MATEMATISKA VETENSKAPER **Hjälpmedel: inga**
Chalmers tekniska högskola Datum: 2019-06-03 kl. 14.00 - 18.00
Examinator: Thomas Bäckdahl Telefonvakt: Barbara Schnitzer
Lösningsförslag ankn 5325

TMA671 Linjär algebra och numerisk analys

Tentan består av 8 uppgifter, med varierande antal deluppgifter.
Maximalt antal poäng är 60, och betygsgränserna är 30 för 3 (godkänd), 42 för 4 och 54 för 5.
Upp till 10 bonuspoäng från 2019 års bonusuppgifter får tillgodoräknas.
Renskriv dina lösningar, lämna ej in kladdpapper! Poängavdrag ges för dåligt motiverade, svårtolkade eller svårläsliga lösningar.
Tentan rättas och bedöms anonymt. Resultat meddelas via Ladok ca. tre veckor efter tentamenstillfället.

---

1. Betrakta följande tabell över funktionsvärden av en funktion $f(t)$.

t	-1	1	3
f	1	-1	5

Antag också att $f \in C^4([-1,3])$ och $|f^{(4)}(t)| \leq \frac{15}{16}$ för alla $t \in [-1,3]$.

(a) Bestäm interpolationspolynomet genom punkterna. (2p)

(b) Simpsons regel approximerar $\int_a^b f(t)dt$ genom att integrera interpolationspolynomet av grad 2 som interpolerar punkterna a, b och $(a+b)/2$. Härled Simpsons regel utan trunkeringsfel. Du får använda $\int_a^b (t-a)(t-b)dt = -\frac{1}{6}(b-a)^3$ och trapetsregeln utan bevis. (4p)

(c) Bestäm en approximation med felgräns av $\int_{-1}^3 f(t)dt$ med Simpsons regel eller resultatet i a) uppgiften. (2p)
Du får använda att Simpsons regel för approximation av $\int_a^b f(t)dt$ har trunkeringsfel $R_T = \frac{1}{45}2^{-6}(b-a)^5 f^{(4)}(\xi)$ där $a \leq \xi \leq b$.

Lösning:

(a) Vi har 3 punkter så vi behöver ett andragradspolynom. Newtons form är $p(t) = c_0 + c_1(t+1) + c_2(t+1)(t-1)$. Interpolationsvillkoren ger

$$\begin{aligned} p(-1) &= c_0 = 1 \\ p(1) &= c_0 + c_1(1+1) = -1 \\ p(3) &= c_0 + c_1(3+1) + c_2(3+1)(3-1) = 5 \end{aligned}$$

Detta ger $c_0 = 1$, $c_1 = -1$, $c_2 = 1$, så interpolationspolynomet blir $p(t) = 1 - (t+1) + (t+1)(t-1) = t^2 - t - 1$.

(b) Simpsons regel ger samma sak som integration av interpolationspolynomet p_2 av grad 2 som intepolerar punkterna a, b och $(a+b)/2$, vilket ger

$$\begin{aligned} p_2(t) &= c_0 + c_1(t-a) + c_2(t-a)(t-b) \\ p_2(a) &= c_0 = f(a) \\ p_2(b) &= c_0 + c_1(b-a) = f(b) \\ p_2(\tfrac{a+b}{2}) &= c_0 + c_1\tfrac{b-a}{2} - c_2\tfrac{(b-a)^2}{4} = f(\tfrac{a+b}{2}) \end{aligned}$$

Vi får $c_0 = f(a)$, $c_1 = \frac{f(b)-f(a)}{b-a}$ och $f(a) + \frac{f(b)-f(a)}{2} - c_2\frac{(b-a)^2}{4} = f(\frac{a+b}{2})$ som ger $c_2 = \frac{2f(a)+2f(a)-4f(\frac{a+b}{2})}{(b-a)^2}$. Vi ser också att $p_2(t) = p_1(t) + c_2(t-a)(t-b)$, där $p_1(t)$ är interpolationspolynomet i punkterna a och b. Integrering ger

$$\begin{aligned}\int_a^b p_2(t)dt &= \int_a^b p_1(t)dt + c_2\int_a^b (t-a)(t-b)dt = \frac{f(a)+f(b)}{2}(b-a) - c_2\frac{1}{6}(b-a)^3\\ &= \frac{b-a}{6}\left(f(a) + 4f\left(\frac{a+b}{2}\right) + f(b)\right),\end{aligned}$$

där trapetsregeln och den givna integralen använts. Så Simpsons regel lyder

$$\int_a^b f(t)dt \approx \int_a^b p_2(t)dt = \frac{b-a}{6}\left(f(a) + 4f\left(\frac{a+b}{2}\right) + f(b)\right)$$

(c) Simpsons regel ger $\int_{-1}^3 f(t)dt \approx \frac{3+1}{6}(1+4(-1)+5) = 4/3$. Trunkeringsfelet uppskattas till $|R_t| \leq \frac{1}{2880}4^5\frac{15}{16} = 1/3$. Alltså får vi $1 \leq \int_{-1}^3 f(t)d \leq 5/3$.

2. Låt $\mathbf{e} = \{1, t, t^2, t^3\}$ vara en bas för ett vektorrum V med skalärprodukt

$$\langle p(t), q(t)\rangle = \int_{-1}^1 p(t)q(t)dt.$$

Låt $F : V \to V$ vara avbildningen som ges av

$$F\big(p(t)\big) = \frac{d}{dt}\Big((t^2-1)\frac{dp}{dt}\Big).$$

(a) Visa att F är en linjär avbildning. (1p)

(b) Bestäm matrisen A för F i basen $\mathbf{e}$. (2p)

(c) Bestäm avbildningens egenvärden. (2p)

(d) Är F symmetrisk med avseende på skalärprodukten? (Tips: Partialintegrera.) Kan du dra samma slutsats från matrisen A? Varför/varför inte? (3p)

Lösning:

(a) F är linjär ty

$$\begin{aligned}F(\alpha p(t) + \beta q(t)) &= \frac{d}{dt}\Big((t^2-1)\frac{d}{dt}(\alpha p(t) + \beta q(t))\Big) = \alpha\frac{d}{dt}\Big((t^2-1)\frac{dp}{dt}\Big) + \beta\frac{d}{dt}\Big((t^2-1)\frac{dq}{dt}\Big)\\ &= \alpha F(p(t)) + \beta F(q(t))\end{aligned}$$

(b) $F(1) = 0$, $F(t) = 2t$, $F(t^2) = 6t^2 - 2$, $F(t^3) = 12t^3 - 6t$. Matrisen A för F i basen $\mathbf{e}$ blir

$$A = \begin{bmatrix} 0 & 0 & -2 & 0\\ 0 & 2 & 0 & -6\\ 0 & 0 & 6 & 0\\ 0 & 0 & 0 & 12\end{bmatrix}$$

(c) Egenvärdena till F ges av egenvärdena till matrisen A. De kan läsas av från diagonalen och blir således $\lambda_1 = 0$, $\lambda_2 = 2$, $\lambda_3 = 6$, $\lambda_4 = 12$.

(d) Partialintegration ger

$$\begin{aligned}
\langle F\big(p(t)\big), q(t)\rangle &= \int_{-1}^{1} \frac{d}{dt}\Big((t^2-1)\frac{dp}{dt}\Big)q(t)dt \\
&= \underbrace{\Big[(t^2-1)\frac{dp}{dt}q(t)\Big]_{-1}^{1}}_{0} - \int_{-1}^{1}(t^2-1)\frac{dp}{dt}\frac{dq}{dt}dt \\
&= -\underbrace{\Big[(t^2-1)\frac{dq}{dt}p(t)\Big]_{-1}^{1}}_{0} + \int_{-1}^{1} p(t)\frac{d}{dt}\Big((t^2-1)\frac{dq}{dt}\Big)dt \\
&= \langle p(t), F\big(q(t)\big)\rangle
\end{aligned}$$

Så F är symmetrisk. A är inte symmetrisk på grund av att $\mathbf{e}$ inte är en ON-bas. T.ex. är $\langle 1, t^2\rangle = \int_{-1}^{1} t^2 dt = \big[\frac{1}{3}t^3\big]_{-1}^{1} = \frac{2}{3} \neq 0$.

3. Betrakta begynnelsevärdesproblemet (P1) som ges av följande system av differentialekvationer

$$\begin{cases} x_1'(t) = & x_1(t) & + & 3x_2(t) \\ x_2'(t) = & -2x_1(t) & - & 4x_2(t) \end{cases}$$

med begynnelsevillkor $x_1(0) = 1$ och $x_2(0) = 0$.

(a) Lös begynnelsevärdesproblemet (P1) med hjälp av diagonalisering. (4p)

(b) Heuns metod på ett generellt problem $y'(t) = f(t, y(t)),\ y(a) = c$ definieras av

$$y_{k+1} = y_k + \frac{h}{2}(k_1 + k_2), \quad k_1 = f(t_k, y_k), \quad k_2 = f(t_k + h, y_k + hk_1).$$

Utför en iteration med Heuns metod på (P1) med steglängd $h = 1/2$. (2p)

(c) Avgör om Heuns metod är stabil för (P1) med steglängd $h = 1/2$. (3p)
(Tips: Metoden applicerad på testproblemet $y'(t) = \lambda y(t)$ ger en tillväxtfaktor som kan användas för stabilitetsanalysen. Alternativt kan diagonalisering användas.)

Lösning:

(a) Med $\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$ och $A = \begin{bmatrix} 1 & 3 \\ -2 & -4 \end{bmatrix}$ blir systemet $\mathbf{x}'(t) = A\mathbf{x}(t)$.

Diagonalisera A. Karakteristiska polynomet är $\det(A - \lambda I) = \begin{vmatrix} 1-\lambda & 3 \\ -2 & -4-\lambda \end{vmatrix} =$ $\lambda^2 + 3\lambda + 2 = (\lambda+1)(\lambda+2)$, så egenvärdena blir $\lambda_1 = -2$ och $\lambda_2 = -1$.
Egenvektorer till $\lambda_1 = -2$:

$$\left[\begin{array}{cc|c} 3 & 3 & 0 \\ -2 & -2 & 0 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right] \Rightarrow \text{ Egenvektor } \mathbf{v}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Egenvektorer till $\lambda_2 = -1$:

$$\left[\begin{array}{cc|c} 2 & 3 & 0 \\ -2 & -3 & 0 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 3/2 & 0 \\ 0 & 0 & 0 \end{array}\right] \Rightarrow \text{ Egenvektor } \mathbf{v}_2 = \begin{bmatrix} -3 \\ 2 \end{bmatrix}$$

Detta ger $A = TDT^{-1}$ med $T = \begin{bmatrix} -1 & -3 \\ 1 & 2 \end{bmatrix}$ och $D = \begin{bmatrix} -2 & 0 \\ 0 & -1 \end{bmatrix}$.

Låt $\mathbf{y}(t) = \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = T^{-1}\mathbf{x}(t)$. Då får vi $\mathbf{y}'(t) = D\mathbf{y}(t)$, dvs $\begin{cases} y_1'(t) = -2y_1(t) \\ y_2'(t) = -y_2(t) \end{cases}$.

De allmänna lösningarna är $\begin{cases} y_1(t) = c_1 e^{-2t} \\ y_2(t) = c_2 e^{-t} \end{cases}$.

c_1 och c_2 bestäms av $T\mathbf{y}(0) = \mathbf{x}(0) \Leftrightarrow \left[\begin{array}{cc|c} -1 & -3 & 1 \\ 1 & 2 & 0 \end{array}\right] \Rightarrow c_1 = 2, c_2 = -1$.

Det ursprungliga systemets lösning blir

$$\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = T\mathbf{y}(t) = \begin{bmatrix} -1 & -3 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 2e^{-2t} \\ -e^{-t} \end{bmatrix} = \begin{bmatrix} -2e^{-2t} + 3e^{-t} \\ 2e^{-2t} - 2e^{-t} \end{bmatrix}.$$

(b) För (P1) med $\mathbf{x}_k$ som approximation efter k iterationer får vi $k_1 = A\mathbf{x}_k$, $k_2 = A(\mathbf{x}_k + hk_1) = A\mathbf{x}_k + hA^2\mathbf{x}_k$ så $\mathbf{x}_{k+1} = \mathbf{x}_k + \frac{1}{2}h(k_1 + k_2) = (I + hA + \frac{1}{2}h^2A^2)\mathbf{x}_k$. Efter en iteration får vi $\mathbf{x}_1 = (I + \frac{1}{2}A + \frac{1}{8}A^2)\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 7/8 \\ -1/4 \end{bmatrix}$.

(c) Heuns metod applicerat på testproblemet $y'(t) = \lambda y(t)$ ger $y_{k+1} = (1 + h\lambda + \frac{1}{2}h^2\lambda^2)y_k$ så tillväxtfaktorn blir $1 + h\lambda + \frac{1}{2}h^2\lambda^2$. Alternativt kan man använda diagonalisering på (P1) som ger $\mathbf{x}_{k+1} = T(I + hD + \frac{1}{2}h^2D^2)T^{-1}\mathbf{x}_k = T(I + hD + \frac{1}{2}h^2D^2)^kT^{-1}\mathbf{x}_0$. Matrisen i mitten växer inte om $|1 + h\lambda_i + \frac{1}{2}h^2\lambda_i^2| \leq 1$. Med vår steglängd och våra egenvärden får vi $1 + h\lambda_1 + \frac{1}{2}h^2\lambda_1^2 = \frac{1}{2} < 1$ och $1 + h\lambda_2 + \frac{1}{2}h^2\lambda_2^2 = \frac{5}{8} < 1$ så metoden är stabil för den här steglängden.

4. Låt $A = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & -1 \\ 1 & 1 & 0 \\ 0 & 1 & -1 \end{bmatrix}$ och $\mathbf{b} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$.

(a) Bestäm en kompakt QR-faktorisering av A med hjälp av Gram-Schmidts ortogonaliseringsprocess. (4p)

(b) Använd QR-faktoriseringen för att hitta en minstakvadratlösning till $A\mathbf{x} = \mathbf{b}$. (2p)

(c) Använd resultatet i b) och lämplig sats för att hitta en icke-trivial vektor i $N(A^T)$. (2p)

Lösning: Låt $A = [\mathbf{a}_1\ \mathbf{a}_2\ \mathbf{a}_3]$.

(a) Vi använder Gram-Schmidts ortogonaliseringsprocess. Låt

$$\mathbf{u}_1 = \mathbf{a}_1, \quad \mathbf{q}_1 = \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix},$$

$$\mathbf{u}_2 = \mathbf{a}_2 - (\mathbf{q}_1 \cdot \mathbf{a}_2)\mathbf{q}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \sqrt{2}\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{q}_2 = \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|} = \frac{1}{\sqrt{2}}\begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix},$$

$$\mathbf{u}_3 = \mathbf{a}_3 - (\mathbf{q}_1 \cdot \mathbf{a}_3)\mathbf{q}_1 - (\mathbf{q}_2 \cdot \mathbf{a}_3)\mathbf{q}_2 = \begin{bmatrix} 2 \\ -1 \\ 0 \\ -1 \end{bmatrix} - \sqrt{2}\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} - (-\sqrt{2})\frac{1}{\sqrt{2}}\begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix},$$

$$\mathbf{q}_3 = \frac{\mathbf{u}_3}{\|\mathbf{u}_3\|} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}.$$

Då är $\{\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3\}$ en ON bas för $V(A)$.

$$Q_1 = [\mathbf{q}_1\ \mathbf{q}_2\ \mathbf{q}_3] = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$$

Med $r_{jj} = \|u_j\|$ och $r_{ij} = (\mathbf{q}_i \cdot \mathbf{a}_j)$ för $i \leq j$ får vi $A = Q_1 R$ med

$$R = \sqrt{2}\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}.$$

(b) Problemet $A\mathbf{x} = \mathbf{b}$ har minstakvadratlösning som ges av $R\mathbf{x} = Q_1^T\mathbf{b}$, dvs

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 0 & 1 & -1 & 3 \\ 0 & 0 & 1 & -1 \end{array}\right] \quad \Rightarrow \quad \mathbf{x} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$$

(c) Residualen $\mathbf{r} = A\mathbf{x} - \mathbf{b} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ -1 \end{bmatrix}$ uppfyller enligt konstruktion $\mathbf{r} \in V(A)^\perp = N(A^T)$ enligt satsen om de fyra fundamentala underrummen.

5. Betrakta problemet att minimera $f(x_1, x_2) = 2x_2^4 + x_1^2 + x_1x_2 - x_2 + 2$ utan bivillkor.

(a) Gör en iteration med brantaste lutningsmetoden (Steepest Descent), med start i origo och exakt linjesökning. (3p)

(b) Visa att brantaste lutningsmetoden alltid ger en descentriktning (avtaganderiktning) i en punkt om punkten inte är en kritisk punkt. (2p)

Lösning:

(a) Låt $\mathbf{x}^{(0)} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ vara startpunkten. Gradientberäkning ger

$$\nabla f(\mathbf{x}) = \begin{bmatrix} 2x_1 + x_2 \\ 8x_2^3 + x_1 - 1 \end{bmatrix} \quad \Rightarrow \quad \nabla f(\mathbf{x}^{(0)}) = \begin{bmatrix} 0 \\ -1 \end{bmatrix} \quad \Rightarrow \quad \mathbf{s}_0 = -\nabla f(\mathbf{x}^{(0)}) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Exakt linjesökning kan utföras genom att minimera $g(\alpha) = f(\mathbf{x}^{(0)} + \alpha\mathbf{s}_0) = 2\alpha^4 - \alpha + 2$. Stationära punkter: $0 = g'(\alpha) = 8\alpha^3 - 1 \Rightarrow \alpha = 1/2$. De andra rötterna är kompexa och descentriktning endast för $\alpha > 0$. Ny punkt $\mathbf{x}^{(1)} = \mathbf{x}^{(0)} + \alpha\mathbf{s}_0 = \begin{bmatrix} 0 \\ 1/2 \end{bmatrix}$.

(b) $\nabla f(\mathbf{x}_k)^T\mathbf{s}_k = -\nabla f(\mathbf{x}_k)^T\nabla f(\mathbf{x}_k) = -\|\nabla f(\mathbf{x}_k)\|^2 \leq 0$ med likhet endast om $\nabla f(\mathbf{x}_k) = 0$. Alltså är $\mathbf{s}_k = -\nabla f(\mathbf{x}_k)$ alltid en descentriktning om $\nabla f(\mathbf{x}_k) \neq 0$.

6. Betrakta det icke-linjära ekvationssystemet $\begin{cases} 2x_2^3 - x_1x_2 = 1 \\ 3x_1 + x_2^2 = 0 \end{cases}$.

(a) Gör en iteration med Newtons metod med start i $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. (3p)

(b) Gör en metodoberoende feluppskattning av resultatet i a) uppgiften. (3p)

Lösning:

(a) Problemet kan formuleras som $0 = f(\mathbf{x}) = \begin{bmatrix} 2x_2^3 - x_1x_2 - 1 \\ 3x_1 + x_2^2 \end{bmatrix}$. Jacobianen blir $J(\mathbf{x}) = \begin{bmatrix} -x_2 & 6x_2^2 - x_1 \\ 3 & 2x_2 \end{bmatrix}$. I startpunkten $\mathbf{x}^{(0)} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ får vi $J(\mathbf{x}^{(0)}) = \begin{bmatrix} 0 & -1 \\ 3 & 0 \end{bmatrix}$. Newtons metod $J(\mathbf{x}^{(0)})\mathbf{s}^{(0)} = -f(\mathbf{x}^{(0)})$ ger

$$\left[\begin{array}{cc|c} 0 & -1 & 1 \\ 3 & 0 & -3 \end{array}\right] \quad \Rightarrow \quad \mathbf{s}^{(0)} = \begin{bmatrix} -1 \\ -1 \end{bmatrix} \quad \Rightarrow \quad \mathbf{x}^{(1)} = \mathbf{x}^{(0)} + \mathbf{s}^{(0)} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}.$$

(b) Metodoberoende feluppskattning av $\mathbf{x}^{(1)}$ ges av $\|\delta\mathbf{x}\| \approx \|J(\mathbf{x}^{(1)})^{-1}f(\mathbf{x}^{(1)})\|$. Vi får $J(\mathbf{x}^{(1)}) = \begin{bmatrix} 1 & 6 \\ 3 & -2 \end{bmatrix}$ och $f(\mathbf{x}^{(1)}) = \begin{bmatrix} -3 \\ 1 \end{bmatrix}$ så

$$\left[\begin{array}{cc|c} 1 & 6 & -3 \\ 3 & -2 & 1 \end{array}\right] \quad \Rightarrow \quad J(\mathbf{x}^{(1)})^{-1}f(\mathbf{x}^{(1)}) = \begin{bmatrix} 0 \\ -1/2 \end{bmatrix} \quad \Rightarrow \quad \|\delta\mathbf{x}\| \approx 1/2.$$

7. (a) Visa att om A är en kvadratisk matris och $PA = LU$ är en LU-faktorisering med pivotering så är $\det(A) = (-1)^s \det(U)$, där s är antal radbyten som motsvarar multiplikationen med P. (3p)

(b) Använd begreppet Rayleighkvot för att bevisa att för en godtycklig $m \times n$ matris A har matrisen A^TA endast icke-negativa egenvärden. (3p)

Lösning:

(a) $\det(P)\det(A) = \det(PA) = \det(LU) = \det(L)\det(U) = \det(U)$ på grund av att L är triangulär med ettor på diagonalen. Determinanen växlar tecken för varje radbyte, så $\det(P) = (-1)^s$, där s är antal radbyten. Vi får därför $\det(A) = (-1)^s \det(U)$.

(b) Om λ är ett egenvärde till A^TA så finns motsvarande egenvektor $\mathbf{u} \neq 0$ så att $A^TA\mathbf{u} = \lambda\mathbf{u} \Rightarrow \mathbf{u}^TA^TA\mathbf{u} = \lambda\mathbf{u}^T\mathbf{u}$ så egenvärdet är Rayleighkvoten

$$\lambda = \frac{\mathbf{u}^TA^TA\mathbf{u}}{\mathbf{u}^T\mathbf{u}} = \frac{(A\mathbf{u})^TA\mathbf{u}}{\mathbf{u}^T\mathbf{u}} = \frac{\|A\mathbf{u}\|^2}{\|\mathbf{u}\|^2} \geq 0.$$

8. En 10×3 matris A har en kompakt singulärvärdesfaktorisering (SVD) $A = U_1\Sigma_1V_1^T$ med $\Sigma_1 = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 10^{-30} \end{bmatrix}$.

(a) Bestäm egenvärdena för A^TA. (1p)

(b) Utgående från singulärvärdesfaktoriseringen beskriv en matris A^+ så att $\mathbf{x} = A^+\mathbf{b}$ ger minstakvadratlösningen till $A\mathbf{x} = \mathbf{b}$ och visa detta. Att minstakvadratlösningar uppfyller normalekvationerna anses välkänt. (4p)

(c) Varför är varken normalekvationerna eller $\mathbf{x} = A^+\mathbf{b}$ lämplig att använda för numerisk lösning av minstakvadratproblemet? Ange lämpliga mått. (3p)

(d) Beskriv ett bättre sätt att numeriskt approximera minstakvadratlösningen. (2p)

Lösning:

(a) Egenvärdena för A^TA ges av kvadraten på de singulära värdena. Vi får alltså $\lambda_1 = 4$, $\lambda_2 = 1$, $\lambda_3 = 10^{-60}$.

(b) Moore-Penrose pseudoinvers ges av $A^+ = V_1\Sigma_1^{-1}U_1^T$. Kolonnerna i U_1 och V_1 är ortogonala vilket ger $U_1^TU_1 = I$ och $V_1^TV_1 = I$. (Obs vi får <u>inte</u> $U_1U_1^T = I$.) Om $\mathbf{x} = A^+\mathbf{b}$ får vi $A^TA\mathbf{x} = (U_1\Sigma_1V_1^T)^TU_1\Sigma_1\underbrace{V_1^TV_1}_{I}\Sigma_1^{-1}U_1^T\mathbf{b} = V_1\Sigma_1\underbrace{U_1^TU_1}_{I}\underbrace{\Sigma_1\Sigma_1^{-1}}_{I}U_1^T\mathbf{b} = (U_1\Sigma_1V_1^T)^T\mathbf{b} = A^T\mathbf{b}$. Alltså uppfyller $\mathbf{x}$ normalekvationerna så det är en minstakvadratlösning.

(c) Konditionstalet $\kappa(A^+)$ är kvoten mellan största och minsta singulära värdet av A^+. Dessa singulära värden kan läsas ur $\Sigma_1^{-1} = \begin{bmatrix} 2^{-1} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 10^{30} \end{bmatrix}$. Konditionstalet är alltså $\kappa(A^+) = 2 \cdot 10^{30}$. De singulära värdena för $A^TA = (U_1\Sigma_1V_1^T)^TU_1\Sigma_1V_1^T = V_1\Sigma_1^2V_1^T$ kan läsas ur $\Sigma_1^2 = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 10^{60} \end{bmatrix}$. Alltså är konditionstalet $\kappa(A^TA) =$

$4 \cdot 10^{60}$. Så stora konditionstal är väldigt problematiskt i ett flyttalsystem, så båda metoderna är olämpliga.

(d) Trunkerad singulärvärdesfaktorisering approximerar A väl. Motsvarande Moore-Penrose pseudoinvers kan användas dvs $A_2^+ = V_1 \begin{bmatrix} 2^{-1} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} U_1^T$ istället. Konditionstalet blir nu $\kappa(A_2^+) = 2$.