \to \mathbb{R} : (x, y) \mapsto x - y$
4. $\mathbb{R}^2 \to \mathbb{R} : (x, y) \mapsto x$

Lineare Abbildungen und Matrizen

Theorem

Jede $\mathbb{R}$-lineare Abbildung α von einem n-dimensionalen in in einen m-dimensionalen VR kann bezüglich von Basen des Bild- und Urbildvektorraums als Matrix $A = (a_{ij})_{i,j=0}^{n-1,m-1}$ dargestellt werden. (i Zeilenindex, j Spaltenindex)

Sei $\{u_j\}_{j=0}^{m-1}$ Basis von U und $\{v_i\}_{i=0}^{n-1}$ Basis von V, dann ist $\alpha(u_j) = \sum_{i=0}^{n-1} a_{ij} v_i$ für gewisse $a_{ij} \in \mathbb{R}$. (Die j-te Spalten von A bekommt man, wenn das Bild des j-ten Basisvektors u_j betrachtet.)
Für einen Vektor $U \ni x = \sum_{j=0}^{m-1} x_j u_j$ mit Koordinaten $x_j \in \mathbb{R}$ gilt

$$\alpha(x) = \sum_{j=0}^{m-1} x_j \alpha(u_j) = \sum_{j=0}^{m-1} x_j \sum_{i=0}^{n-1} a_{ij} v_i = \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} a_{ij} x_j v_i$$

Damit sind $y_i := \sum_{j=0}^{m-1} a_{ij} x_j$ $(y = Ax)$ die Koordinaten des Bildvektors bzgl. der Basis von V, d.h. $\alpha(x) = \sum_{i=0}^{n-1} y_i v_i$.

Lineare Abbildungen und Matrizen: Begriffe und Fragestellungen

- $\mathrm{kern}(A) = \{v \in V \mid Av = 0\} \subseteq V$
- $\mathrm{bild}(A) = \{w \in W \mid \exists v \in V, Av = w\} \subseteq W$
- Basen für $\mathrm{kern}(A)$ und $\mathrm{bild}(A)$
- $\mathrm{rang}(A) = \dim(\mathrm{bild}(A))$
- Falls $n = m$, ist $A \in \mathrm{Gl}(n)$? Existiert A^{-1}?
- Zu $A \in \mathrm{Gl}(n)$ löse
$$Ax = b, \quad b \in \mathbb{R}^n$$
- Zu $A \in \mathbb{R}^{m \times n}$ finde $x \in \mathbb{R}^n$, so dass
$$\|Ax - b\|_2 \to \min$$

Lösung mit Hilfe der Zeilenstufenform

Die Lösung von $Ax = b$ ist gegeben durch $x = A^{-1}b$.

Ansatz: Berechne A^{-1} mit Hilfe der spez. Zeilenstufenform

$$\left[\begin{array}{c|c} A & I \end{array}\right] \longrightarrow \left[\begin{array}{c|c} Z_A & E_l \cdots E_1 E_0 I \end{array}\right] = \left[\begin{array}{c|c} I & A^{-1} \end{array}\right]$$

Berechne $x = A^{-1}b$.

Da $A^{-1} = E_l \cdots E_1 E_0$ (E_i Elementarmatrix), ist

$$A^{-1}b = E_l \cdots E_1 E_0 b.$$

Statt A^{-1}, berechne direkt spez. Zeilenstufenform von

$$\left[\begin{array}{c|c} A & b \end{array}\right] \longrightarrow \left[\begin{array}{c|c} Z_A & E_l \cdots E_1 E_0 b \end{array}\right] = \left[\begin{array}{c|c} I & A^{-1}b \end{array}\right]$$

Elementarmatrizen

Definition Matrizen der Form $E = I - uv^T$, $u, v \in \mathbb{R}^n$ mit $u^T v \neq 1$ heißen Elementarmatrizen.

$$uv^T = \begin{bmatrix} u_0 \\ \vdots \\ u_{n-1} \end{bmatrix} \begin{bmatrix} v_0, \ldots v_{n-1} \end{bmatrix} = \begin{bmatrix} u_0 v_0 & \cdots & u_0 v_{n-1} \\ \vdots & & \vdots \\ u_{n-1} v_0 & \cdots & u_{n-1} v_{n-1} \end{bmatrix}$$

$$v^T u = \begin{bmatrix} v_0, \ldots, v_{n-1} \end{bmatrix} \begin{bmatrix} u_0 \\ \vdots \\ u_{n-1} \end{bmatrix} = \sum_{i=0}^{n-1} v_i u_i$$

E ist invertierbar, $E^{-1} = I - \frac{1}{v^T u - 1} uv^T$

Spezielle Elementarmatrizen (elementare Zeilentransformationen)

Sei $e_i{}_{i=0}^{n-1}$ die Vektoren der Standardbasis (1 an Stelle i sonst Null)

- Multiplikation der i-ten Zeile mit $\alpha \in \mathbb{R}$ oder $\mathbb{C}$

$$E = I - (1 - \alpha) e_i e_i^T$$

- Addition der i-ten Zeile zur j-ten Zeile

$$E = I + e_j e_i^T$$

- (Addition des α-fachen der i-ten Zeile zur j-ten Zeile

$$E = I + \alpha e_j e_i^T)$$

- Vertausche i-te und j-te Zeile

$$E = I - uu^T, \qquad u = e_i - e_j$$

Zeilenstufenform

Erinnerung: Jede Matrix kann durch elementare Zeilenumformungen (oder äquivalent durch Multiplikation von links mit Elementarmatrizen) auf Zeilenstufenform Z transformiert werden:

$$A \to Z = \begin{bmatrix} 0 & \otimes & \times & \cdots & \cdots & \cdots & \cdots & \cdots & \times \\ 0 & 0 & 0 & \otimes & & & & & \vdots \\ \vdots & & \vdots & 0 & \otimes & \times & \times & & \vdots \\ \vdots & & \vdots & \vdots & 0 & 0 & 0 & \otimes & \times \\ \vdots & & \vdots & \vdots & \vdots & \vdots & \vdots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$\otimes$ = Pivotelemente $\neq 0$

Zeilenstufenform II

Eigenschaften der Zeilenstufenform Z von A

- Pivotpositionen sind eindeutig durch Elemente von A bestimmt
- Einträge in Z sind nicht eindeutig bestimmt
- Spalten von A, welche Pivotpositionen enthalten heißen Basisspalten
- $$\begin{aligned} \text{Rang}(A) &= \#\ \text{Basisspalten von } A \\ &= \#\ \text{von Null verschiedene Zeilen von } Z \\ &= \#\ \text{Pivotelemente} \end{aligned}$$

Beispiel

Aufgabe: Bestimme den Rang und identifiziere die Basisspalten von

$$A = \begin{bmatrix} 1 & 2 & 1 & 1 \\ 2 & 4 & 2 & 2 \\ 3 & 6 & 3 & 4 \end{bmatrix}$$

Lösung: Zeilenstufenform

$$Z = \begin{bmatrix} 1 & 2 & 1 & 1 \\ \hline 0 & 0 & 0 & \multicolumn{1}{|c}{1} \\ \cline{4-4} 0 & 0 & 0 & 0 \end{bmatrix}$$

damit: Rang $A = 2$, Pivotelemente in 1. und 4. Spalte

Basisspalten: 1. und 4. Spalte von A $(A_{:,0}, A_{:,3})$

Spezielle Zeilenstufenform

Erinnerung: Z ist in spezieller Zeilenstufenform Z_A, wenn

- Z in Zeilenstufenform ist
- alle Pivotelemente eins sind
- alle Einträge oberhalb der Pivotelemente Null sind

$$Z_A = \begin{bmatrix} 1 & \times & 0 & 0 & \times & \times & 0 & \times \\ 0 & 0 & 1 & 0 & \times & \times & 0 & \times \\ \vdots & \vdots & 0 & 1 & \times & \times & 0 & \times \\ \vdots & \vdots & \vdots & 0 & 0 & 0 & 1 & \times \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Z_A ist eindeutig bestimmt

Beispiel

Aufgabe: Bestimme die spezielle Zeilenstufenform Z_A, den Rang und die Basisspalten von

$$A = \begin{bmatrix} 1 & 2 & 2 & 3 & 1 \\ 2 & 4 & 4 & 6 & 2 \\ 3 & 6 & 6 & 9 & 6 \\ 1 & 2 & 4 & 5 & 3 \end{bmatrix}$$

Lösung:

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

damit: Rang $A = 3$, Basisspalten: $A_{:,0}, A_{:,2}, A_{:,4}$

Berechnung der Inversen einer Matrix

Erinnerung: Ist $A \in \mathrm{GL}(n)$, also invertierbar, dann ist die spezielle Zeilenstufenform die Identität: $Z_A = I_n$

$$E^l \cdots E^2 E^1 A = I \iff E^l \cdots E^2 E^1 = A^{-1} = E^l \cdots E^2 E^1 I$$

Gauß-Jordan-Verfahren: Transformiere die erweiterte Matrix

$$\left[\begin{array}{c|c} A & I \end{array}\right] \longrightarrow \left[\begin{array}{c|c} I & A^{-1} \end{array}\right]$$

Beispiel:

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{bmatrix} \longrightarrow A^{-1} = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$

Inverse einer Matrix

Eigenschaften der Inversen: Es seien $A, B \in \mathbb{K}^{n\times n}$ nicht singulär, dann gilt

- $(A^{-1})^{-1} = A$
- AB ist nicht singulär
- $(AB)^{-1} = B^{-1}A^{-1}$
- $(A^{-1})^T = (A^T)^{-1}$, $(A^{-1})^* = (A^*)^{-1}$,

Gauß-Elimination (Zeilenstufenform für $A \in Gl(n)$)

Nach $k-1$ Schritten Gauß-Elimination erhält man

$$A \longrightarrow A_{k-1} = \begin{bmatrix} \times & \cdots & \times & \alpha_1 & \times & \cdots & \times \\ & \ddots & & \vdots & \vdots & & \vdots \\ & & \times & \alpha_{k-1} & \times & \cdots & \times \\ & & & \alpha_k & \times & \cdots & \times \\ & & & \vdots & \vdots & & \vdots \\ & & & \alpha_n & \times & \cdots & \times \end{bmatrix}$$

k-ter Schritt

```
if A[k,k] = 0
   Finde Index l, so dass A[l,k] ~= 0
   Tausche Zeilen A[k,:] und A[l,:]

Eliminiere Eintraege A[k+1:,:]
```

Eliminationsschritt I

Erinnerung: Subtraktion des $\ell_{jk} = \alpha_j / \alpha_k$-fachen der k-ten Zeile von der j-ten Zeile entspricht Links-Multiplikation mit

$$I - \ell_{jk} e_j e_k^T$$

Elimination von `A[k+1:,k]` $\Leftrightarrow$ Links-Multiplikation mit L_k

$$L_k = (I - \ell_{n-1,k} e_{n-1} e_k^T) \cdots (I - \ell_{k+2,k} e_{k+2} e_k^T)(I - \ell_{k+1,k} e_{k+1} e_k^T)$$

Wegen $\langle e_i, e_j \rangle_2 = e_j^T e_i = 0, \quad i \neq j$

$$\begin{aligned} L_k &= I - \ell_{n,k} e_n e_k^T - \ldots - \ell_{k+2,k} e_{k+2} e_k^T - \ell_{k+1,k} e_{k+1} e_k^T \\ &= I - \ell_k e_k^T, \quad \ell_k = \begin{bmatrix} 0 & \cdots & 0 & \ell_{k+1,k} & \cdots & \ell_{n-1,k} \end{bmatrix}^T \end{aligned}$$

Eliminationsschritt II

Nach Konstruktion von L_k gilt dann

$$A_k = L_k \cdot A_{k-1} = \begin{bmatrix} \times & \cdots & \times & \alpha_0 & \times & \cdots & \times \\ & \ddots & & \vdots & \vdots & & \vdots \\ & & \times & \alpha_{k-1} & \times & \cdots & \times \\ & & & \alpha_k & \times & \cdots & \times \\ & & & 0 & \vdots & & \vdots \\ & & & \vdots & \vdots & & \vdots \\ & & & 0 & \times & \cdots & \times \end{bmatrix}$$

$$= (I - \ell_k e_k^T) A_{k-1} = A_{k-1} - \ell_k e_k^T A_{k-1},$$

Beobachtung

$L_k \cdot$ ändert die ersten $k-1$ Zeilen von A_{k-1} nicht

Algorithmus (ohne Vertauschungen)

Mit $A_0 = A$ und $A_k = L_k A_{k-1}$ erhält man

$$L_{n-1} \cdots L_1 \cdot L_0 A = R \quad \text{obere Dreiecksmatrix}$$

Wegen $L_k^{-1} = I + \ell_k e_k^T$ ist dann

$$A = L_0^{-1} L_1^{-1} \cdots L_{n-1}^{-1} R = LR$$

$$\begin{aligned} \text{Mit} \quad L &= (I + \ell_0 e_0^T)(I + \ell_1 e_1^T) \cdots (I + \ell_{n-1} e_{n-1}^T) \\ &= \begin{bmatrix} 1 & & & \\ \ell_{1,0} & 1 & & \\ \vdots & \ddots & \ddots & \\ \ell_{n-1,0} & \cdots & \ell_{n-1,n-2} & 1 \end{bmatrix} \end{aligned}$$

Beispiel

Die Elemente von L und R können auf den Speicherplätzen von A gespeichert werden.

$$A = \begin{bmatrix} 2 & 2 & 2 \\ 4 & 7 & 7 \\ 6 & 18 & 22 \end{bmatrix} \longrightarrow \left[\begin{array}{c|cc} 2 & 2 & 2 \\ \hline 2 & 3 & 3 \\ 3 & 12 & 16 \end{array}\right] \longrightarrow \left[\begin{array}{cc|c} 2 & 2 & 2 \\ \hline 2 & 3 & 3 \\ \hline 3 & 4 & 4 \end{array}\right]$$

Wir erhalten also die LR-Zerlegung

$$A = LR \quad \text{mit} \quad L = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 4 & 1 \end{bmatrix}, R = \begin{bmatrix} 2 & 2 & 2 \\ 0 & 3 & 3 \\ 0 & 0 & 4 \end{bmatrix}$$

LR-Zerlegung – Algorithmus

```
L = np.eye(n);

for k in range(n-1):








  for i in range(k+1, n):
    L[i, k] = A[i, k] / A[k,k];
    for j in range(k, n):
      A[i, j] = A[i, j] - L[i, k]*A[k, j];
R = A;
```

LR-Zerlegung – Algorithmus

```
L = np.eye(n);
P = np.eye(n);
for k in range(n-1):
  if A[k,k] == 0:
    Suche Index l > k, so dass  A[l,  k] != 0
    Tausche Zeilen k und l von A, L und P und
    Spalten k und l von L

  for i in range(k+1, n):
    L[i, k] = A[i, k] / A[k, k];
    for j in range(k, n):
      A[i, j] = A[i, j] - L[i, k]*A[k, j];
R = A;
```

LR-Zerlegung – Platzsparende Variante

```
p = range(n);
for k in range(n - 1):
  if A[k,k] == 0:
    Suche Index l > k, so dass  A[l,k] != 0
    Tausche Zeilen k und l von A
    vertausche Elemente k und l von p

  for i in range(k+1, n):
    A[i, k]     = A[i, k] / A[k, k];
    A[i, k+1:] = A[i, k+1:] - A[i, k] * A[k, k+1:];
```

Dann:

```
R = np.triu(A);
L = np.eye(n) + np.tril(A,-1);
P = Permutationsmatrix, die p[k] auf k abbildet, k = 0,
```

- j-Schleife vektorisiert
- A zum Speichern von L und R verwendet

Satz (LR-Zerlegung)

Sei $A \in \mathbb{K}^{n \times n}$ so, dass bei der Reduktion auf Zeilenstufenform ohne Zeilenvertauschung kein Pivotelement 0 auftritt. Dann gibt es eine Zerlegung $A = LR$ mit

1. L ist untere und R ist obere Dreiecksmatrix
2. $\ell_{i,i} = 1$ und $r_{i,i} \neq 0$ für $i = 1, 2, \ldots, n$
3. L und R sind damit eindeutig bestimmt.

Anwendungen der LR-Zerlegung

Lösung linearer Gleichungssysteme

Sei $A = LR$ die LR-Zerlegung

$$
\begin{aligned}
& Ax = b \\
\iff & L(Rx) = b \\
\iff & Ly = b, \quad Rx = y
\end{aligned}
$$

- L und R haben Dreiecksform
- $Ly = b \longrightarrow$ Vorwärtssubstitution
- $Rx = y \longrightarrow$ Rückwärtssubstitution

Aufwand

$$\begin{aligned}
\sum_{k=0}^{n-2} \sum_{i=k+1}^{n-1} (n-k+1) &= \sum_{k=0}^{n-2} (n-k-1)(n-k+1) \\
&= \sum_{j=1}^{n-1} j(j+2) \\
&= \frac{1}{6}(n-1)n(2n-1) + n(n-1) \\
&= \frac{1}{3}n^3 + \frac{1}{2}n^2 - \frac{5}{6}n \\
&= \frac{1}{3}n^3 + \mathcal{O}(n^2)
\end{aligned}$$

Erinnerung: $\displaystyle\sum_{j=1}^{m} j^2 = \frac{1}{6}m(m+1)(2m+1)$

Vorwärtssubstitution

$$Ly = b \Longleftrightarrow \begin{bmatrix} 1 & & & \\ \ell_{2,1} & 1 & & \\ \vdots & \ddots & \ddots & \\ \ell_{n,1} & \cdots & \ell_{n,n-1} & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

Lösung

```
for i in range(n):
    y[i] = b[i];
    for j in range(i):
        y[i] -= L[i,j]*y[j];
```

Aufwand $\displaystyle\sum_{i=0}^{n-1} i = \frac{1}{2}n(n-1) = \frac{1}{2}n^2 + O(n)$

Rückwärtssubstitution

$$Rx = y \iff \begin{bmatrix} r_{1,1} & r_{1,2} & \cdots & r_{1,n} \\ & \ddots & \ddots & \vdots \\ & & r_{n-1,n-1} & r_{n-1,n} \\ & & & r_{n,n} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

Lösung

```
for i in range(n-1, -1, -1): # Rückwärts von n-1, ..., 0
    x[i] = y[i];
    for j in range(i+1, n):
        x[i] -= R[i,j]*x[j];
    x[i] = x[i]/R[i,i];
```

Aufwand $\displaystyle\sum_{i=0}^{n-1}(n-i) = n^2 - \frac{n(n-1)}{2} = \frac{1}{2}n^2 + O(n)$

Bemerkungen zur LR-Zerlegung

- Ist die LR-Zerlegung von A schon berechnet und möchte man ein weiteres Gleichungssystem $Ax = c$ lösen, so kann man die teure LR-Zerlegung ($\sim \frac{1}{3}n^3$ Operationen) wiederverwenden und muss nur die billigen Vorwärts-/Rückwärtseliminationen (je $\sim \frac{1}{2}n^2$ Operationen) neu rechnen.
- Wendet man das Gauß-Verfahren direkt auf die erweiterte Matrix $[A \mid b]$ an, so ergibt sich $[R \mid L^{-1}b]$, die Vorwärtselimination kann also einfach mitgerechnet werden.

Bilinearformen

Sei V ein $\mathbb{R}$-Vektorraum.

Definition

Eine Abbildung $a : V \times V \to \mathbb{R}$ heißt Bilinearform auf V, falls

- $a(u, v + w) = a(u, v) + a(u, w)$
- $a(u + w, v) = a(u, v) + a(w, v)$
- $a(\alpha u, v) = \alpha a(u, v)$
- $a(u, \alpha v) = \alpha a(u, v)$

für alle Vektoren u, v, w und $\alpha \in \mathbb{R}$ gilt.

Die Abbildungen $x \mapsto a(x, v)$ und $x \mapsto a(u, x)$ sind beide linear.

Definition

a heißt positiv (semi-) definit, falls $\mathbb{R} \ni a(u, u) > 0$ bzw. ≥ 0 für alle $0 \neq u \in V$.

Bilinearformen

Sei V ein $\mathbb{R}$-Vektorraum.

Definition

Eine Abbildung $a : V \times V \to \mathbb{R}$ heißt Bilinearform auf V, falls

- $a(u, v + w) = a(u, v) + a(u, w)$
- $a(u + w, v) = a(u, v) + a(w, v)$
- $a(\alpha u, v) = \alpha a(u, v)$
- $a(u, \alpha v) = \alpha a(u, v)$

für alle Vektoren u, v, w und $\alpha \in \mathbb{R}$ gilt.

Die Abbildungen $x \mapsto a(x, v)$ und $x \mapsto a(u, x)$ sind beide linear.

Definition

a heißt positiv (semi-) definit, falls $\mathbb{R} \ni a(u, u) > 0$ bzw. ≥ 0 für alle $0 \neq u \in V$.

Sesquilinearformen

Sei V ein $\mathbb{C}$-Vektorraum.

Definition

Eine Abbildung $a : V \times V \to \mathbb{C}$ heißt Sesquilinearform auf V, falls

- $a(u, v + w) = a(u, v) + a(u, w)$
- $a(u + w, v) = a(u, v) + a(w, v)$
- $a(\alpha u, v) = \alpha a(u, v)$
- $a(u, \alpha v) = \overline{\alpha} a(u, v)$

für alle Vektoren u, v, w und $\alpha \in \mathbb{C}$ gilt.

D.h. $a(\cdot, v)$ ist linear und $a(u, \cdot)$ ist semilinear.

Definition

a heißt positiv (semi-) definit, falls $\mathbb{R} \ni a(u, u) > 0$ bzw. ≥ 0 für alle $0 \neq u \in V$.

Sesquilinearformen

Sei V ein $\mathbb{C}$-Vektorraum.

Definition

Eine Abbildung $a : V \times V \to \mathbb{C}$ heißt Sesquilinearform auf V, falls

- $a(u, v + w) = a(u, v) + a(u, w)$
- $a(u + w, v) = a(u, v) + a(w, v)$
- $a(\alpha u, v) = \alpha a(u, v)$
- $a(u, \alpha v) = \overline{\alpha} a(u, v)$

für alle Vektoren u, v, w und $\alpha \in \mathbb{C}$ gilt.

D.h. $a(\cdot, v)$ ist linear und $a(u, \cdot)$ ist semilinear.

Definition

a heißt positiv (semi-) definit, falls $\mathbb{R} \ni a(u, u) > 0$ bzw. ≥ 0 für alle $0 \neq u \in V$.

Skalarprodukt / inneres Produkt

Sei V ein $\mathbb{K}$-Vektorraum (reell oder komplex).

Definition

Eine Bi-/Sesquilinearform a auf V heißt symmetrisch/hermitesch, falls $a(u, v) = \overline{a(v, u)}$ für alle $u, v \in V$.

Definition

Eine symmetrische/hermitesche, positiv definite Bi-/Sesquilinearform ist ein Skalarprodukt.

Beispiel:
Für $u, v \in \mathbb{C}^n$ ist

$$(u, v) \mapsto \sum_{k=1}^{n} u_k \overline{v_k} = v^H u =: \langle u, v \rangle$$

ein Skalarprodukt (das euklidische Skalarprodukt des $\mathbb{C}^n$).

Skalarprodukt / inneres Produkt

Sei V ein $\mathbb{K}$-Vektorraum (reell oder komplex).

Definition

Eine Bi-/Sesquilinearform a auf V heißt symmetrisch/hermitesch, falls $a(u, v) = \overline{a(v, u)}$ für alle $u, v \in V$.

Definition

Eine symmetrische/hermitesche, positiv definite Bi-/Sesquilinearform ist ein Skalarprodukt.

Beispiel:
Für $u, v \in \mathbb{R}^n$ ist

$$(u, v) \mapsto \sum_{k=1}^{n} u_k v_k = v^T u =: \langle u, v \rangle$$

ein Skalarprodukt (das euklidische Skalarprodukt des $\mathbb{R}^n$).

Beispiele Skalarprodukte

1. $(A, B) \mapsto \sum_{i,j=1}^{n} a_{ij} b_{ij}$ ist ein Skalarprodukt auf dem Vektorraum der reellen $n \times n$ Matrizen.
2. $(p, q) \mapsto \int_0^1 p(t)\overline{q(t)}dt$ ist ein Skalarprodukt auf $\mathbb{P}_m$.

Normen

Definition

Sei V ein $\mathbb{K}$-Vektorraum. Eine Abbildung $||\cdot|| : V \to \mathbb{R}^+$ heißt Norm, wenn

1. $||x|| = 0 \ \Rightarrow x = 0$
2. $||x + y|| \leq ||x|| + ||y||$ (Dreiecksungleichung)
3. $||\alpha x|| = |\alpha| \ ||x||$

Induzierte Norm

Definition

Sei V ein $\mathbb{K}$-Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$. Für ein $x \in V$ ist $||x|| := \sqrt{\langle x, x \rangle}$, die von $\langle \cdot, \cdot \rangle$ induzierte Norm.

Theorem

Eine von einem Skalarprodukt induzierte Norm ist eine Norm.

Beispiel:
Für $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ ist

$$||x||^2 = \sum_{i=1}^{n} x_i^2 = x^T x$$

die euklidische Länge von x, die vom euklidischen Skalarprodukt induzierte Norm.

Induzierte Normen

Definition

Seien $||\cdot||_V$ und $||\cdot||_W$ Normen über $\mathbb{K}$-Vektorräumen V bzw. W, $A : V \to W$ linear. Durch

$$||A||_{W \leftarrow V} = \sup_{x \in V \setminus \{0\}} \frac{||A(x)||_W}{||x||_V} = \sup_{x \in V, ||x||_V = 1} ||A(x)||_W$$

wird eine Norm auf dem Vektorraum der linearen Abbildungen induziert.

Theorem

$||A||_{W \leftarrow V}$ *ist eine Norm.*

Theorem

Es gilt

$$\| \operatorname{Id} \|_{V \leftarrow V} = 1.$$

Beispiele: $||A||_1$ und $||A||_\infty$

Sei $A = [a_1 \mid a_2 \mid \cdots \mid a_n] \in \mathbb{C}^{m\times n}$ und $||x||_1 = 1$:

$$||Ax||_1 = ||\sum_{j=1}^{n} x_j a_j||_1 \leq \sum_{j=1}^{n} |x_j|||a_j||_1 \leq \max_{1\leq j\leq n} ||a_j||_1$$

Damit: $||A||_1 \leq \max_{1\leq j\leq n} ||a_j||_1$

Für $x = e_k$, wobei k so, dass $||a_k||_1 = \max_{1\leq j\leq n} ||a_j||_1$, gilt Gleichheit, also

$$||A||_1 = \max_{1\leq j\leq n} ||a_j||_1 \qquad \text{maximale Spaltensummennorm}$$

analog:

$$||A||_\infty = \max_{1\leq i\leq n} ||A(i,:)^T||_1 \qquad \text{maximale Zeilensummennorm}$$

Unitäre und Orthogonale Abbildungen

Definition

Sei A eine lineare Abbildung auf einem $\mathbb{R}$-Vektorraum mit Skalarprodukt $(V, \langle\cdot,\cdot\rangle)$ und davon induzierter Norm. Falls $||Ax|| = ||x||$ für alle $x \in V$ gilt, so heißt A orthogonal.

Definition

Sei A eine lineare Abbildung auf einem $\mathbb{C}$-Vektorraum mit Skalarprodukt $(V, \langle\cdot,\cdot\rangle)$ und davon induzierter Norm. Falls $||Ax|| = ||x||$ für alle $x \in V$ gilt, so heißt A unitär.

Beispiele

1. In $\mathbb{R}^2$ ist die Drehung gegeben durch die Matrix

$$\begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix}$$

2. In $\mathbb{C}^3$ ist die Spiegelung an der (x, y)-Ebene

$$(x, y, z) \mapsto (x, y, -z)$$

3. In $\mathbb{R}^3$ ist die Projektion auf die (x, y)-Ebene

$$(x, y, z) \mapsto (x, y, 0)$$

Orthogonal- und Orthonormalbasen

Definition

Sei V ein $\mathbb{K}$-Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle_V$ (reell oder komplex).
Eine Menge $\{v_1, \dots, v_n\} \subseteq V$ heißt orthogonal, falls

$$\langle v_i, v_j \rangle_V = 0, \qquad \text{falls } i \neq j.$$

Eine orthogonale Menge $\{v_1, \dots, v_n\} \subseteq V$ heißt orthonormal, falls

$$\langle v_i, v_i \rangle_V = 1, \qquad \text{für } 1 \leq i \leq n.$$

Entsprechende Basen heißen Orthogonalbasis bzw. Orthonormalbasis.

Seien $\{v_1, \ldots, v_n\} \subseteq \mathbb{R}^m$ paarweise ortho*normale* Vektoren, so ist die Matrix

$$Q = [v_1 \mid \cdots \mid v_n]$$

ortho*gonal*.

Umgekehrt sind die Spalten einer ortho*gonalen* Matrix ortho*normale* Vektoren.

Gram-Schmidt-Orthogonalisierung

Erinnerung:

Linear unabhängige Vektoren $a_1, \dots, a_n$ in $\mathbb{R}^m$ können zu paarweise orthogonalen Vektoren $q_1, \dots, q_n$ orthonormalisiert werden, sodass für $k = 1, \dots, n$

$$\text{span}\,\{a_1, \dots, a_k\} = \text{span}\,\{q_1, \dots, q_k\}$$

Gram-Schmidt-Orthogonalisierung: Für $k = 1, \dots, n$

$$\tilde{q}_k = a_k - \sum_{j=1}^{k-1} r_{jk} q_j, \qquad r_{jk} = \langle q_j, a_k \rangle$$

$$q_k = \tilde{q}_k / r_{kk}, \qquad r_{kk} = ||\tilde{q}_k||$$

Gram-Schmidt-Orthogonalisierung

Erinnerung:
Linear unabhängige Vektoren $a_1, \ldots, a_n$ in $\mathbb{R}^m$ können zu paarweise orthogonalen Vektoren $q_1, \ldots, q_n$ orthonormalisiert werden, sodass für $k = 1, \ldots, n$

$$\text{span } \{a_1, \ldots, a_k\} = \text{span } \{q_1, \ldots, q_k\}$$

Gram-Schmidt-Orthogonalisierung: Für $k = 1, \ldots, n$

$$\tilde{q}_k = a_k - \sum_{j=1}^{k-1} r_{jk} q_j, \qquad r_{jk} = \langle q_j, a_k \rangle$$

$$q_k = \tilde{q}_k / r_{kk}, \qquad r_{kk} = ||\tilde{q}_k||$$

QR-Zerlegung für $A \in \mathbb{R}^{m\times n}$

Gram-Schmidt-Orthogonalisierung: für $k = 1, \ldots, n$

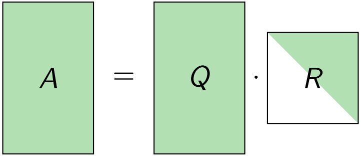
$$\tilde{q}_k = a_k - \sum_{j=1}^{k-1} r_{jk} q_k, \qquad q_k = \tilde{q}_k / r_{kk}$$

$$\iff a_k = \sum_{j=1}^{k} r_{jk} q_j = \begin{bmatrix} q_1 \mid q_2 \mid \ldots \mid q_k \end{bmatrix} \begin{bmatrix} r_{1k} \\ \vdots \\ r_{kk} \end{bmatrix}$$

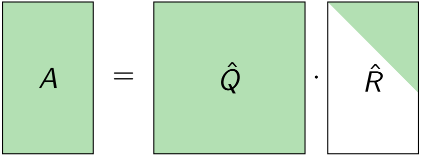
$$\iff \begin{bmatrix} a_1 \mid a_2 \mid \ldots \mid a_n \end{bmatrix} = \begin{bmatrix} q_1 \mid q_2 \mid \ldots \mid q_n \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ & r_{22} & \cdots & r_{2n} \\ & & \ddots & \vdots \\ & & & r_{nn} \end{bmatrix}$$

Reduzierte und volle QR-Zerlegung

Reduzierte QR-Zerlegung



Volle QR-Zerlegung



Reduzierte QR-Zerlegung

Definition

Sei $A \in \mathbb{R}^{m \times n}$ mit $\text{Rang}(A) = n < m$ gegeben. Dann heißt

- $A = QR$ mit $Q \in \mathbb{R}^{m \times n}$ (spalten-) orthogonal und oberer Dreiecksmatrix $R \in \mathbb{R}^{n \times n}$ heißt reduzierte QR-Zerlegung von A.
- $A = \hat{Q}\hat{R}$ mit $\hat{Q} = [Q \ Q_\perp] \in \mathbb{R}^{m \times m}$ orthogonal und oberer Dreiecksmatrix $\hat{R} = \begin{bmatrix} R \\ 0_{m-n,n} \end{bmatrix} \in \mathbb{R}^{m \times n}$ heißt volle QR-Zerlegung.

Theorem

Die reduzierte QR-Zerlegung von A ist eindeutig, wenn man $r_{ii} > 0$ voraussetzt.

Modifiziertes Gram-Schmidt-Verfahren

Klassische Gram-Schmidt-Orthogonalisierung:

$$\tilde{q}_k = a_k - \sum_{j=1}^{k-1} r_{jk} q_j, \qquad r_{jk} = \langle q_j, a_k \rangle$$

$$q_k = \tilde{q}_k / r_{kk}, \qquad r_{kk} = ||\tilde{q}_k||$$

Interpretation (falls die Summe in der Reihenfolge 1,2,... berechnet wird):

- orthogonalisiere a_k (gleichzeitig) gegen $q_1, \ldots, q_{k-1}$

Modifikation:

- orthogonalisiere $\tilde{q}_k$ nacheinander gegen $q_1, \ldots, q_{k-1}$

Modifiziertes Gram-Schmidt-Verfahren

- orthogonalisiere a_j zunächst gegen q_1:

$$\tilde{q}_{1k} = a_k - r_{1j} q_1, \qquad r_{1k} = \langle q_1, a_k \rangle$$

- anschließend orthogonalisiere $\tilde{q}_{1k}$ gegen q_2:

$$\tilde{q}_{2k} = \tilde{q}_{1k} - r_{2k} q_2, \qquad r_{2k} = \langle q_2, \tilde{q}_{1k} \rangle$$

- anschließend orthogonalisiere $\tilde{q}_{2k}$ gegen q_3:

$$\tilde{q}_{3k} = \tilde{q}_{2k} - r_{3k} q_3, \qquad r_{3k} = \langle q_3, \tilde{q}_{2k} \rangle$$

- usw.

In exakter Arithmetik sind beide Algorithmen äquivalent, denn

$$r_{jk} = \langle q_j, a_k \rangle = \left\langle q_j, \left(a_k - \sum_{i=1}^{j-1} r_{ik} q_i\right) \right\rangle$$

da $\langle q_j, q_i \rangle = 0$ für $j \neq i$.

Vor- und Nachteile

- Orthogonalisierungen *müssen* beim modifizierten Verfahren nacheinander durchgeführt werden
- in Gleitkommaarithmetik ist der modifizierte Gram-Schmidt-Algorithmus stabiler

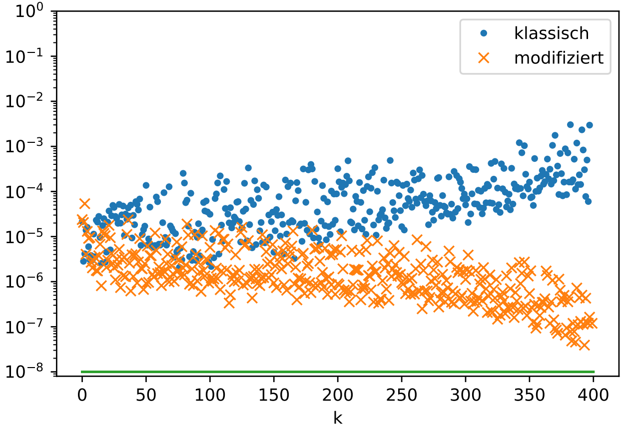
Klassisches und Modifiziertes Gram Schmidt Verfahren

Gegeben Basisvektoren $[a_1, .., a_n]$
for $k = 1 \dots n$ **do**
  $q_k = a_k$
  if $k \neq 1$ **then**
    $r_{j,k} = \langle q_j, q_k \rangle$, für $j = 1 \dots k-1$
    $q_k = q_k - \sum_{j=1}^{k-1} q_j r_{j,k}$
  end if
  $r_{k,k} = ||q_k||$
  $q_k = q_k / r_{k,k}$
end for
Erhalte Basisvektoren $q_1, \dots, q_n$ und Matrix R

Gegeben Basisvektoren $[a_1, .., a_n]$
for $k = 1, \dots, n$ **do**
  $q_k = a_k$
  for $i = 1..k-1$ **do**
    $r_{i,k} = \langle q_i, q_k \rangle$
    $q_k = q_k - q_i r_{i,k}$
  end for
  $r_{k,k} = ||q_k||$
  $q_k = q_k / r_{k,k}$
end for
Erhalte Basisvektoren $q_1, \dots, q_n$ und Matrix R

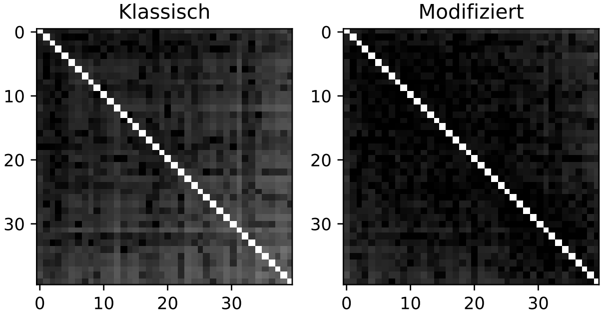
Stabilitätsvergleich

Klassisches und modifiziertes Gram-Schmidt-Verfahren
Orthogonalität von Q: $\max_{j>k} |q_k^T q_j|$



Stabilitätsvergleich

Klassisches und modifiziertes Gram-Schmidt-Verfahren
Einträge von $\log_{10}(|Q^T Q|)$. Weiß $\hat{=}\ 0 = \log_{10}(1)$,
Schwarz $\hat{=}\ -8 = \log_{10}(10^{-8}) = \log_{10}(eps)$,



Lösung linearer Gleichungssysteme mit QR-Zerlegung

Für $A \in \mathbb{R}^{n \times n}$ nicht singulär ist

$$Ax = b \Longleftrightarrow QRx = b$$

Wie löst man mit Q? Tipp: $Q^T Q = Id$

I.A. teurer aber (mit mod. Gram-Schmidt oder Householder-Transformationen) stabiler als Gauß-Elimination.

Lösung linearer Gleichungssysteme mit QR-Zerlegung

Für $A \in \mathbb{R}^{n \times n}$ nicht singulär ist

$$Ax = b \Longleftrightarrow QRx = b \Longleftrightarrow Rx = Q^T b$$

Lösung eines Systems mit Dreiecksstruktur (Rückwärtssubstitution)

I.A. teurer aber (mit mod. Gram-Schmidt oder Householder-Transformationen) stabiler als Gauß-Elimination.

Lösung linearer Gleichungssysteme mit QR-Zerlegung

Für $A \in \mathbb{R}^{n \times n}$ nicht singulär ist

$$Ax = b \Longleftrightarrow QRx = b \Longleftrightarrow Rx = Q^T b$$

Lösung eines Systems mit Dreiecksstruktur (Rückwärtssubstitution)

I.A. teurer aber (mit mod. Gram-Schmidt oder Householder-Transformationen) stabiler als Gauß-Elimination.