Fachbereich Mathematik
Prof. Dr. P. Spellucci
Sommersemester 2009

TECHNISCHE
UNIVERSITÄT
DARMSTADT

15.6.2009

Numerische Lineare Algebra
Übung 8

Präsenzübung

Ü 21 Die Konditionszahl einer $m \times n$-Matrix vom Rang n ist definiert als

$$\mathrm{cond}(A) \stackrel{def}{=} \frac{\max\{||Ax|| \ : \ ||x|| = 1\}}{\min\{||Ax|| \ : \ ||x|| = 1\}}$$

Welcher Wert ergibt sich für die euklidische Norm unter Anwendung der SVD?

Ü 22 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 $s_{i,i}$ durch ihre Reziprokwerte.

Ü 23 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 $\mathrm{Rang}(A + xy^T) = 2$.
Hinweis: H6

Hausübung

H 22 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:

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

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

H 23 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 24 $U\Sigma V^H$ 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 $||A - B||_2 < \sigma_{r+1}$.

Hinweis: H6

H 25 Gegeben seien $N \gg 3$ Punkte im $\mathbb{R}^3$. Gesucht ist eine Ebene im $\mathbb{R}^3$, sodaß die Summe der Quadrate der orthogonalen Abstände dieser Ebene von den gegebenen Punkten minimal wird. Zeigen Sie, dass man dieses Problem als eine homogene Kleinste-Quadrate Aufgabe mit der Matrix A mit der $i-$ten Zeile

$$[\xi_i, \eta_i, \zeta_i, 1] \ , \ x^i \ = \ (\xi_i, \eta_i, \zeta_i)^T$$

lösen kann, d.h. in der Form

$$||Ac||_2 \stackrel{!}{=} \min_c \ \text{mit} \ ||c||_2 \ = \ 1 \ .$$

Wie kann man nun die Singulärwertzerlegung von A benutzen, um die Lösung "direkt" anzugeben?
Hinweis: Wie drückt sich der Abstand des Punktes x^i von der gesuchten Ebene in der Hessenormalform der Ebene aus?

Numerische Lineare Algebra
Übung 8, Lösungsvorschlag

Präsenzübung

Ü 21 Die Konditionszahl einer $m \times n$-Matrix vom Rang n ist definiert als

$$\text{cond}(A) \stackrel{def}{=} \frac{\max\{||Ax|| \ : \ ||x|| = 1\}}{\min\{||Ax|| \ : \ ||x|| = 1\}}$$

Welcher Wert ergibt sich für die euklidische Norm unter Anwendung der SVD?

Unter Benutzung der Singulärwertzerlegung ergibt sich mit der Substitution $y = V^H x$ *und* $\{x \ : \ ||x|| = 1\} \ = \ \{y = V^H x \ : ||x|| = 1\}$

$$||Ax|| \ = \ ||U\Sigma V^H x|| \ = \ ||\Sigma y||$$

also

$$\max\{||Ax|| \ : \ ||x|| = 1\} \ = \ \sigma_1 \text{ und } \min\{||Ax|| \ : \ ||x|| = 1\} \ = \ \sigma_n$$

wobei wie üblich die Singulärwerte absteigend numeriert angenommen sind.

Ü 22 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 $s_{i,i}$ 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 23 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 $\mathrm{Rang}(A + xy^T) = 2$.
Hinweis: H6

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 22 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

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

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

H 23 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 $I - A^\# A$:

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

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 24 $U\Sigma V^H$ 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 $||A - B||_2 < \sigma_{r+1}$.

Hinweis: H6

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.

H 25 Gegeben seien $N \gg 3$ Punkte im $\mathbb{R}^3$. Gesucht ist eine Ebene im $\mathbb{R}^3$, sodaß die Summe der Quadrate der orthogonalen Abstände dieser Ebene von den gegebenen Punkten minimal wird. Zeigen Sie, dass man dieses Problem als eine homogene Kleinste-Quadrate Aufgabe mit der Matrix A mit der $i-$ten Zeile

$$[\xi_i, \eta_i, \zeta_i, 1] \ , \ x^i \ = \ (\xi_i, \eta_i, \zeta_i)^T$$

lösen kann, d.h. in der Form

$$||Ac||_2 \overset{!}{=} \min_c \text{ mit } ||c||_2 = 1 .$$

Wie kann man nun die Singulärwertzerlegung von A benutzen, um die Lösung "direkt" anzugeben?
Hinweis: Wie drückt sich der Abstand des Punktes x^i von der gesuchten Ebene in der Hessenormalform der Ebene aus?

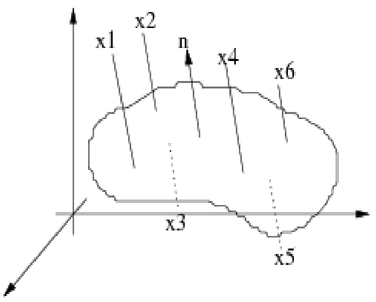
Die Ebenengleichung in Hessenormalform lautet

$$n^T x + \delta = 0; \qquad \textit{mit dem Normalenvektor } n = \begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix}$$

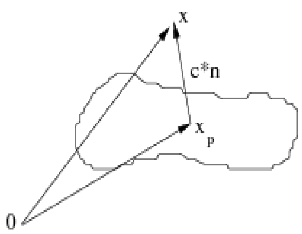
und der Normierung $||n||_2 = 1$. *Es ist dann*

$$\delta = -x_0^T n$$

und x_0 *der dem Nullpunkt am nächsten gelegenen Ebenenpunkt.*



Bestimmung des Abstands eines Punktes x^i *von der Ebene: Hierzu sei* x_p^i *die Projektion von* x^i *auf die Ebene.*



Damit haben wir $x^i = x_p^i + \lambda_x n$ *und der Abstand von* x^i *zur Ebene ergibt sich zu*

$$\| x^i - x_p^i \| = |\lambda|$$

Weiter haben wir wegen der Ebenengleichung für x_p^i:

$$n^T x^i = n^T x_p^i + \lambda n^T n = -\delta + \lambda .$$

Also

$$|\lambda| = \| \, n^T x^i + \delta \, \|$$

Das Problem besteht also darin, die Summe der Abstandsquadrate

$$\sum_{i=1}^{N} (n_1 x_1^i + n_2 x_2^i + n_3 x_3^i + \delta)^2$$

zu minimieren über n und δ unter der Bedingung $||n||_2 = 1$. Mit der angegebenen Form der Matrix und

$$x^i \; = \; (\xi_i, \eta_i, \zeta_i)^T$$

haben wir das Problem

$$||Ac||_2 \; \stackrel{!}{=} \; \min_c$$

mit der Nebenbedingung $c_1^2 + c_2^2 + c_3^2 = 1$ und $c_4 = \delta$. Da aber die Richtung des Lösungsvektors des homogenen Problems ohnehin nur bis auf einen Faktor bestimmt ist, lösen wir zunächst dieses homogene Problem mit der willkürlichen Normierung $||c||_2 = 1$ und skalieren die ersten drei Komponenten nachträglich. Mit der `SVD` *von A*

$$A \; = \; USV^T \;\; O.B.d.A \;\; s_{11} \; \geq \; \ldots \geq s_{44}$$

wird die Norm von Ac offenbar minimal für $c = V(:,4)$. Mit nachträglicher Normierung von $c(1:3)$ auf Länge 1 erhalten wir dann die Hesse-Normalform.