

# Lineare Algebra Zusammenfassung

African Hit  
Winter 2016/17

## Generell:

Axiome:

$$(\alpha B)A = \alpha(BA)$$

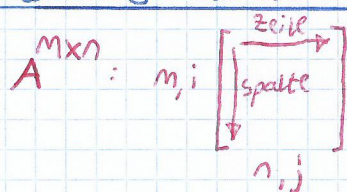
$$(\alpha A)B = \alpha(AB) = A(\alpha B)$$

$$(\alpha + \beta)A = (\alpha A) + (\beta A)$$

$$\alpha(A+B) = (\alpha A) + (\alpha B)$$

$$A+B = B+A$$

$$A(BC) = (AB)C$$



Transponiert:  $A \in \mathbb{R}$ ,  $A^T$  ist A „an der Diagonalen gespiegelt“  
 Adjungiert:  $A \in \mathbb{C}$ ,  $A^H$  ist transponiert + Vorzeichenwechsel bei den Imaginärteilen.

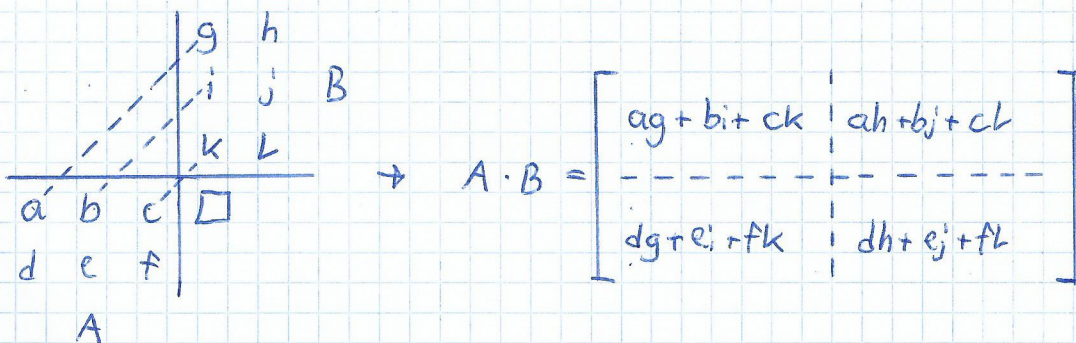
Kommutativität: Falls gilt  $AB = BA$ , so sind A, B quadratisch und A, B symmetrisch.

## Matrixmultiplikation und Gauss



A sei eine  $m \times n$  Matrix  
 B eine  $k \times l$  Matrix.  
 Es muss  $n=k$  gelten.  
 Resultat ist  $m \times l$  gross.

Berechnungsschema:



## Gauss-Elimination:

Finden der Lösung von  $Ax=b$ , bzw. eine Matrix auf Zeilenstufenform bringen.

„Erlaubte“ Operationen: Addieren von Zeilen bzw. deren Vielfachern zu anderen, Vertauschen von Zeilen.

Beachte: Im Allgemeinen Fall muss die Anzahl Gleichungen und die Anzahl Variablen nicht übereinstimmen! Die überflüssigen Gleichungen werden dabei alle 0. (Nur 0 als Koeff.)

Rang: Der Rang eines LGS ist die Anzahl „Nicht-Null“-Zeilen = Anzahl Pivotelemente nach dem Gaussen.

## Spezielle Matrizen, Skalarprodukt, Inverse

Einheits-/ Identitätsmatrix:  $n \times n$  Matrix  $I_n = \text{diag}(1, 1, \dots, 1)$

Nullmatrix:  $m \times n$  Matrix, alle Einträge sind 0. Rang = 0.

Quadratische Matrix:  $n \times n$  Matrix (beide Seiten gleich gross)

Symmetrisch / Hermitesch: Eine quadratische Matrix ist genau dann symmetrisch / hermitesch, wenn sie ihrem Transponierten / Adjungierten entspricht:

Adjungierten entspricht:

$$A = A^T \Rightarrow A \text{ symmetrisch}$$

$$A = A^H \Rightarrow A \text{ Hermitesch, } A^H = (\bar{A})^T = \overline{(A^T)}$$

Obere Dreiecksmatrix: Quadratische Matrix A, wo  $a_{ij} = 0$  für  $i > j$  (d.h. alle Einträge unter Diagonalen sind 0.)

Diagonalmatrix: Auch nur für quadratische Matrizen definiert, alles ausser Diagonale sind 0.

Unitär / Orthogonal:  $U^H U = U U^H = I$  bzw.  $U^T U = U U^T = I$

## Regulär

Quadratisch ( $m=n$ )

$\text{Rang}(A) = n$  (keine Lin. Abh.)

$\det(A) \neq 0$

$\text{EW} \neq 0$

$Ax=b$  hat 1 Lsg pro  $b$ .

Invertierbar

$Ax=0 \Rightarrow x=0$

## Singular

$\text{Rang}(A) < n$

$\det(A) = 0$

$\text{EW} = 0$

$Ax=b$ :

- Keine eind. Lsg für alle  $b$

- Unendl. Lsg für einige  $b$

- Keine Lsg für einige  $b$

$Ax=0 \Rightarrow x \neq 0$

## LR / LU - Zerlegung

Algorithmus:  $Ax=b$  gegeben.  $x$  sei unbekannt.  $A$  regulär.

1) Schreibe  $(I_n | A)$ .

2) Gauss  $A$  und schreibe Mult. Faktoren mit umgekehrtem

Vorzeichen in  $I_n$  an Stelle wo bei  $A$  eine Null entsteht.

3) Erhalte  $(L | R)$ .

4) Berechne  $Lc = b$ .

5) Berechne  $Rx = c$ .

6) Überprüfe  $Ax=b \rightarrow$  Profit!

Dies gilt nur ohne Zeilenumtauschungen!

Mit Zeilenumtauschungen / Pivotisieren:

Wie oben, aber setze  $P = I_n$  und wenn bei  $A$  beim Gauss Zeilen vertauscht werden, dann tausche diese auch in  $P$ !

Beziehungen:

$A = LR$ ;  $Lc = b$ ;  $Rx = c$

bzw.

$PA = LR$ ;  $Lc = Pb$ ;  $Rx = c$

Wichtige Regeln:

$(A^T)^T = A$ ,  $(A^H)^H = A$ ;  $(\alpha A)^T = \alpha \cdot A^T$ ,  $(\alpha A)^H = \alpha \cdot A^H$

$(A+B)^T = A^T + B^T$ ;  $(A+B)^H = A^H + B^H$

$(AB)^T = B^T A^T$ ;  $(AB)^H = B^H A^H$  für  $A^{m \times n}$  und  $B^{n \times p}$

Skalarprodukt:  $\langle x, y \rangle = x^T y$  bzw.  $x^H y$

• Lin. im 2. Faktor:  $\langle x, \alpha(y+z) \rangle = \alpha \langle x, y \rangle + \alpha \langle x, z \rangle$

• Symmetrie:  $\langle x, y \rangle = \langle y, x \rangle$

• Positiv Definit:  $\langle x, x \rangle \geq 0$  und  $\langle x, x \rangle = 0 \Leftrightarrow x = 0$

Länge von Vektoren:  $|a| = \|a\| = \sqrt{a_1^2 + a_2^2 + a_3^2} = \sqrt{\langle a, a \rangle}$

Inverse:  $A^{-1}$ , so dass  $A \cdot A^{-1} = A^{-1} \cdot A = I$  gilt.

Regeln:  $(A^{-1})^{-1} = A$ ;  $(AB)^{-1} = B^{-1} A^{-1}$ ;  $(A^T)^{-1} = (A^{-1})^T$ ;  $(A^H)^{-1} = (A^{-1})^H$

Berechnung: Sei  $A$  reguläre  $n \times n$  Matrix.

1) Schreibe  $(A | I_n)$

2) Gauss so dass  $(I_n | \square)$

3) Ablesen von  $A^{-1}$

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

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

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

$$\begin{pmatrix} 1 & 0 & -2 & 1 \\ 0 & 1 & 3/2 & -1/2 \end{pmatrix} \cdot A^{-1}$$

## Vektorräume, Span, Lin. Abhängigkeit, Erzeugendensystem, Basis

Vektorraumaxiome:

V1 (Assoziativ) :  $(x+y)+z = x+(y+z)$

V2 (Neutralement) :  $x+0 = 0+x = x$

V3 (Inverses) :  $\forall x \in V \exists x^{-1} \in V: x+x^{-1} = 0$

V4 (Kommutativ) :  $x+y = y+x$

S1 (Distributiv) :  $\alpha(x+y) = \alpha \cdot x + \alpha \cdot y$

S2 (Distributiv) :  $(\alpha+\beta)x = \alpha \cdot x + \beta \cdot x$

S3 (Assoziativ):  $(\alpha \cdot \beta) \cdot x = \alpha \cdot (\beta \cdot x)$

S4 (Neutralement):  $1 \cdot x = x$

Polynome  $\mathcal{P}_n$ : Polynome vom Grad  $n$ ,  $p(t) = a_0 t^0 + a_1 t^1 + \dots + a_n t^n$

Untervektorraum: nichtleere Teilmenge von  $V$  mit:

U1 (Abgeschlossenheit):  $\forall u, v \in U: u + v \in U$

U2 (Abgeschlossenheit):  $\forall \lambda \in \mathbb{K}, u \in U: \lambda \cdot u \in U$

→ Jeder UVR ist auch ein VR und es ist i. d. R. leichter nur die zwei Eigenschaften des UVR zu prüfen!

Span:  $\text{span}\{a_1, \dots, a_n\}$  ist die Menge der Linearkombinationen aus  $a_1, \dots, a_n$ . Er bildet einen UVR.

Lineare (Un)Abhängigkeit:  $a_1, \dots, a_n$  sind genau dann linear unabhängig wenn sich kein Vektor als Linearkombination der anderen darstellen lässt.

Erzeugendensystem: Eine Menge aus denen man alle  $v \in V$  als Linearkombination dieser Vektoren darstellen kann.

Basis: Sei  $\mathcal{B}$  eine Teilmenge von  $V$ .  $\mathcal{B}$  ist Basis von  $V$

$\Leftrightarrow V = \text{span}(\mathcal{B})$  und alle  $b \in \mathcal{B}$  sind linear unabh. voneinander.

$\Leftrightarrow \mathcal{B}$  ist eine max. lin. unabh. Teilmenge von  $V$

$\Leftrightarrow \mathcal{B}$  ist ein minimales Erzeugendensystem von  $V$ .

Dimension des VR: Anzahl Basisvektoren des VR.

### Lineare Abbildungen, Basiswechsel

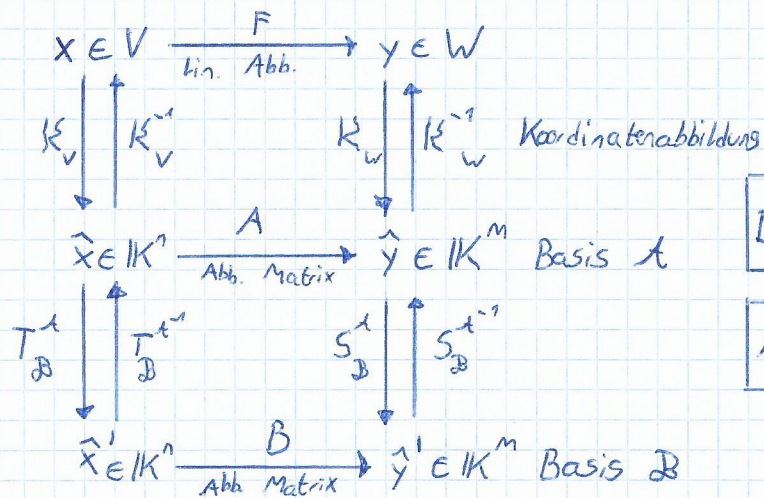
Seien  $V, W$  VR über  $\mathbb{K}$ .

Lineare Abbildung:  $F: V \rightarrow W, x \mapsto Fx$  stellt ein Element aus einem VR in einem anderen VR dar.

Abbildungsmatrix:  $m \times n$ -Matrix (für  $V$   $n$ -dimensionaler VR und  $W$   $m$ -dimensionaler VR). Stellt lin. Abb. in einer Basis dar. Berechnung: Basisvektoren in lineare Abbildung einsetzen, „Basis abbilden“.

Transformationsmatrix: Von alter Basis  $\mathcal{A}$  in neue Basis  $\mathcal{B}$ :  $T_{\mathcal{B}}^{\mathcal{A}}$ .

Berechnung:  $(\mathcal{B} | \mathcal{A}) \rightarrow \text{Gauß} \rightarrow (I_n | T_{\mathcal{B}}^{\mathcal{A}})$



$$B = S_{\mathcal{B}}^{\mathcal{A}} \cdot A \cdot T_{\mathcal{A}}^{\mathcal{B}}$$

$$A = S_{\mathcal{B}}^{\mathcal{A}^{-1}} \cdot B \cdot T_{\mathcal{B}}^{\mathcal{A}}$$

Koordinatenabbildung: Wie kann ein Vektor als Linearkombination der Basisvektoren dargestellt werden?  $\Rightarrow$  Vektoren normalerweise in  $\mathcal{B}$  dargestellt, d.h.  $\begin{pmatrix} 4 \\ 1 \end{pmatrix} = \underline{1} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \underline{4} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

Für lin. Abbildungen gilt:

1)  $F(x + y) = F(x) + F(y)$

2)  $F(\alpha \cdot x) = \alpha \cdot F(x)$

## Kern, Rang einer Matrix

Sei  $F: X \rightarrow Y$  eine lin. Abbildung. Der Kern von  $F$  ist dann die Menge aller  $x \in X$  so dass  $Fx = 0$ :

$$\ker(F) := \{x \in X \mid Fx = 0\}.$$

Der Kern wird oft als Nullspace  $\mathcal{N}$  bezeichnet.

Eine Matrix  $A$  hat nur einen Kern wenn ihre Determinante  $= 0$  ist.

Der Kern ist die Lösungsmenge des homogenen Gleichungssystem

$$Ax = 0, \text{ d.h. um den Kern zu bestimmen löst man einfach } Ax = 0.$$

Der Kern ist dann der span dieser Lösungsvektoren!

Der Rang einer Matrix  $A$  ist die Anzahl linear unabhängiger Vektoren in  $A$ . Bei einer Abbildung  $F$  ist der Rang als Dimension des Bildes von  $F$  definiert:  $\text{rang}(F) = \dim(\text{Bild}(F))$ .

Ist der Rang einer Quadratischen Matrix  $=$  ihrer Größe ( $r = n$ ), so hat sie vollen Rang und ist regulär (invertierbar). Dies trifft genau dann zu, wenn ihre Determinante  $\neq 0$  ist.

Es gilt:

$$1) \text{ Rang}(A) = \text{Rang}(A^T)$$

$$2) \text{ Rang}(A+B) \leq \text{Rang}(A) + \text{Rang}(B)$$

## Norm, Orthogonalbasen und Gram-Schmidt

Sei  $V$  ein VR über  $\mathbb{K}$ . Eine Norm in  $V$  ist eine Funktion

$$\|\cdot\|: V \rightarrow \mathbb{R}, x \mapsto \|x\| \quad \text{mit folgenden Eigenschaften:}$$

$$1) \text{ positiv definit: } \forall x \in V: \|x\| \geq 0, \quad \|x\| = 0 \Rightarrow x = 0$$

$$2) \text{ homogen: } \|\alpha x\| = |\alpha| \cdot \|x\| \quad \forall x \in V \wedge \forall \alpha \in \mathbb{K}$$

$$3) \Delta\text{-Ungleichung: } \forall x, y \in V: \|x+y\| \leq \|x\| + \|y\|.$$

$$\cdot \text{Eins-Norm: } \|v\|_1 = \|(v_1, \dots, v_n)\| := |v_1| + \dots + |v_n|$$

$$\cdot \text{Zwei-Norm: } \|v\|_2 = \sqrt{v_1^2 + \dots + v_n^2} = \sqrt{\langle v, v \rangle} \text{ Euklidische- / Standardnorm}$$

$$\cdot \text{Maximum-Norm: } \|v\|_\infty = \max(|v_i|, 1 \leq i \leq n)$$

$$\text{Cauchy-Schwarz-Ungleichung: } \forall x, y \in V: |\langle x, y \rangle| \leq \|x\| \cdot \|y\|$$

Orthogonale Basis: Basisvektoren stehen paarweise normal zueinander:

$$|\langle b_k, b_l \rangle| = 0, \text{ falls } k \neq l$$

Orthonormale Basis: Ist eine Basis, die orthogonal ist und zusätzlich Länge 1 hat ( $\|b_k\| = 1, \forall k$ ).

$$M \text{ orthogonal} \rightarrow M^{-1} = M^T$$

Koordinatenbestimmung  $V$  ein VR mit  $\dim = n$ ,  $\{b_1, \dots, b_n\}$  eine ONB.

$$1) \text{ Da } \{b_1, \dots, b_n\} \text{ eine Basis: } \forall x \in V: x = \sum_{k=1}^n \lambda_k \cdot b_k$$

$$2) \text{ Da } \{b_1, \dots, b_n\} \text{ eine ONB: } \forall x \in V: \lambda_k = \langle b_k, x \rangle \Rightarrow x = \sum_{k=1}^n \langle b_k, x \rangle \cdot b_k$$

Parsevalsche Formel: Sei  $V$  ein VR und  $\{b_1, \dots, b_n\}$  eine ONB. Seien

$$x, y \in V: \lambda_k = \langle b_k, x \rangle, \mu_k = \langle b_k, y \rangle, \lambda = (\lambda_1, \dots, \lambda_n), \mu = (\mu_1, \dots, \mu_n)$$

$$\text{so gilt nun } \langle x, y \rangle = \langle \lambda, \mu \rangle \Rightarrow \text{Ist Längentreu: } \|x\| = \|\lambda\|$$

$$\text{und ist winkeltreu: } \angle(x, y) = \angle(\lambda, \mu) \Rightarrow x \perp y = \lambda \perp \mu$$

Gram-Schmidt-Verfahren: Mit GS können wir eine Menge an

Vektoren  $\{a_1, \dots, a_n\}$  in eine orthonormale Menge  $\{b_1, \dots, b_n\}$

umwandeln. Ist  $\{a_1, \dots, a_n\}$  eine Basis ist  $\{b_1, \dots, b_n\}$  eine ONB

Algorithmus: Berechnet aus Menge  $\{a_1, \dots, a_n\}$  eine Orthonormale

Menge  $\{b_1, \dots, b_n\}$ :

$$b_1 := \frac{a_1}{\|a_1\|}$$

$$\tilde{b}_k := a_k - \sum_{j=1}^{k-1} \langle b_j, a_k \rangle \cdot b_j \quad \left. \vphantom{\tilde{b}_k} \right]_{k=2, \dots, n}$$

$$b_k := \frac{\tilde{b}_k}{\|\tilde{b}_k\|}$$

GS bezüglich Skalarprodukt:

Falls Basis in ONB bezüglich

eines speziellen Skalarprodukts  $s$

berechnet wird, ersetze einfach

SP in  $\langle b_j, a_k \rangle$  und  $\|b_k\|$  durch

$s$ !

## Projektion, Methode der kleinsten Quadrate, QR-Zerlegung

**Projektion:** Lineare Abbildung  $P: \mathbb{E}^m \rightarrow \mathbb{E}^m$ , so dass  $P = P^2$  gilt, d.h. alle Vektoren die in ihrem Bild liegen auf sich selbst abbilden. Gilt zusätzlich  $\mathcal{N}(P) \perp \mathcal{R}(P)$  (der Nullraum steht senkrecht zum Bildraum), so sprechen wir von einer Orthogonalprojektion.

**Berechnung der Orthogonalprojektion:** Sei  $A$  eine  $m \times n$ -Matrix. Eine orthogonale Projektion  $P_A$  auf den Bildraum von  $A$  lässt sich folgendermassen berechnen:

$$P_A := A(A^H A)^{-1} A^H$$

Für eine Orthogonalprojektion  $P$  gilt:  $\|y - Py\|_2 = \min_{z \in \mathcal{L}(P)} \|y - z\|_2$ . Dies besagt, dass für einen Punkt  $y$  der nicht in der Projektionsebene liegt ( $y \notin \mathcal{L}(P)$ ) die Abbildung  $P \cdot y$  einen Punkt in der Projektionsebene ( $Py \in \mathcal{L}(P)$ ) liefert, welcher am nächsten bei  $y$  liegt (d.h. Distanz von  $y$  zu  $Py$  ist minimal).

**QR-Zerlegung:** Zerlege  $m \times n$ -Matrix  $A$  in eine orthogonale Matrix  $Q$  und eine rechte Obere Dreiecksmatrix  $R$  so dass  $A = QR$  gilt.

**Algorithmus für Reduzierte QR-Zerlegung:**

1) Berechne mittels Gram-Schmidt aus  $A = (a_1 | a_2 | \dots | a_n)$   $n$  orthonormale Vektoren  $q_1, \dots, q_n \Rightarrow Q = (q_1 | \dots | q_n)$

2) Berechne  $R$  wie folgt:

$$r_{11} := \|a_1\|$$

$$r_{jk} := \langle q_j, a_k \rangle \quad j=1, \dots, k-1$$

$$r_{kk} := \| \tilde{q}_k \|$$

$$\Rightarrow R = \begin{pmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ 0 & r_{22} & \dots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & r_{nn} \end{pmatrix}$$

**Algorithmus für Vollständige QR-Zerlegung:**

- 1) Erweitere  $Q$  um  $(m-n)$  orthonormale Vektoren, so dass  $Q$  eine quadratische  $(m \times m)$  orthonormale Matrix wird.  $q_{n+1}, \dots, q_m$  können dabei mit den Vektoren des Kerns und Gram-Schmidt berechnet werden:  $\tilde{Q} = (Q | Q_\perp)$
- 2) Füge  $(m-n)$  Nullzeilen in  $R$  hinzu:  $\tilde{R} = \begin{pmatrix} R \\ 0 \end{pmatrix}$   
Nun gilt  $A = \tilde{Q} \cdot \tilde{R}$

**Methode der kleinsten Quadrate:**

Betrachte das überbestimmte LGS  $Ax = y$  mit einer „hohen“ Matrix  $A$  ( $m \times n$ , wobei  $m > n \Rightarrow$  mehr Gleichungen als Unbekannte). Dieses LGS hat im Allgemeinen keine Lösung, aber wir können ein  $x$  so bestimmen, dass der „Fehler“ zwischen  $y$  und  $Ax$  (das Residuum bzw. der Residuenvektor  $r = y - Ax$ ) möglichst klein wird, d.h. dass die Euklidische Norm möglichst minimal ist.

$x$  heisst dabei Lösung im Sinne der kleinsten Quadrate.

Wir wissen:  $Ax = P_A y = A(A^H A)^{-1} A^H y$  gilt. Daraus können wir die Normalgleichungen  $x = (A^H A)^{-1} A^H y$  bzw.  $A^H A x = A^H y$  bestimmen.  $(A^H A)^{-1} A^H$  wird dabei Pseudoinverse genannt.

Achtung: Für grosse Matrizen  $A$  ist das Verfahren mittels QR-Zerlegung meistens besser:

**Verfahren mittels QR-Zerlegung:**

$$A = \tilde{Q} \tilde{R} \Rightarrow \text{Pseudoinverse: } (A^H A)^{-1} A^H \Leftrightarrow ((\tilde{Q} \tilde{R})^H (\tilde{Q} \tilde{R}))^{-1} (\tilde{Q} \tilde{R})^H \Leftrightarrow (\tilde{R}^H \tilde{Q}^H \tilde{Q} \tilde{R})^{-1} (\tilde{R}^H \tilde{Q}^H) \Leftrightarrow \tilde{R}^{-1} \tilde{R}^H \tilde{R}^{-1} \tilde{R}^H \tilde{Q}^H \Leftrightarrow \tilde{R}^{-1} \tilde{Q}^H$$

$\Rightarrow Ax = y$  kann also durch  $x = \tilde{R}^{-1} \tilde{Q}^H y$  approximiert werden.

## Determinante

Die Determinante einer quadratischen  $n \times n$ -Matrix  $A$  ordnet dieser einen Skalar zu und wird mit  $\det(A)$  oder  $|A|$

bezeichnet:  $\det: A \in \mathbb{E}^{n \times n} \rightarrow \mathbb{E}$

1x1 Matrix:  $\det(a) = a$

2x2 Matrix:  $\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$

3x3 Matrix:  $\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \rightsquigarrow \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{22} & a_{23} & a_{33} \\ a_{31} & a_{32} & a_{33} \\ a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}$

$= a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23}$

$- a_{31}a_{22}a_{13} - a_{21}a_{12}a_{33} - a_{11}a_{32}a_{23}$

Zeilenvertauschen ändert das Vorzeichen der Determinante!

Wichtige Eigenschaften der Determinante:

1)  $\det(A) = \det(A^T) \quad \forall A \in \mathbb{R}^{n \times n}$

2)  $A \in \mathbb{R}^{n \times n}$  regulär:  $\det(A^{-1}) = \frac{1}{\det(A)}$

3) Determinante von oberen / unteren Dreiecksmatrizen ist das Produkt der Diagonalelemente.

4) Zeilenvertauschen ändert Vorzeichen:  $\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = -\det \begin{pmatrix} c & d \\ a & b \end{pmatrix}$

5) Multiplikation einer Zeile mit einer Konstanten multipliziert die Determinante um das selbe:  $\det \begin{pmatrix} \lambda a & \lambda b \\ c & d \end{pmatrix} = \lambda \cdot \det \begin{pmatrix} a & b \\ c & d \end{pmatrix}$

6) Determinante von Matrixprodukten:  $\det(AB) = \det(A) \cdot \det(B)$

Achtung:  $\det(A+B) \neq \det(A) + \det(B)$

7)  $A \in \mathbb{R}^{n \times n}$ ,  $\lambda \in \mathbb{R}$ ,  $\det(\lambda A) = \lambda^n \cdot \det(A)$

8) Blockdreiecksmatrizen:  $A, B, C, D \in \mathbb{R}^{n \times n}$ ,  $\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det \begin{pmatrix} A & 0 \\ C & D \end{pmatrix} = \det(A) \cdot \det(D)$

9) Determinante = 0  $\Leftrightarrow$  Matrix ist singulär

10) Determinante  $\neq 0 \Leftrightarrow$  Matrix ist regulär

Determinante via Gauss-Algorithmus:

1) Bringe  $n \times n$ -Matrix  $A$  durch Gauss-Algorithmus in Zeilenstufenform.

2) Berechne  $(-1)^v \prod_{k=1}^n r_{kk}$  wobei  $v$  die Anzahl Zeilenvertauschungen ist (d.h. fast wie bei normaler Dreiecksmatrix).

Ähnliche Matrizen haben die gleiche Determinante.

Bei einer Dreiecksmatrix ist die Determinante ebenfalls das Produkt der Diagonalelemente.

## Laplacescher Entwicklungssatz, Cramersche Regel

Mit dem Laplaceschen Entwicklungssatz können Determinanten von beliebigen  $n \times n$ -Matrizen berechnet werden. Man kann dabei nach der  $i$ -ten Zeile oder der  $j$ -ten Spalte entwickeln, wobei es sich anbietet, eine Zeile / Spalte mit möglichst vielen Nullen zu suchen.

• Entwicklung nach  $i$ -ter Zeile:  $\det(A) = \sum_{j=1}^n (-1)^{i+j} \cdot a_{ij} \cdot \det(A_{ij})$

• Entwicklung nach  $j$ -ter Spalte:  $\det(A) = \sum_{i=1}^n (-1)^{i+j} \cdot a_{ij} \cdot \det(A_{ij})$

wobei  $A_{ij}$  die  $(n-1) \times (n-1)$ -Untermatrix von  $A$  ist, die man durch Streichen der  $i$ -ten Zeile und  $j$ -ten Spalte erhält.

Für Kofaktor  $(-1)^{i+j}$  gilt:

Beispiel:  $A = \begin{pmatrix} 1 & 0 & 4 \\ 6 & 1 & 3 \\ 0 & 0 & 6 \end{pmatrix} \rightarrow$  Entw. nach 2. Spalte

$\Rightarrow \det(A) = (-1)^{1+2} \cdot 0 \cdot \det \begin{pmatrix} 6 & 3 \\ 0 & 6 \end{pmatrix} + (-1)^{2+2} \cdot 1 \cdot$

$\det \begin{pmatrix} 1 & 4 \\ 0 & 6 \end{pmatrix} + (-1)^{3+2} \cdot 0 \cdot \det \begin{pmatrix} 1 & 4 \\ 6 & 3 \end{pmatrix}$

$= \det \begin{pmatrix} 1 & 4 \\ 0 & 6 \end{pmatrix} = \underline{\underline{6}}$

## Cramer'sche Regel:

Hat man ein reguläres LGS gegeben, kann man mit der Cramer'schen Regeln einzelne Unbekannte direkt berechnen.

Erklärung anhand eines Beispiels:

$$\begin{array}{l} \text{Gegeben sei LGS: } 2x + 1y + 1z = 3 \\ 1x - 1y - 1z = 0 \\ 1x + 2y + 1z = 0 \end{array} \rightsquigarrow \begin{pmatrix} 2 & 1 & 1 \\ 1 & -1 & -1 \\ 1 & 2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \\ 0 \end{pmatrix}$$

Um nun beispielsweise  $y$  zu berechnen, berechne

$$D_y = \det \begin{pmatrix} 2 & 3 & 1 \\ 1 & 0 & -1 \\ 1 & 0 & 1 \end{pmatrix} = -6 \quad \text{und} \quad D = \det \begin{pmatrix} 2 & 1 & 1 \\ 1 & -1 & -1 \\ 1 & 2 & 1 \end{pmatrix} = 3$$

$$y \text{ ist dann } y = \frac{D_y}{D} = \frac{-6}{3} = -2.$$

Dies funktioniert für beliebig grosse LGS.  $D_i$  ist einfach die Matrix die man erhält, wenn man die  $i$ -te Unbekannte (die, die man finden will) durch die rechte Seite ersetzt.

## Eigenwerte, Eigenvektoren, Eigenwertzerlegung, Ähnlichkeitstransformation

Wir betrachten lineare Abbildungen mit  $F: V \rightarrow V; x \mapsto Fx$  eines endlich-dimensionalen VR in sich und interessieren uns für Vektoren, deren Richtung durch die Abbildung nicht ändert, d.h. sie werden nur um einen Faktor skaliert.

Solche Vektoren heißen Eigenvektoren. Der Faktor um welcher der EV skaliert wird heißt Eigenwert.

Formell:  $\lambda \in \mathbb{K}$  heißt Eigenwert der linearen Abbildung  $F: V \rightarrow V$ , falls es einen Eigenvektor  $v \in V, v \neq 0$  gibt, so dass  $Fv = \lambda v$  gilt.

Ist  $\lambda$  ein EW von  $F$ , so ist der dazugehörige Eigenraum  $E_\lambda$  gleich der um den Nullvektor erweiterten Menge der Eigenvektoren zu  $\lambda$ :

$$E_\lambda := \{v \in V \mid Fv = \lambda v\}$$

Die Menge aller Eigenvektoren von  $F$  heißt Spektrum von  $F$  und wird mit  $\sigma(F)$  bezeichnet.

Die Eigenwerte der linearen Abbildung  $F$  entsprechen den Eigenwerten der Abbildungsmatrix  $A$ .

### Berechnung von Eigenwerten:

Sei  $A$  die Abbildungsmatrix der linearen Abbildung  $F$ .

- 1) Berechne das charakteristische Polynom  $\chi_A := \det(A - \lambda I)$  (d.h. setze auf der Diagonalen überall  $-\lambda$ , berechne Determinante)
- 2) Berechne die Nullstellen von  $\chi_A$  (Achtung: NS können mehrmals vorkommen)  $\rightarrow$  NS sind Eigenwerte.
- 3) Notiere zu jedem EW, wie oft er vorkommt  $\rightarrow$  algebraische Vielfachheit. (Wie oft kommt  $\lambda$  als NS vor)

### Berechnung von Eigenvektoren:

- 1) Berechne alle EW.
- 2) Bestimme zu jedem  $\lambda_i$  den dazugehörigen Eigenraum  $E_{\lambda_i}$ ,  $E_{\lambda_i} = \ker(A - \lambda_i \cdot I)$ . Berechne den Kern, d.h. finde alle  $x$  für die gilt:  $(A - \lambda_i I) \cdot x = 0$ .

Die geometrische Vielfachheit von  $\lambda$  ist dann  $= \dim(E_\lambda)$ , d.h. wie viele EV der EW  $\lambda$  hat.

Beachte: Algebraische Vielfachheit kann, muss aber nicht der algebraischen Vielfachheit entsprechen!

Beachte: EV zu verschiedenen EW sind immer lin. unabhängig.

Charakteristische Gleichung:  $\chi_A(\lambda) = 0$

Spur: Summe der Diagonalelemente von A.

Eine quadratische Matrix A ist genau dann singular, wenn sie 0 als EW hat: A ist singular  $\Leftrightarrow 0 \in \sigma(A)$

Zwei Matrizen A, B heissen ähnlich wenn sie bei

Verwendung unterschiedlicher Basen die gleiche lineare Abbildung beschreiben, d.h. zwei quadratische Matrizen

$A, B \in \mathbb{E}^{n \times n}$  heissen ähnlich wenn es eine reguläre Matrix  $S \in \mathbb{E}^{n \times n}$  gibt so dass

$$B = S^{-1}AS \quad \text{bzw.} \quad SB = AS.$$

Im Falle des kommutativen Diagramms (Seite 3) gilt dann auch  $A \mapsto B = S_B^{-1} \cdot A \cdot T_B^{-1}$ , was auch als Ähnlichkeits-  
transformation bezeichnet wird.

Ähnliche Matrizen haben das gleiche charakteristische Polynom, gleiche Determinanten, gleiche Spur und gleiche Eigenwerte. Die geometrische & algebraische Vielfachheit ist dann ebenfalls gleich.

Eine Basis von A, die nur aus Eigenvektoren besteht, nennt man Eigenbasis von A.

### Eigenwertzerlegung:

Sei  $F: V \rightarrow V$  eine lineare Abbildung mit Abbildungsmatrix A. Sei  $\dim(V) = n$ .  $\lambda_1, \dots, \lambda_k$  seien die k Eigenw. und  $v_1, \dots, v_k$  die EV. Falls nun eine (oder beide) der folgenden Aussagen zutrifft, ist die Matrix A diagonalisierbar.

1)  $k = n$  und alle n EW sind verschieden.

2) Für alle EW & dazugeh. EV gilt: algebraische V. = geometrische V.

Gilt keine dieser Aussagen ist A nicht diagonalisierbar.

Fall 1:  $n = k$ , alle n EW verschieden:

Da alle n EV linear unabhängig sind, bilden diese n Vektoren eine Eigenbasis V. Daraus folgt für  $1 \leq i \leq n$ :

$$Av_i = \lambda_i v_i = v_i \cdot \lambda_i \quad \text{und somit}$$

$$A \cdot \left( v_1 \mid v_2 \mid \dots \mid v_n \right) = \left( v_1 \mid v_2 \mid \dots \mid v_n \right) \cdot \begin{pmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \ddots \\ & & & \lambda_n \end{pmatrix}$$

und da  $v_1, \dots, v_n$  linear unabhängig sind kann man diese invertieren:

$A \cdot V = V \cdot \Lambda \rightsquigarrow A = V \cdot \Lambda \cdot V^{-1}$ , wobei man  $A = V \cdot \Lambda \cdot V^{-1}$  die Eigenwertzerlegung bzw. Spektralzerlegung von A nennt.

Fall 2: algebraische V. = geometrische V.

1) Erweitere die k EV um  $n-k$  (linear unabhängigen) EV zu einer Eigenbasis.

2) Berechne  $\Lambda = V^{-1}AV$

3)  $A = V \cdot \Lambda \cdot V^{-1}$  ist die Eigenwertzerlegung

### Eigenwerte und Eigenvektoren symmetrischer und hermitescher Matrizen

Die meisten Eigenwertprobleme die in der Praxis vorkommen sind selbstadjungiert, d.h. die Matrizen sind symmetrisch bzw. hermitesch.

In diesem Fall wird viel vereinfacht:  $A \in \mathbb{C}^{n \times n}$  Hermitesch:

1) Alle Eigenwerte  $\lambda_1, \dots, \lambda_n$  sind reell

2) Die EV zu versch. EW sind paarweise orthogonal in  $\mathbb{C}^n$ .

3) Es gibt eine orthonormale Basis aus EV  $u_1, \dots, u_n$  von A.

4) Für unitäre Matrix  $U := (u_1, \dots, u_n)$  gilt:  $U^H A U = \Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$  8



Ist  $A \in \mathbb{R}^{n \times n}$  symmetrisch, so gilt:

- 1) Alle Eigenwerte  $\lambda_1 \dots \lambda_n$  sind reell.
- 2) Die EV zu verschiedenen EW sind paarweise orthogonal.
- 3) Es gibt eine orthonormale Basis des  $\mathbb{R}^n$  aus EV  $u_1 \dots u_n$  von  $A$ .
- 4) Für die orthogonale Matrix  $U = (u_1 \dots u_n)$  gilt:  
 $U^T A U = \Lambda = \text{diag}(\lambda_1 \dots \lambda_n)$

### Die Spektralnorm

Die Spektralnorm entspricht geometrisch dem Grösstmöglichen Streckungsfaktor, der durch Anwendung der Matrix auf einen Vektor der Länge 1 entsteht.

Die Spektralnorm einer Matrix  $A \in \mathbb{E}^{n \times n}$  ist gleich der Wurzel aus dem grössten Eigenwert aus  $A^H A$ :

$$\|A\|_2 := \max\{\sqrt{\lambda}, \lambda \text{ Eigenwert von } A^H A\}$$

Ist  $A$  reell gilt natürlich  $A^H A = A^T A$ .

Die Spektralnorm einer Hermiteschen (oder reell symmetrischen) Matrix  $A \in \mathbb{E}^{n \times n}$  ist gleich dem Betrag des grössten Eigenwerts dieser Matrix:

$$\|A\|_2 = \max(|\lambda|, \lambda \text{ Eigenwert von } A)$$

Bei einer Unitären (bzw. orthogonalen) Matrix  $U$  gilt wegen  $U^H U = I_n$  stets  $\|U\|_2 = 1$ .

Die Spektralnorm der Inversen  $A^{-1}$ :

$$\|A^{-1}\|_2 = \max\left\{\frac{1}{\sqrt{\lambda}}, \lambda \text{ Eigenwert von } A^H A\right\}$$

Ist  $A$  Hermitesch oder reell:

$$\|A^{-1}\|_2 = \max\left\{\frac{1}{|\lambda|}, \lambda \text{ Eigenwert von } A\right\}$$

Die 2-Norm-Konditionszahl einer Matrix  $A \in \mathbb{E}^{n \times n}$  ist  
$$K_2(A) = \frac{\max\{\sqrt{\lambda}, \lambda \text{ ist EW von } A^H A\}}{\min\{\sqrt{\lambda}, \lambda \text{ ist EW von } A^H A\}}$$

Ist  $A$  hermitesch oder reell gilt:

$$K_2(A) = \frac{\max\{|\lambda|, \lambda \text{ EW von } A\}}{\min\{|\lambda|, \lambda \text{ EW von } A\}}$$

Ist  $A$  singulär, so ist  $K_2(A) = \infty$ .

Hilbert-Matrix:

$$H_n := \begin{pmatrix} 1 & 1/2 & 1/3 & \dots & 1/n \\ 1/2 & 1/3 & 1/4 & \dots & 1/(n+1) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1/n & 1/(n+1) & 1/(n+2) & \dots & 1/(2n-1) \end{pmatrix}$$
 Hat sehr hohe Konditionszahl, d.h. schlechte Kondition.

### Definitheit von Matrizen

Sei  $A$  eine  $n \times n$ -Matrix und  $x$  ein  $n$ -zeiliger Spaltenvektor  $x \in V, x \neq 0$ .  $A$  ist:

positiv definit, falls  $x^T A x > 0$

positiv semidefinit, falls  $x^T A x \geq 0$

negativ definit, falls  $x^T A x < 0$

negativ semidefinit, falls  $x^T A x \leq 0$

Kriterien:

positiv definit  $\rightarrow$  Alle EW  $> 0$

positiv semidefinit  $\rightarrow$  Alle EW  $\geq 0$

negativ definit  $\rightarrow$  Alle EW  $< 0$

negativ semidefinit  $\rightarrow$  Alle EW  $\leq 0$

indefinit  $\rightarrow$  positive & negative EW existieren.

## Singularwertzerlegung (SVD)

Die Singularwertzerlegung einer Matrix zerlegt beliebige (d.h. auch nicht-quadratische) Matrizen in ein Produkt dreier spezieller Matrizen, so dass gilt:

$$\text{Singularwertzerlegung: } A = U \Sigma V^H$$

wobei  $U$  und  $V$  orthogonal / unitär und  $\Sigma$  diagonal sind. Die Größen der Matrizen sind dabei:

$A = m \times n$  - Matrix,  $r = \text{Rang von } A$

$U = m \times m$  - Matrix

$\Sigma = m \times n$  - Matrix

$V = n \times n$  - Matrix

Um diese (komplette) SVD zu berechnen, mache folgendes:

- 1) Berechne die  $n \times n$  - Matrix  $A^T A$ .
- 2) Berechne die EW inkl. alg. Vielfalt von  $A^T A$ , bzw. da  $A^T A$  orthogonal ist die Eigenwertzerlegung  $A^T A = V \Lambda V^H$  wobei  $V$  ja die EV und  $\Lambda$  die EW von  $A^T A$  enthält. (also ganz normale Eigenwertzerlegung).
- 3) Der Singularwert  $\sigma_i$  ist die Wurzel des  $i$ -ten EW  $\lambda_i$  von  $A^T A$ . Ordne diese Werte so, dass  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > \sigma_{r+1} = \dots = \sigma_n = 0$  gilt. Ordne die ersten  $r$  Werte in der Matrix  $\Sigma_r = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_r)$  an, und erhalte so die  $r \times r$  - Matrix  $\Sigma_r$ . Um  $\Sigma$  zu erhalten, erweitere diese bis sie die Größe  $m \times n$  hat und fülle den Rest mit 0 (falls  $r = n$  hat man  $\Sigma_r = \Sigma$ ).
- 4) Berechne die Matrix  $U_r = A V_r \Sigma_r^{-1}$ , wobei  $V_r$  die ersten  $r$  Spaltenvektoren von  $V = (V_r | V_\perp)$  sind.

Man kann die einzelnen  $u_i = \frac{1}{\sigma_i} A v_i$  berechnen und  $U_r$  so erstellen.

Erweitere  $U_r$  dann zu einer orthogonalen  $m \times m$  - Matrix  $U$ .

$$\Rightarrow A = U \Sigma V^H$$

Bei der ökonomischen bzw. Reduzierten SVD gilt:

I) Für  $m < n \rightarrow A = U \Sigma_m V_m^H$

II) Für  $m > n \rightarrow A = U_n \Sigma_n V^H$

Singularwerte von  $A$ : alle  $\sigma_i$

Linke Singulärvektoren: Vektoren  $u_1 \dots u_m$

Rechte Singulärvektoren: Vektoren  $v_1 \dots v_n$

## Folgerungen & Anwendungen der SVD

$R(A)$  und  $\mathcal{N}(A^H)$  sind zueinander orthogonal und spannen den Bildraum  $\mathbb{E}^m$  auf.

$R(A^H)$  und  $\mathcal{N}(A)$  sind zueinander orthogonal und spannen den Bildraum  $\mathbb{E}^n$  auf.

Es sei  $A \in \mathbb{E}^{m \times n}$ ,  $\text{Rang } A = r$ . Dann sind die  $r$  positiven EW von  $A^H A$  und  $A A^H$  gleich, aber die Vielfachheit des EW 0 ist  $n-r$  bzw.  $m-r$ .

Für die Spektralnorm einer Matrix  $A \in \mathbb{E}^{n \times n}$  gilt:

$$\|A\|_2 = \sigma_1. \text{ Ist } A \text{ regulär, so gilt für die Konditionszahl}$$

$$\kappa_2(A) = \frac{\sigma_1}{\sigma_n}.$$

## Hauptkomponentenanalyse:

„Reduktion“ der Datenmenge durch Weglassen „unwichtiger“

Komponenten:

$$A = \begin{pmatrix} | & | & \dots & | \\ a_1 & a_2 & \dots & a_n \\ | & | & \dots & | \end{pmatrix}, a_j \in \mathbb{R}^m, m \text{ ist „gross“ d.h. } m \gg n, \text{ „hohe“ Matrix } A.$$

$a_j$  „leben in der Nähe“ von einem Unterraum von  $\mathbb{R}^m$  mit niedriger Dimension.

$$A = U \cdot \begin{pmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \dots & \\ & & & \sigma_k & & \\ & & & & \dots & \\ & & & & & \dots & \\ & & & & & & 0 \end{pmatrix} \cdot V^T \approx U \cdot \begin{pmatrix} \sigma_1 & & & \\ & \sigma_k & & \\ & & \dots & \\ & & & \dots & \\ & & & & \dots & \\ & & & & & \dots & \\ & & & & & & 0 \end{pmatrix} \cdot V^T$$

$= \sum_{j=1}^k u_j \sigma_j v_j^T$  wobei  $\{u_1, \dots, u_k\}$  die Basis für den reduzierten Unterraum ist:

$$\begin{pmatrix} | & \dots & | \\ a_1 & \dots & a_n \\ | & \dots & | \end{pmatrix} \approx \underbrace{\begin{pmatrix} | & \dots & | \\ u_1 & \dots & u_k \\ | & \dots & | \end{pmatrix}}_{m \times k} \underbrace{\begin{pmatrix} \sigma_1 & & & \\ & \sigma_k & & \\ & & \dots & \\ & & & 0 \end{pmatrix}}_{k \times n} \underbrace{\begin{pmatrix} | & \dots & | \\ v_1 & \dots & v_n \\ | & \dots & | \end{pmatrix}}_{n \times n}$$

Basisvektoren  
vom Unterraum

Koordinaten im  
Unterraum

$$\Rightarrow \tilde{a}_j := a_j - \underbrace{\frac{1}{n} \sum_{t=1}^n a_t}_{\text{Durchschnitt}} \Rightarrow \tilde{A} = \begin{pmatrix} | & \dots & | \\ \tilde{a}_1 & \dots & \tilde{a}_n \\ | & \dots & | \end{pmatrix}$$

Durchschnitt

Anwendung z.B. in der Gesichtserkennung.

## Cauchy-Schwarz-Ungleichung Beweis

$$|\langle x, y \rangle| \leq \|x\| \cdot \|y\|$$

Beweis: Für beliebiges  $\alpha \in \mathbb{E}$  ist

$$0 \leq \langle \alpha x + y, \alpha x + y \rangle = \alpha \bar{\alpha} \langle x, x \rangle + \bar{\alpha} \langle x, y \rangle + \alpha \overline{\langle x, y \rangle} + \langle y, y \rangle$$

Für  $x=0$  gilt die Ungleichung, und zwar mit dem Gleichheitszeichen. Wir können  $x \neq 0$  annehmen und  $\alpha = -\frac{\langle x, y \rangle}{\langle x, x \rangle}$  wählen,

womit nach Multiplikation mit  $\langle x, x \rangle$  folgt:

$$0 \leq |\langle x, y \rangle|^2 - |\langle x, y \rangle|^2 - |\langle x, y \rangle|^2 + \langle x, x \rangle \langle y, y \rangle$$

was gleich dem Quadrat der CSU ist:

$$|\langle x, y \rangle|^2 \leq \langle x, x \rangle \langle y, y \rangle, \text{ d.h. die CSU ist bewiesen.}$$