MATEMATISKA VETENSKAPER **Hjälpmedel: inga**

Chalmers tekniska högskola Datum: 2022-08-25 kl. 14.00 - 18.00

Examinator: Thomas Bäckdahl Telefonvakt: Thomas Bäckdahl

Lösningsförslag 031-772 5396

TMA672/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 2021 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 funktionen $f(x) = \dfrac{6}{x+1}$ för $0 \le x \le 5$. Låt $p(x)$ vara interpolationspolynomet som interpolerar f i punkterna $x = 0$, $x = 3$, $x = 5$.

(a) Bestäm interpolationspolynomet $p(x)$. (2p)

(b) Generellt gäller för interpolationsfelet att

$$p_n(x) - f(x) = -\frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_0)(x - x_1)\cdots(x - x_n),$$

där ξ ligger någonstans mellan $\{x_0, x_1, \ldots, x_n, x\}$. Använd detta tillsammans med lämplig skattning av $f^{(n+1)}(\xi)$ för att uppskatta interpolationsfelet i $x = 1$ och jämför med exakt $|p_n(1) - f(1)|$. (2p)

(c) Vad innebär Runges fenomen? (1p)

(d) Uppskatta $\int_0^5 f(x)dx$ med hjälp av funktionsvärdena i $x = 0$, $x = 1$, $x = 3$, $x = 5$. Använd Trapetsregeln på intervallet $[0, 1]$ och Simpsons regel på intervallet $[1, 5]$. (3p)

Lösning:

(a) Vi har 3 punkter så vi behöver ett andragradspolynom. Newtons form är $p(x) = c_0 + c_1x + c_2x(x - 3)$. Interpolationsvillkoren ger

$$\begin{aligned} 6 &= f(0) = p(0) = c_0 \\ \tfrac{3}{2} &= f(3) = p(3) = c_0 + c_1 3 \\ 1 &= f(5) = p(5) = c_0 + c_1 5 + c_2 5(5-3) \end{aligned}$$

Detta ger $c_0 = 6$, $c_1 = -3/2$, $c_2 = 1/4$, så interpolationspolynomet blir $p(x) = 6 - \frac{3}{2}c_1x + \frac{1}{4}x(x - 3) = 6 - \frac{9}{4}x + \frac{1}{4}x^2$.

(b) För vårt fall får vi

$$p_2(x) - f(x) = -\frac{f^{(3)}(\xi)}{3!}x(x - 3)(x - 5).$$

Derivatorna blir $f^{(1)}(x) = -6(x+1)^{-2}, f^{(2)}(x) = 12(x+1)^{-3}, f^{(3)}(x) = -36(x+1)^{-4}$ som är till beloppet avtagande på intervallet. Därför får vi $|f^{(3)}(x)| \le 36$ vilket ger $|p_2(3) - f(3)| = \frac{|f^{(3)}(\xi)|}{3!}|1(1-3)(1-5)| \le \frac{36}{6} \cdot 8 = 48$. Riktiga felet blir $|p_2(1) - f(1)| = |4 - 3| = 1$. Alltså ger feluppskattningsformeln en ganska grov överskattning av felet.

(c) Runges fenomen innebär att interpolationsfelet mellan interpolationspunkterna kan bli stort framförallt nära ändpunkterna. Felet blir större ju högre gradtal man väljer. Problemet kan motverkas genom att hålla gradtalet lågt och fördela interpolationspunkterna tätare nära ändpunkterna.

(d) Trapetsregeln $\int_a^b f(x)dx \approx \frac{b-a}{2}(f(a)+f(b))$ på $[0,1]$ ger

$$\int_0^1 f(x)dx \approx \frac{1}{2}(f(0)+f(1)) = \frac{9}{2}$$

Simpsons regel $\int_a^b f(x)dx \approx \frac{b-a}{6}(f(a)+4f\Big((a+b)/2\Big)+f(b))$ på $[1,5]$ ger

$$\int_1^5 f(x)dx \approx \frac{4}{6}(f(1)+4f(3)+f(5)) = \frac{20}{3}$$

Totalt får vi $\int_0^5 f(x)dx \approx \frac{9}{2}+\frac{20}{3} = \frac{67}{6}$.

2. Låt $V = C[0,1]$ vara vektorrummet av reella kontinuerliga funktioner på intervallet $[-1,1]$. Låt U vara underrummet som spänns upp av $\{1, t, 3t^2-1, t^2\}$.

(a) Visa att

$$\langle p(t), q(t)\rangle = \frac{1}{2}\int_{-1}^1 p(t)q(t)dt.$$

är en skalärprodukt på V. (3p)

(b) Avgör ifall $\{1, t, 3t^2-1, t^2\}$ utgör en ON-bas för U. Om inte, bestäm en ON-bas. (3p)

(c) Bestäm ortogonalprojektionen av $p(t) = t^4$ på underrummet U. (2p)

Lösning:

(a) Den är symmetrisk: $\langle p(t), q(t)\rangle = \frac{1}{2}\int_{-1}^1 p(t)q(t)dt = \langle q(t), p(t)\rangle$, linjär i första argumentet: $\langle ap(t)+br(t), q(t)\rangle = \frac{1}{2}\int_{-1}^1 (ap(t)+br(t))q(t)dt = a\frac{1}{2}\int_{-1}^1 p(t)q(t)dt + b\frac{1}{2}\int_{-1}^1 r(t)q(t)dt = a\langle p(t), q(t)\rangle + b\langle r(t), q(t)\rangle$. Vi har även $\langle p(t), p(t)\rangle = \frac{1}{2}\int_{-1}^1 p(t)^2dt \geq 0$. På grund av kontinuitet får vi även att likhet endast kan gälla om $p(t) = 0$ på hela intervallet. Utan kontinuiteten hade $p(t)$ kunnat vara nollskild på t.ex. enstaka punkter. Vi drar slutsatsen att det är en skalärprodukt.

(b) Vi ser direkt att t^2 är en linjärkombination av $3t^2-1$ och 1. Därför är den givna listan inte en ON-bas, men de tre första skulle kunna utgöra en bas. Beräkning av resterande skalärprodukter ger

$$\langle 1, t\rangle = 0, \qquad \langle 1, 3t^2-1\rangle = 0, \qquad \langle t, 3t^2-1\rangle = 0,$$
$$\langle 1, 1\rangle = 1, \qquad \langle t, t\rangle = \frac{1}{3}, \qquad \langle 3t^2-1, 3t^2-1\rangle = \frac{4}{5}.$$

Ortogonala nollskillda vektorer är linjärt oberoende, så $\{1, t, 3t^2-1\}$ en ortogonal bas för U. Normering ger en ON-bas $\{e_1, e_2, e_3\} = \{1, \sqrt{3}t, \frac{\sqrt{3}}{2}(3t^2-1)\}$.

(c) Projektionen av $p(t) = t^4$ på U är

$$\langle p(t), e_1\rangle e_1 + \langle p(t), e_2\rangle e_2 + \langle p(t), e_3\rangle e_3 = \frac{1}{5}e_1 + 0e_2 + \frac{4}{7\sqrt{5}}e_3 = \frac{1}{5} + \frac{2}{7}(3t^2-1) = \frac{6}{7}t^2 - \frac{3}{35}.$$

3. Betrakta begynnelsevärdesproblemet som ges av följande differentialekvation

$$x''(t) + 5x'(t) + 6x(t) = 0$$

med begynnelsevillkor $x(0) = -1$ och $x'(0) = 4$.

(a) Formulera om problemet som ett första ordnings system. (1p)

(b) Lös begynnelsevärdesproblemet från (a) uppgiften med hjälp av diagonalisering. (4p)

(c) Är problemet stabilt? Motivering krävs. (1p)

(d) Använd diagonaliseringen från (b) uppgiften för att beräkna vad Eulers bakåt-metod ger efter k iterationer med steglängd $h = 1$. (3p)

Lösning:

(a) Låt $y_1 = x$ och $y_2 = x'$, så $y_1' = x' = y_2$ och $y_2' = x'' = -5x' - 6x = -5y_2 - 6y_1$. På matrisform får vi $\mathbf{y}' = A\mathbf{y}$ med

$$A = \begin{bmatrix} 0 & 1 \\ -6 & -5 \end{bmatrix}, \qquad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}, \qquad \mathbf{y}(0) = \begin{bmatrix} -1 \\ 4 \end{bmatrix}$$

(b) Diagonalisera A. Karakteristiska polynomet är $\det(A - \lambda I) = \begin{vmatrix} -\lambda & 1 \\ -6 & -5-\lambda \end{vmatrix} = \lambda^2 + 5\lambda + 6 = (\lambda+3)(\lambda+2)$, så egenvärdena blir $\lambda_1 = -3$ och $\lambda_2 = -2$.
Egenvektorer till $\lambda_1 = -3$:

$$\left[\begin{array}{cc|c} 3 & 1 & 0 \\ -6 & -2 & 0 \end{array}\right] \sim \left[\begin{array}{cc|c} 3 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right] \Rightarrow \text{ Egenvektor } \mathbf{v}_1 = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$$

Egenvektorer till $\lambda_2 = -2$:

$$\left[\begin{array}{cc|c} 2 & 1 & 0 \\ -6 & -3 & 0 \end{array}\right] \sim \left[\begin{array}{cc|c} 2 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right] \Rightarrow \text{ Egenvektor } \mathbf{v}_2 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

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

Låt $\mathbf{z}(t) = \begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix} = T^{-1}\mathbf{y}(t)$. Då får vi $\mathbf{z}'(t) = D\mathbf{z}(t)$, dvs $\begin{cases} z_1'(t) = -3z_1(t) \\ z_2'(t) = -2z_2(t) \end{cases}$.

De allmänna lösningarna är $\begin{cases} z_1(t) = c_1 e^{-3t} \\ z_2(t) = c_2 e^{-2t} \end{cases}$.

c_1 och c_2 bestäms av $T\mathbf{z}(0) = \mathbf{y}(0) \Leftrightarrow \left[\begin{array}{cc|c} -1 & -1 & -1 \\ 3 & 2 & 4 \end{array}\right] \Rightarrow c_1 = 2, c_2 = -1$.

Lösningarna uttryckt i $\mathbf{y}$ blir

$$\begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = T\mathbf{z}(t) = \begin{bmatrix} -1 & -1 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} 2e^{-3t} \\ -e^{-2t} \end{bmatrix} = \begin{bmatrix} -2e^{-3t} + e^{-2t} \\ 6e^{-3t} - 2e^{-2t} \end{bmatrix}.$$

Svar: $x(t) = y_1(t) = -2e^{-3t} + e^{-2t}$.

(c) Båda egenvärdena är negativa, så problemet är stabilt.

(d) Eulers bakåt-metod med steglängd h ger $\mathbf{y}^{(k+1)} = \mathbf{y}^{(k)} + hA\mathbf{y}^{(k+1)} \Rightarrow (I - hA)\mathbf{y}^{(k+1)} = \mathbf{y}^{(k)} \Rightarrow \mathbf{y}^{(k+1)} = (I - hA)^{-1}\mathbf{y}^{(k)}$, där $\mathbf{y}^{(k)}$ är lösningen efter steg k. Efter k iterationer får vi $\mathbf{y}^{(k)} = (I - hA)^{-k}\mathbf{y}^{(0)} = (TT^{-1} - hTDT^{-1})^{-k}\mathbf{y}^{(0)} = T(I - hD)^k T^{-1}\mathbf{y}^{(0)}$. Från (b) har vi att $T^{-1}\mathbf{y}^{(0)} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$. Vi får med $h = 1$

$$\begin{aligned} \mathbf{y}^{(k)} &= \begin{bmatrix} -1 & -1 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} (1+3)^{-k} & 0 \\ 0 & (1+2)^{-k} \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 & -1 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} 2 \cdot 4^{-k} \\ -3^{-k} \end{bmatrix} \\ &= \begin{bmatrix} -2 \cdot 4^{-k} + 3^{-k} \\ 6 \cdot 4^{-k} - 3^{-k} \end{bmatrix}. \end{aligned}$$

4. Låt $A = \begin{bmatrix} 2 & -1 & 1 \\ 0 & 2 & -1 \\ 0 & -2 & 3 \\ 2 & 3 & -3 \end{bmatrix}$ och $\mathbf{b} = \begin{bmatrix} 5 \\ -4 \\ 6 \\ -1 \end{bmatrix}$.

(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) Bestäm minsta avståndet mellan $\mathbf{b}$ och $\mathrm{V}(A)$. (2p)

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

(a) Vi använder Gram-Schmidts ortogonaliseringsprocess. Låt

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

$$\mathbf{u}_2 = \mathbf{a}_2 - (\mathbf{q}_1 \cdot \mathbf{a}_2)\mathbf{q}_1 = \begin{bmatrix}-1\\2\\-2\\3\end{bmatrix} - \sqrt{2}\frac{1}{\sqrt{2}}\begin{bmatrix}1\\0\\0\\1\end{bmatrix} = \begin{bmatrix}-2\\2\\-2\\2\end{bmatrix}, \quad \mathbf{q}_2 = \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|} = \frac{1}{2}\begin{bmatrix}-1\\1\\-1\\1\end{bmatrix},$$

$$\mathbf{u}_3 = \mathbf{a}_3 - (\mathbf{q}_1 \cdot \mathbf{a}_3)\mathbf{q}_1 - (\mathbf{q}_2 \cdot \mathbf{a}_3)\mathbf{q}_2 = \begin{bmatrix}1\\-1\\3\\-3\end{bmatrix} - (-\sqrt{2})\frac{1}{\sqrt{2}}\begin{bmatrix}1\\0\\0\\1\end{bmatrix} - (-4)\frac{1}{2}\begin{bmatrix}-1\\1\\-1\\1\end{bmatrix} = \begin{bmatrix}0\\1\\1\\0\end{bmatrix},$$

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

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

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

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

$$R = \begin{bmatrix}2\sqrt{2} & \sqrt{2} & -\sqrt{2}\\ 0 & 4 & -4\\ 0 & 0 & \sqrt{2}\end{bmatrix}.$$

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

$$\left[\begin{array}{ccc|c} 2\sqrt{2} & \sqrt{2} & -\sqrt{2} & 2\sqrt{2}\\ 0 & 4 & -4 & -8\\ 0 & 0 & \sqrt{2} & \sqrt{2}\end{array}\right] \Rightarrow \left[\begin{array}{ccc|c} 1 & \frac{1}{2} & -\frac{1}{2} & 1\\ 0 & 1 & -1 & -2\\ 0 & 0 & 1 & 1\end{array}\right] \quad \hat{\mathbf{x}} = \begin{bmatrix}2\\-1\\1\end{bmatrix}$$

(c) Enligt konstruktion minimeras $\|\mathbf{r}\|$ av minstakvadratlösningen $\hat{\mathbf{x}}$. Residualen blir $\mathbf{r} = A\hat{\mathbf{x}} - \mathbf{b} = \begin{bmatrix}1\\1\\-1\\-1\end{bmatrix}$. Så $\|\mathbf{r}\| = \sqrt{1+1+1+1} = 2$.

Svar: Minsta avståndet är 2.

5. Låt $f(x_1, x_2) = \frac{1}{4}x_1^4 - x_1 + x_2 - x_1x_2 + x_2^2$. Vi vill minimera $f(\mathbf{x})$.

(a) Bestäm en sökriktning $\mathbf{s}$ med hjälp av Newtons metod med startpunkt $\mathbf{x}^{(0)} = \begin{bmatrix}1\\-1\end{bmatrix}$. (3p)

(b) Utför exakt linjesökning i riktningen $\mathbf{s}$ från (b) uppgiften. (2p)

Lösning:

(a) $\nabla f(\mathbf{x}) = \begin{bmatrix} -1 + x_1^3 - x_2 \\ 1 - x_1 + 2x_2 \end{bmatrix}$, $H(\mathbf{x}) = \begin{bmatrix} 3x_1^2 & -1 \\ -1 & 2 \end{bmatrix}$ Newtons metod ger $H(\mathbf{x}^{(0)})\mathbf{s} = -\nabla f(\mathbf{x}^{(0)}) \Leftrightarrow$ $\left[\begin{array}{cc|c} 3 & -1 & -1 \\ -1 & 2 & 2 \end{array}\right] \Rightarrow \mathbf{s} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

(b) Linjesökningen innebär att vi vill minimera $g(t) = f(\mathbf{x}^{(0)} + t\mathbf{s}) = -\frac{3}{4} + (t-1)^2$. Vi söker stationär punkt $0 = g'(t) = 2(t-1)$. Så $g'(1) = 0$, $g'(t) < 0$ för $t < 1$ och $g'(t) > 0$ för $t > 1$, så $t = 1$ ger minpunkten för linjesökningen. Så $\mathbf{x}^{(1)} = \mathbf{x}^{(0)} + \mathbf{s} = [1, 0]^T$.

6. Betrakta ekvationen $0 = f(x) = 4^x - 2 - x$.

(a) Visa att ekvationen har minst en rot på intervallet $[0, 1]$. (1p)

(b) Använd sekantmetoden för att bestämma x_3 om initialvärdena är $x_0 = 0$ och $x_1 = 1$. (2p)

(c) Som alternativ till sekantmetoden kan man tänka sig att skriva om problemet som $x = g_1(x) = 4^x - 2$ eller $x = g_2(x) = \dfrac{\ln(2+x)}{\ln(4)}$ och använda sig av fixpunktsiteration. Vilken av funktionerna g_1 eller g_2 är lämpligast på intervallet $[0, 1]$? Motivera väl! (Tips: $\ln(2) < \ln(e) < \ln(4)$.) (3p)

Lösning:

(a) Vi har att $f(0) = -1 < 0$ och $f(1) = 1 > 0$. Funktionen är kontinuerlig så satsen om mellanliggande värde ger att minst en rot existerar.

(b) Sekantmetoden ger

$$\begin{aligned} x_{k+1} &= x_k - f(x_k)\frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})}, \\ x_2 &= x_1 - f(x_1)\frac{x_1 - x_0}{f(x_1) - f(x_0)} = \frac{1}{2}, \qquad f(1/2) = -1/2 \\ x_3 &= x_2 - f(x_2)\frac{x_2 - x_1}{f(x_2) - f(x_1)} = \frac{2}{3}. \end{aligned}$$

(c) En fixpunktsiteration $x_{n+1} = g(x_n)$ är lokalt konvergent mot en fixpunkt x^* om $|g(x^*)| < 1$. För våra funktioner får vi $g_1'(x) = \ln(4)4^x$ och $g_2'(x) = \dfrac{1}{(2+x)\ln(4)}$. Funktionen $g_1'(x) \geq \ln(4) > \ln(e) = 1$ på hela intervallet vilket inte ger konvergens. Funktionen $g_2'(x)$ är avtagande på intervallet så $0 < \dfrac{1}{3\ln(4)} = g_2'(1) \leq g_2'(x) \leq g_2'(0) = \dfrac{1}{2\ln(4)} < \dfrac{1}{2}$ för $0 \leq x \leq 1$ vilket ger konvergens. Således måste vi välja g_2 för att få en konvergent metod.

7. (a) Låt A vara en godtycklig $m \times n$ matris. Visa att $\mathrm{N}(A^T) = \mathrm{V}(A)^\perp$. (3p)

(b) Låt A vara en $n \times n$ matris och anta att matrisen har egenvärden $\lambda_1 \neq \lambda_2$ med motsvarande egenvektorer $\mathbf{u}_1$ och $\mathbf{u}_2$. Det kan eventuellt finnas fler egenvärden. Visa att $\mathbf{u}_1$ och $\mathbf{u}_2$ är linjärt oberoende. (3p)

Lösning:

(a) Se sats 2.11 på sida 77.

(b) Antag att de är linjärt beroende. Egenvektorer är inte nollvektorer, så därför finns $\mu \in \mathbb{R}$ så att $\mathbf{u}_1 = \mu\mathbf{u}_2 \Rightarrow A\mathbf{u}_1 = \mu A\mathbf{u}_2 \Rightarrow \lambda_1\mathbf{u}_1 = \mu\lambda_2\mathbf{u}_2$, men $\mu\mathbf{u}_2 = \mathbf{u}_1$, så $\lambda_1\mathbf{u}_1 = \lambda_2\mathbf{u}_1 \Rightarrow \lambda_1 = \lambda_2$ vilket motsäger $\lambda_1 \neq \lambda_2$, så $\mathbf{u}_1$ och $\mathbf{u}_2$ är linjärt oberoende.

8. En 6×4 matris A har en kompakt singulärvärdesfaktorisering (SVD) $A = U_1 \Sigma_1 V_1^T$ med $\Sigma_1 = \begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2^{-40} \end{bmatrix}$.

(a) Bestäm egenvärdena för $A^T A$. (1p)

(b) Använd matriserna U_1, Σ_1 och V_1 för att konstruera 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) Beräkna konditionstalen för lösning med hjälp av normalekvationerna respektive A^+ och förklara varför ingen av metoderna är lämplig för numeriska beräkningar. (3p)

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

Lösning:

(a) Egenvärdena för $A^T A$ ges av kvadraten på de singulära värdena. Vi får alltså $\lambda_1 = 16$, $\lambda_2 = 4$, $\lambda_3 = 1$, $\lambda_4 = 2^{-80}$.

(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 $U_1^T U_1 = I$ och $V_1^T V_1 = I$. (Obs vi får <u>inte</u> $U_1 U_1^T = I$.) Om $\mathbf{x} = A^+\mathbf{b}$ får vi $A^T A\mathbf{x} = (U_1 \Sigma_1 V_1^T)^T U_1 \Sigma_1 \underbrace{V_1^T V_1}_{I} \Sigma_1^{-1} U_1^T \mathbf{b} = V_1 \Sigma_1 \underbrace{U_1^T U_1}_{I} \underbrace{\Sigma_1 \Sigma_1^{-1}}_{I} U_1^T \mathbf{b} = (U_1 \Sigma_1 V_1^T)^T \mathbf{b} = A^T \mathbf{b}$. 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 $\Sigma_1^{-1} = \begin{bmatrix} 4^{-1} & 0 & 0 & 0 \\ 0 & 2^{-1} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2^{40} \end{bmatrix}$. Konditionstalet är alltså $\kappa(A^+) = 4 \cdot 2^{40} = 2^{42}$. De singulära värdena för $A^T A = (U_1 \Sigma_1 V_1^T)^T U_1 \Sigma_1 V_1^T = V_1 \Sigma_1^2 V_1^T$ kan läsas ur $\Sigma_1^2 = \begin{bmatrix} 16 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2^{-80} \end{bmatrix}$. Alltså är konditionstalet $\kappa(A^T A) = 16 \cdot 2^{80} = 2^{82}$. 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 $A_2^+ = V_1 \begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} U_1^T$ istället. Konditionstalet blir nu $\kappa(A_2^+) = 4$.