Lineare Algebra III – Singulärwertzerlegung

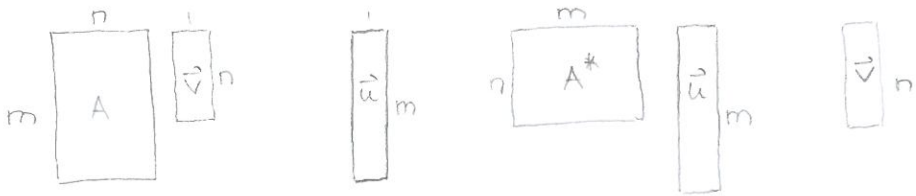
Mit der sogenannten Singulärwertzerlegung wollen wir heute noch einmal ein numerisches Verfahren aus der linearen Algebra kennenlernen.

Definition 1

Gegeben sei eine komplexe Matrix A mit m Reihen und n Spalten. Wir nennen σ einen Singulärwert von A, wenn σ nicht negativ ist und es zwei von $\emptyset$ verschiedene Vektoren $\vec{u}$ (m-Vektor) und $\vec{v}$ (n-Vektor) gibt, so daß

Adjungation $a_{ij}^* = \overline{a_{ji}}$

$$A\vec{v} = \sigma \cdot \vec{u} \qquad A^* \vec{u} = \sigma \cdot \vec{v}$$



$\vec{u}$ und $\vec{v}$ heißen linker und rechter Singulärvektor.

Definition 2

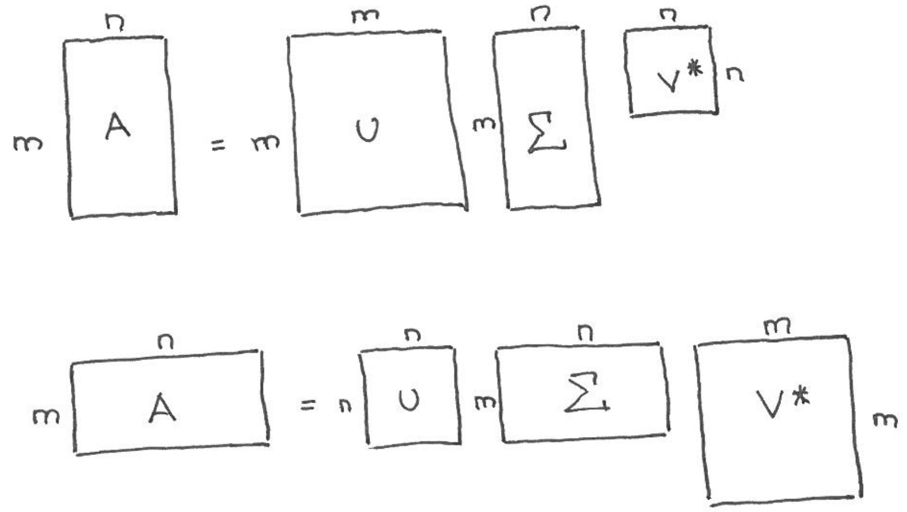
Gegeben sei eine komplexe Matrix A mit m Reihen und n Spalten. Das Matrixprodukt

$$\boxed{A = U \cdot \Sigma \cdot V^*}$$

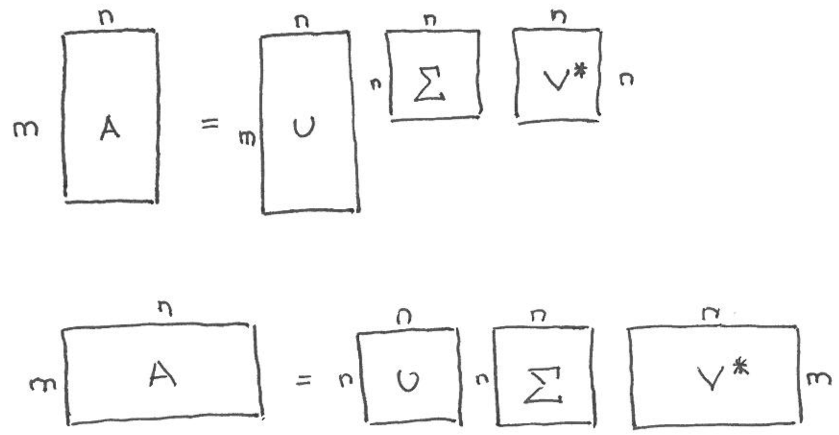
ist eine Singulärwertzerlegung von A wenn U und V jeweils orthonormale Spalten besitzen und Σ nur nicht-negative Elemente auf seiner Diagonalen besitzt und überall sonst Nullen.

Es gibt 3 verschiedene Formen der Singulärwertzerlegung

Form 1 A und Σ haben dieselben Dimensionen.

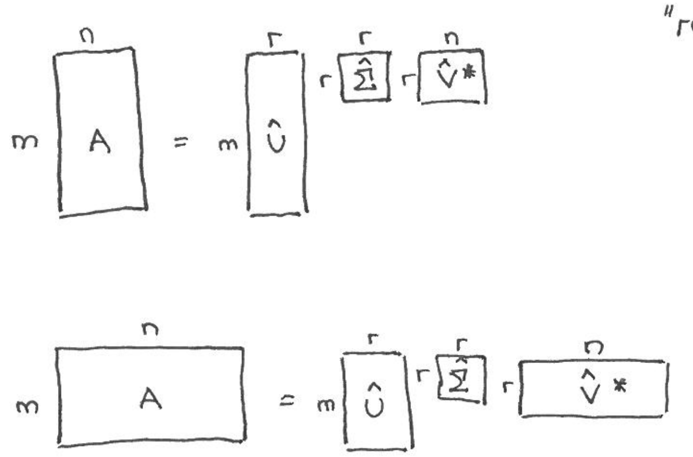


Form 2 Σ ist quadratisch.



Form 3 $\hat{\Sigma}$ ist quadratisch und enthält nur positive Singulärwerte (keine Nullen)

"reduzierte Singulärwertzerlegung"



Berechnung der Singulärwertzerlegung

Theoretisch ließe sich die Singulärwertzerlegung der Matrix A auf die Eigenwertzerlegung ihrer Kovarianzmatrix $A^* A$ abbilden, denn es gilt

$$A^* A = V \Sigma^* \Sigma V^*$$

Ein Algorithmus könnte also so aussehen, daß wir die folgenden Schritte durchführen

1. Bilde die Kovarianzmatrix $A^* A$
2. Berechne deren Eigenwertzerlegung $A^* A = V \Lambda V^*$
3. Σ ist eine $m \times n$ Matrix, deren Diagonaleinträge die positiven Wurzeln von Λ sind
4. Löse $U \Sigma = A \cdot V$ für unitäres U

In der Praxis wird dieser Algorithmus allerdings nicht benutzt, da die Abbildung auf dieses Eigenwertproblem zusätzliche numerische Instabilitäten mit sich bringt, die sich vermeiden lassen.

Nehmen wir der Einfachheit halber an, daß A quadratisch ($m = n$) sei. Betrachte nun die hermitesche $2m \times 2m$ Matrix

$$H = \begin{pmatrix} 0 & A^* \\ A & 0 \end{pmatrix}$$

$(A^* = V \Sigma^* U^*)$

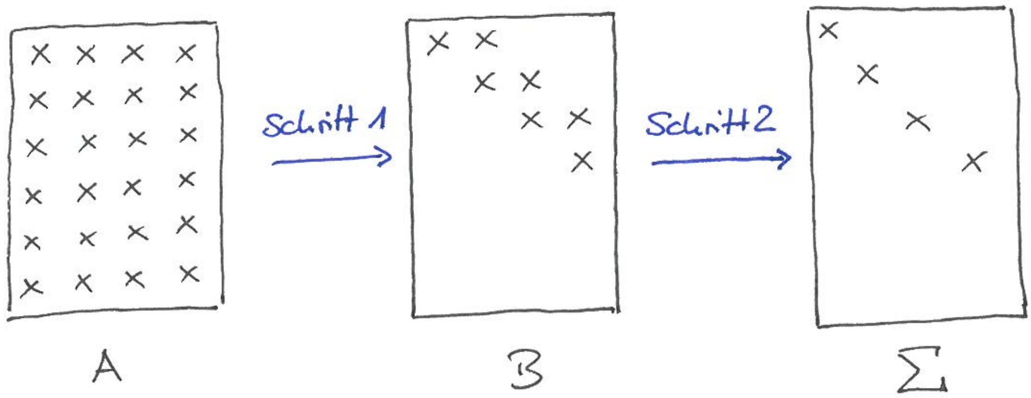
Wegen $A = U \Sigma V^* \rightarrow AV = U\Sigma$ und $A^* U = V \Sigma^* = V \Sigma$ gilt

$$\begin{pmatrix} 0 & A^* \\ A & 0 \end{pmatrix} \begin{pmatrix} V & V \\ U & -U \end{pmatrix} = \begin{pmatrix} V & V \\ U & -U \end{pmatrix} \begin{pmatrix} \Sigma & 0 \\ 0 & -\Sigma \end{pmatrix}$$

was einer Eigenwertzerlegung von H entspricht. Die Singulärwerte von A entsprechen den absoluten Eigenwerten von H und die Singulärvektoren von A lassen sich aus den Eigenvektoren von H konstruieren.

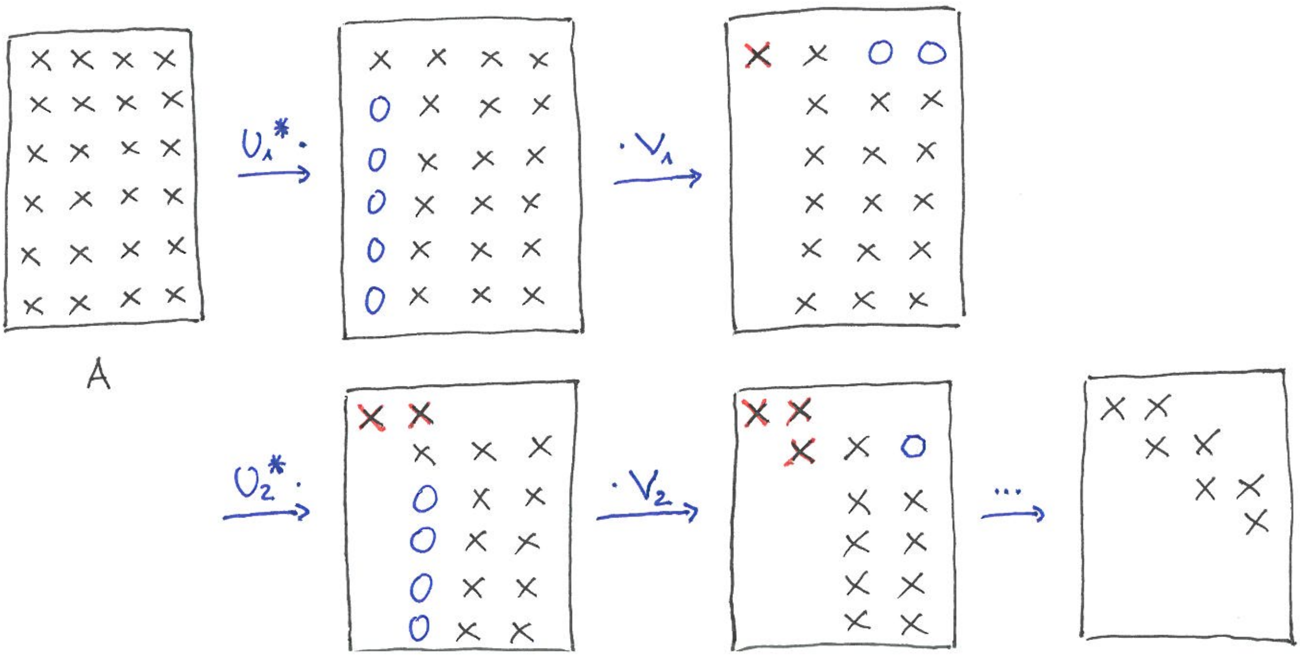
Die Standardalgorithmen zur Berechnung der Singulärwertzerlegung basieren auf dieser zweiten Abbildung auf ein Eigenwertproblem, welches numerisch stabil ist. Die tatsächlichen Verfahren sind dabei derart elegant formuliert, daß die $2m \times 2m$ Matrix H niemals explizit berechnet werden muß.

Typischerweise gehen die Algorithmen in zwei Schritten vor.
Im ersten Schritt wird die Matrix A auf Bidiagonalform gebracht, und im zweiten Schritt die bidiagonale Matrix diagonalisiert.



Schritt 1: Bidiagonalisierung.

Die einfachste Weise den ersten Schritt der Bidiagonalisierung zu sehen ist der Golub-Kahan Algorithmus, der aus einer Sequenz von Householder-Spiegelungen besteht.



Die benutzten Householder-Spiegelungen wollen wir an dieser Stelle nur ganz kurz besprechen, für eine längere Beschreibung sei etwa auf das Buch von Trefethen & Bau verwiesen.

Ziel der Householder-Transformation ist die Spiegelung einer Spalte (oder Zeile) der Matrix A auf ein Vielfaches des entsprechenden Einheitsvektors, deren Berechnung wie folgt gegeben ist

$$\boxed{U_1^* = H_1 = \mathbb{1} - 2 \cdot \frac{\vec{u} \cdot \vec{u}^T}{\vec{u}^T \cdot \vec{u}}}$$

Einheitsvektor

wobei $\vec{u} = \vec{a} + \delta \cdot \|\vec{a}\| \cdot \hat{e}$ mit $\delta = \begin{cases} +1 & \text{für } a_1 > 0 \\ -1 & \text{für } a_1 < 0. \end{cases}$

Spaltenvektor von A

Schritt 2: Diagonalisierung

Für die Diagonalisierung der bidiagonalen Matrix B stehen eine Reihe von Algorithmen zur Verfügung. Bis in die 90er Jahre wurde eine Variante des sogenannten QR-Verfahrens verwandt. Heute benutzt man sogenannte divide-and-conquer Algorithmen, um diese Diagonalisierung besonders schnell berechnen zu können.

Alternatives Verfahren – Jacobi Transformationen

Ein alternatives, etwas weniger effizientes Verfahren zur Singulärwertzerlegung beruht auf sogenannten Jacobi-Rotationen. Auch dieses Verfahren sei hier kurz skizziert.

Zentrales Element des Jacobi-Verfahrens ist die gezielte Elimination eines Matrixelements a_{pq} der Matrix A via eine Jacobi-Transformation der Form

$$P_{pq} = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & c & \cdots & s & & \\ & & \vdots & 1 & \vdots & & \\ & & -s & \cdots & c & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}$$

Eine solche Transformation enthält eine 1 auf der Hauptdiagonalen und ist sonst überall 0 bis auf 4 zusätzliche Einträge

$$(p,p) = (q,q) = c = \cos\phi \qquad \begin{pmatrix} pp & pq \\ qp & qq \end{pmatrix} = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix}$$
$$(p,q) = s \quad (q,p) = -s = -\sin\phi$$

Das heißt wir nehmen eine Rotation um einen noch zu bestimmenden Winkel ϕ vor, welche auf diesen 2×2 Subblock der Matrix angewandt wird. Daraus ergibt sich folgende Ähnlichkeitstransformation

$$P_{pq}^t \cdot A \cdot P_{pq} = A' \qquad (*)$$

Dabei ändern sich nur die Elemente in der p-ten und q-ten Zeile bzw. Spalte.

Speziell gilt für die folgenden Elemente

$$a'_{pp} = c^2 \cdot a_{pp} + s^2 \cdot a_{qq} - 2sc \cdot a_{pq}$$

$$a'_{qq} = c^2 \cdot a_{pp} + s^2 \cdot a_{qq} + 2sc \cdot a_{pq}$$

$$a'_{pq} = (c^2 - s^2)\, a_{pq} + sc \cdot (a_{pp} - a_{qq})$$

Um das Matrixelement a'_{pq} zu eliminieren, wollen wir den Winkel ϕ so wählen, das gilt:

$$a'_{pq} = 0 \quad \Leftrightarrow \quad \frac{a_{qq} - a_{pp}}{a_{pq}} = \frac{\cos^2\phi - \sin^2\phi}{\sin\phi \, \cos\phi} = \frac{1 - \tan^2\phi}{\tan\phi}$$

$$\Leftrightarrow \quad \frac{a_{qq} - a_{pp}}{2 \cdot a_{pq}} = \frac{1 - \tan^2\phi}{2 \cdot \tan\phi} = \cot 2\phi$$

$$\rightarrow \quad \phi = \frac{1}{2} \operatorname{arc\,cot}\left(\frac{a_{qq} - a_{pp}}{2 a_{qp}}\right)$$

Mit einer solchen Transformation läßt sich also a'_{pq} eliminieren. Das Element a'_{qp} verschwindet simultan nur dann wenn auch vorher schon $a_{pq} = a_{qp}$ gegolten hat. Ist dies nicht der Fall muß die Ähnlichkeitstransformation (*) ein wenig angepaßt werden

$$c = \cos\phi \qquad s = \sin\varphi$$

so daß beide Elemente simultan verschwinden.

Wendet man nun eine Sequenz von Ähnlichkeitstransformationen an, so erhält man eine Diagonalform, die genau der Singulärwertmatrix Σ entspricht

$$V = P^1 \cdot P^2 \cdot P^3 \ldots P^k$$

$$\rightarrow \quad \Sigma = V^t \cdot A \cdot V$$

Eine exakte Diagonalform erhält man nur nach sehr hoher Zahl k von Ähnlichkeitstransformationen, da ein auf 0 gesetztes Matixelement in späteren Transformationen [illegible] eine kleine endlichen Wert [illegible]