Fachbereich Mathematik
Prof. Dr. P. Spellucci
Wintersemester 2007/2008

TECHNISCHE
UNIVERSITÄT
DARMSTADT

28.1.2008

Numerik des Matrizeneigenwertproblems
Übung 13

Präsenzübung

Ü 37 Zeigen Sie: jede komplexe $m \times n$-Matrix A kann faktorisiert werden als

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

mit einer unitären $m \times m$–Matrix U, einer $m \times n$–Matrix S mit $s_{1,1} \;\geq\; \ldots \geq s_{r,r} \;>\; 0$, $r =$Rang(A) und $s_{i,j} \;= 0$ sonst und einer unitären $n \times n$–Matrix V^H. (Diese Zerlegung wird als Singulärwertzerlegung bezeichnet, `svd`)
Hinweis: Nehmen Sie o.b.d.A. $m \;\geq\; n$ und betrachten Sie zunächst die Spektralzerlegungen von $A^H A$ und AA^H

Ü 38 Sei A eine beliebige komplexe $m \times n$-Matrix. Eine $n \times m$–Matrix $A^\#$ mit den Eigenschaften

$$\begin{aligned} A^\# A &= (A^\# A)^H \\ AA^\# &= (AA^\#)^H \\ A^\# A A^\# &= A^\# \\ AA^\# A &= A \end{aligned}$$

heisst Moore–Penrose–Pseudoinverse von A. Man kann zeigen, daß sie eindeutig bestimmt ist. Zeigen Sie: Ist

$$A \;=\; USV^H$$

eine Singulärwertzerlegung von A, dann gilt

$$A^\# \;=\; VS^\# U^H$$

wobei $S^\#$ aus S entsteht durch Transposition und Ersetzung der Nichtnullwerte σ_i durch ihre Reziprokwerte.

Ü 39 Es sei

$$A = \frac{1}{2}\begin{pmatrix} 1 & 1 & -1 & 1 \\ 1 & -1 & -1 & -1 \\ 1 & 1 & 1 & -1 \\ 1 & -1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0.8 & 0 & 0.6 \\ 0 & 1 & 0 \\ 0.6 & 0 & -0.8 \end{pmatrix} = U\Sigma V^T$$

eine Singulärwertzerlegung von A.

Geben Sie explizit $x \in \mathbb{R}^4$ und $y \in \mathbb{R}^3$ an, so dass $\|xy^T\|_2$ minimal ist und Rang$(A + xy^T) = 2$.

Hausübung

H 36 Zeigen Sie:
Ist A eine beliebige reelle $m \times n$–Matrix mit $m \geq n$ und $\alpha > 0$, dann ist

$$x(\alpha) = (\alpha I + A^T A)^{-1} A^T b$$

wohldefiniert und es existiert

$$x^* = \lim_{\alpha \to 0^+} x(\alpha) \ .$$

Ferner gilt:

$$\|x^*\|_2 = \min\{\|y\|_2 : \|Ay - b\|_2 \leq \|Az - b\|_2 \text{ für alle } z \in \mathbb{R}^n\} \ .$$

Lässt sich $x(\alpha)$ auch als Lösung einer linearen Ausgleichsaufgabe darstellen und so eventuell numerisch besser berechnen?

H 37 Es sei $A \in \mathbb{R}^{m \times n}$ mit $m \geq n$ und $b \in \mathbb{R}^m$. Zeigen Sie, daß die allgemeine Lösung der linearen Ausgleichsaufgabe

$$\text{Minimiere } \|Ax - b\|_2$$

in der Form

$$x^* = A^\# b + (I - A^\# A)z, \quad z \in \mathbb{R}^n \text{ beliebig}$$

geschrieben werden kann.

Hinweis: Es sei $A = U\binom{\Sigma}{0}V^T$ die (SVD) von A und $A^\# = V(\Sigma^\#, 0)U^T$, wobei die Elemente der Diagonalmatrix $\Sigma^\#$ entweder die Kehrwerte der Singulärwerte σ_i sind, falls $\sigma_i \neq 0$, oder 0 falls $\sigma_i = 0$.

H 38 $U\Sigma V^H$ sei die Singulärwertzerlegung der Matrix $A \in \mathbb{C}^{m,n}$ mit den singulären Werten $\sigma_1 \geq \sigma_2 \geq \ldots \geq 0$. Daraus bilde man die Matrizen Σ_r und $A_r := U\Sigma_r V^H$, indem man nur die ersten r singulären Werte beibehält, also σ_k durch Null ersetzt für alle $k > r$.

a) Was sind die singulären Werte von $A - A_r$ und berechne $||A - A_r||_2$.
Zeige Rang$A_r \leq r$.

b) Es gibt keine Matrix B vom Rang höchstens r mit $||A - B||_2 < \sigma_{r+1}$.

Numerik des Matrizeneigenwertproblems
Übung 13, Lösungsvorschlag

Präsenzübung

Ü 37 Zeigen Sie: jede komplexe $m \times n$-Matrix A kann faktorisiert werden als

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

mit einer unitären $m \times m$–Matrix U, einer $m \times n$–Matrix S mit $s_{1,1} \geq \ldots \geq s_{r,r} > 0$, $r =$Rang(A) und $s_{i,j} = 0$ sonst und einer unitären $n \times n$–Matrix V^H. (Diese Zerlegung wird als Singulärwertzerlegung bezeichnet, `svd`)
Hinweis: Nehmen Sie o.b.d.A. $m \geq n$ und betrachten Sie zunächst die Spektralzerlegungen von $A^H A$ und AA^H

Die Matrizen $A^H A$ und AA^H sind beide hermitisch und positiv semidefinit wegen

$$x^H A^H A x = ||Ax||_2^2 \quad \text{und} \quad x^H A A^H x = ||A^H x||_2^2$$

Mit der Spektralzerlegung von $A^H A$ kann man schreiben

$$A^H A = V \text{diag}(\sigma_1^2, \ldots, \sigma_n^2) V^H$$

wobei die Eigenwerte wegen ihrer Positivität gleich als Quadrate geschrieben wurden. Es ist Rang($A^H A$) = Rang(A) = Rang(diag($\sigma_1^2, \ldots, \sigma_n^2$)) und o.B.d.A. kann man die σ_i sortiert annehmen. Ist nun

$$AA^H y = \lambda y \,, \quad y \neq 0$$

dann gilt

$$A^H y = 0 \Leftrightarrow \lambda = 0$$

und dieses lineare Gleichungssystem hat $m-$Rang(A) linear unabhängige Lösungen y. Ist aber $A^H y \neq 0$, dann

$$AA^H y = \lambda y \Rightarrow A^H A (A^H y) = \lambda (A^H y)$$

d.h. λ ist auch ein Eigenwert von $A^H A$, also unter den σ_i^2 und $A^H y$ ein Eigenvektor dazu. Die Eigenvektormatrix U von AA^H lässt sich unitär wählen und so, daß

$$U^H A A^H U = \begin{pmatrix} \Sigma^2 & O \\ O & O \end{pmatrix}$$

Dabei sei

$$\Sigma^2 = \text{diag}(\sigma_1^2, \ldots, \sigma_n^2)$$

Mit

$$B = A^H U$$

gilt also : die ersten r Spalten von B sind paarweise orthogonal und haben die Längen σ_i, die übrigen sind null. Also kann man schreiben (mit $\sigma_{r+1} = \ldots \sigma_n = 0$)

$$B = (V\Sigma, O)$$

wobei V die Eigenvektormatrix von $A^H A$ ist. Schliesslich wird

$$A = UB^H = U \begin{pmatrix} \Sigma \\ O \end{pmatrix} V^H$$

Wir identifizeiren nun die σ_i mit den $s_{i,i}$ aus der Aufgabenstellung.

Ü 38 Sei A eine beliebige komplexe $m \times n$-Matrix. Eine $n \times m$–Matrix $A^\#$ mit den Eigenschaften

$$\begin{aligned} A^\# A &= (A^\# A)^H \\ AA^\# &= (AA^\#)^H \\ A^\# AA^\# &= A^\# \\ AA^\# A &= A \end{aligned}$$

heisst Moore–Penrose–Pseudoinverse von A. Man kann zeigen, daß sie eindeutig bestimmt ist. Zeigen Sie: Ist

$$A = USV^H$$

eine Singulärwertzerlegung von A, dann gilt

$$A^\# = VS^\# U^H$$

wobei $S^\#$ aus S entsteht durch Transposition und Ersetzung der Nichtnullwerte σ_i durch ihre Reziprokwerte.

Wir beachten, daß $S^\# S$ und $SS^\#$ beides Diagonalmatrizen sind mit 1 auf den ersten r Positionen und null sonst. Wir setzen die Darstellung von $A^\#$ in die Bedingungen ein und multiplizieren aus:

$$\begin{aligned} A^\# A &= VS^\# U^H USV^H = VS^\# SV^H = (VS^\# SV^H)^H \\ AA^\# &= USV^H VS^\# U^H = USS^\# U^H = (USS^\# U^H)^H \\ AA^\# A &= USV^H VS^\# U^H USV^H = USS^\# SV^H = USV^H = A \\ A^\# AA^\# &= VS^\# U^H USV^H VS^\# U^H = VS^\# SS^\# U^H = VS^\# U^H = A^\# \end{aligned}$$

Ü 39 Es sei

$$A = \frac{1}{2} \begin{pmatrix} 1 & 1 & -1 & 1 \\ 1 & -1 & -1 & -1 \\ 1 & 1 & 1 & -1 \\ 1 & -1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0.8 & 0 & 0.6 \\ 0 & 1 & 0 \\ 0.6 & 0 & -0.8 \end{pmatrix} = U\Sigma V^T$$

eine Singulärwertzerlegung von A.

Geben Sie explizit $x \in \mathbb{R}^4$ und $y \in \mathbb{R}^3$ an, so dass $\|xy^T\|_2$ minimal ist und $\text{Rang}(A + xy^T) = 2$.

Es muss $\|A - \tilde{A}\|_2 = \|xy^T\|_2 \geq 1$, *damit der Rang der Matrix* $= 2$ *wird (*1 *ist der betragsmäßig kleinste Singulärwert der Matrix* A*). Mit*

$$\Delta A = U \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix} V^T = -\frac{1}{2} \begin{pmatrix} -1 \\ -1 \\ 1 \\ 1 \end{pmatrix} \begin{pmatrix} 0.6 & 0 & -0.8 \end{pmatrix}$$

wird das Gewünschte offenbar erreicht mit $\|xy^T\|_2 = 1$.

Hausübung

H 36 Zeigen Sie:

Ist A eine beliebige reelle $m \times n$–Matrix mit $m \geq n$ und $\alpha > 0$, dann ist

$$x(\alpha) = (\alpha I + A^T A)^{-1} A^T b$$

wohldefiniert und es existiert

$$x^* = \lim_{\alpha \to 0^+} x(\alpha) \ .$$

Ferner gilt:

$$\|x^*\|_2 = \min\{\|y\|_2 : \|Ay - b\|_2 \leq \|Az - b\|_2 \text{ für alle } z \in \mathbb{R}^n\} \ .$$

Lässt sich $x(\alpha)$ auch als Lösung einer linearen Ausgleichsaufgabe darstellen und so eventuell numerisch besser berechnen?

$A^T A$ ist positiv semidefinit und damit $\alpha I + A^T A$ positiv definit. Damit ist $x(\alpha)$ wohldefiniert.

Singulärwertzerlegung

$$\begin{aligned} A &= U \begin{pmatrix} \Sigma \\ 0 \end{pmatrix} V^T \\ \alpha I + A^T A &= V(\Sigma^2 + \alpha I) V^T \\ x(\alpha) &= V(\Sigma^2 + \alpha I)^{-1} V^T V \begin{pmatrix} \Sigma & 0 \end{pmatrix} U^T b = \\ &= V diag\left(\frac{\sigma_i}{\sigma_i^2 + \alpha}\right) U^T b \end{aligned}$$

Fallunterscheidung:

$$\sigma_i = 0 \quad \Longrightarrow \quad \alpha \ > 0, \ \frac{\sigma_i}{\sigma_i^2 + \alpha} = 0 \quad \Longrightarrow \quad \lim_{\alpha \to 0^+} \ldots = 0$$

$$\sigma_i \neq 0 \quad \Longrightarrow \quad \lim_{\alpha \to 0^+} \frac{\sigma_i}{\sigma_i^2 + \alpha} = \frac{1}{\sigma_i}$$

Damit existiert der Grenzwert und es gilt:

$$\begin{aligned} \lim_{\alpha \to 0^+} x(\alpha) &= V \Sigma^+ U^T b \\ \Sigma^+ &= diag(\sigma_i^+) \\ \sigma_i^+ &= \begin{cases} 0 & \sigma_i = 0 \\ \frac{1}{\sigma_i} & \sigma \neq 0 \end{cases} \end{aligned}$$

Das ist die Lösung des linearen Ausgleichsproblems von minimaler Länge wie im Skript beschrieben und damit ist die Behauptung gezeigt. Offensichtlich ist $x(\alpha)$ Lösung der linearen Ausgleichsaufgabe

$$|| \begin{pmatrix} A \\ \sqrt{\alpha} I \end{pmatrix} x - \begin{pmatrix} b \\ 0 \end{pmatrix} || \stackrel{!}{=} \min_x$$

In dieser Formulierung wird die Bildung von A^TA und die Gleichungslösung mit der eventuell schlecht konditionierte Matrix $A^TA + \alpha I$ vermieden.

H 37 Es sei $A \in \mathbb{R}^{m\times n}$ mit $m \geq n$ und $b \in \mathbb{R}^m$. Zeigen Sie, daß die allgemeine Lösung der linearen Ausgleichsaufgabe

$$\text{Minimiere } \|Ax - b\|_2$$

in der Form

$$x^* = A^\# b + (I - A^\# A)z, \quad z \in \mathbb{R}^n \text{ beliebig}$$

geschrieben werden kann.

Hinweis: Es sei $A = U\binom{\Sigma}{0}V^T$ die (SVD) von A und $A^\# = V(\Sigma^\#, 0)U^T$, wobei die Elemente der Diagonalmatrix $\Sigma^\#$ entweder die Kehrwerte der Singulärwerte σ_i sind, falls $\sigma_i \neq 0$, oder 0 falls $\sigma_i = 0$.

Setzt man in der linearen Ausgleichsaufgabe die (SVD) von A ein ergibt sich

$$\|Ax - b\|_2^2 = \| \begin{pmatrix} \Sigma \\ 0 \end{pmatrix} V^T x - U^T b\|_2^2.$$

Mit $y = V^T x$ und $U^T b = c$ folgt

$$\|Ax - b\|_2^2 = \sum_{i=1}^{r} (\sigma_i y_i - c_i)^2 + \sum_{i=r+1}^{m} c_i^2.$$

Dabei ist $r \leq n$ der Rang der Matrix A, also die Anzahl der nicht verschwindenden Singulärwerte.

Die Lösung y^* hat demnach in den ersten r Komponenten die Werte $y_i^* = \frac{c_i}{\sigma_i}$ und ist in den letzten $n - r$ Komponenten (falls existent) beliebig wählbar da

$$\Sigma = \begin{pmatrix} \Sigma_r & 0 \\ 0 & 0 \end{pmatrix}, \quad \Sigma_r \textit{ invertierbar.}$$

Für die Lösung x^* folgt damit

$$x^* = Vy^* = \sum_{i=1}^{r} v_i \frac{c_i}{\sigma_i} + \sum_{i=r+1}^{n} v_i y_i \quad (y_i \textit{ beliebig}).$$

Die erste Summe ist aber identisch mit $A^\# b$ und um die zweite Summe zu identifizieren untersuchen wir $I - A^\# A$:

$$\begin{aligned}
I - A^\# A &= VV^T - V(\Sigma^\#, 0)\tbinom{\Sigma}{0}V^T \\
&= V\left(I - (\Sigma^\#, 0)\tbinom{\Sigma}{0}\right)V^T \\
&= V\left(I - \left(\begin{array}{c|c} \Sigma_r^{-1} & 0 \\ \hline 0 & 0 \end{array}\right)\left(\begin{array}{c|c} \Sigma_r & 0 \\ \hline 0 & 0 \end{array}\right)\right)V^T \\
&= V\left(\begin{array}{c|c} 0 & 0 \\ \hline 0 & I_{n-r} \end{array}\right)V^T \\
&= (0, \ldots, 0, v_{r+1}, \ldots, v_n)V^T.
\end{aligned}$$

Für beliebiges $z \in \mathbb{R}^n$ ist also

$$(I - A^{\#}A)z = (0, \ldots, 0, v_{r+1}, \ldots, v_n)\underbrace{V^T z}_{=y} = \sum_{i=r+1}^{n} v_i y_i,$$

mit y_i beliebig. Zusammenfassend erhalten wir

$$x^* = \sum_{i=1}^{r} v_i \frac{c_i}{\sigma_i} + \sum_{i=r+1}^{n} v_i y_i = A^{\#}b + (I - A^{\#}A)z$$

für beliebiges $z \in \mathbb{R}^n$.

H 38 $U\Sigma V^H$ sei die Singulärwertzerlegung der Matrix $A \in \mathbb{C}^{m,n}$ mit den singulären Werten $\sigma_1 \geq \sigma_2 \geq \ldots \geq 0$. Daraus bilde man die Matrizen Σ_r und $A_r := U\Sigma_r V^H$, indem man nur die ersten r singulären Werte beibehält, also σ_k durch Null ersetzt für alle $k > r$.

a) Was sind die singulären Werte von $A - A_r$ und berechne $||A - A_r||_2$.
Zeige Rang$A_r \leq r$.

b) Es gibt keine Matrix B vom Rang höchstens r mit $||A - B||_2 < \sigma_{r+1}$.

a) *Die Singulärwertzerlegung von $A - A_r = U(\Sigma - \Sigma_r)V^H$ erhält man, in dem man zu U und V jeweils noch eine Permutation hinzunimmt, so daß die Singulärwerte $\sigma_{r+1} \geq \sigma_{r+2} \geq \ldots \geq 0$ am Anfang stehen. Das sind dann die Singulärwerte von $A - A_r$.*
Man verwende $||A - A_r||_2 = \max(\sigma_k, k > r) = \sigma_{r+1}$
Der Rang von A_r ist die Anzahl der Singulärwerte ungleich 0 ist somit kleiner gleich r. In der Praxis wählt man r so, daß $\sigma_r > \varepsilon > 0$ und damit Rang$A_r = r$.

b) *Es gilt (vgl. H6)*

$$||A - B||_2 \geq \max_i |\sigma_i(A) - \sigma_i(B)|$$

Da B Rang r haben soll, hat B höchstens r Singulärwerte ungleich 0. Damit gilt auf jeden Fall

$$||A - B||_2 \geq |\sigma_{r+1}(A) - \sigma_{r+1}(B)| = \sigma_{r+1}(A)$$

Daraus folgt die Behauptung.