UMEÅ UNIVERSITET
Institutionen för datavetenskap

TENTAMEN

Tentamen

Teknisk-vetenskapliga beräkningar

16 mars 2007

Examinator: Pedher Johansson
Skrivtid: 16 – 22
Hjälpmedel: Inga böcker
Ingen miniräknare

- Maxpoäng 45 p

  Betyg som ges är U, 3, 4 och 5, med följande betygsgränser:

	U	< 23 p
23 p $\leq$	3	< 29 p
29 p $\leq$	4	< 36 p
36 p $\leq$	5	≤ 45 p

- Uppgifterna är slumpvis ordnade
- Börja varje uppgift på ett nytt blad och skriv uppgiftsnummer längst uppe till höger
- Skriv namn på *varje* blad
- Läs frågan noggrannt och besvara de sökta frågeställningarna.
- Uttryck dig *mycket tydligt.* Alla *otydligheter* tolkas till din *nackdel.*
- Även om du inte klarar hela uppgiften kan du få poäng, lös därför så mycket du kan.
- Är frågan uppdelad i flera deluppgifter kan man klara senare deluppgifter även om man inte klarat första. Försök därför att klara *alla* deluppgifter.
- Resultat kommer att tillkännages via webben.
- Lärare kommer förbi kl. 18.00 och kl. 20.00
- Fråga om något är oklart!

Lycka till!

Uppgift 1 (4 p) – *Talrepresentation*

Du står inför problemet att addera tre tal med hjälp av en ”primitiv” dator. Datorn i fråga använder avrundad aritmetik och de reella talen lagras på formen $\pm .d_1 d_2 \ldots d_3 \cdot \beta^p$ där de olika d_i:na utgör den s.k. taldelen och p motsvarar exponenten. I fallet med den ”primitiva” datorn gäller att $\beta = 10$, $t = 3$ samt $0.1 \leq |\pm .d_1 d_2 d_3| < 1$. Talen som ska adderas är $a = 1$, $b = 10^{-3}$ samt $c = -1$. *På vilket av de två nedan föreslagna sätten* skulle du föredra att utföra additionen? *Motivera* utförligt ditt val.

Alt. 1: $(a + b) + c$
Alt. 2: $a + (b + c)$

Det första alternativet innebär att tal a och b först adderas och därefter adderas detta delresultat med c. Det andra alternativet innebär att talen b och c först adderas och att detta delresultat därefter adderas med a.

Uppgift 2 (4 + 4 p) – *Matrisfaktorisering*

Antag att vi har en $m \times n$ matris A ($A \in \mathbb{R}^{m \times n}$), betänk i frågeställningarna nedan, följande tre faktoriseringar,

- LU-faktorisering med pivotering,
- QR-faktorisering, samt
- Cholesky-faktorisering.

a)

Givet de tre faktoriseringarna ovan, *vilka egenskaper* måste vara uppfyllda hos matrisen A, för att *respektive* faktorisering ska kunna utföras.

b)

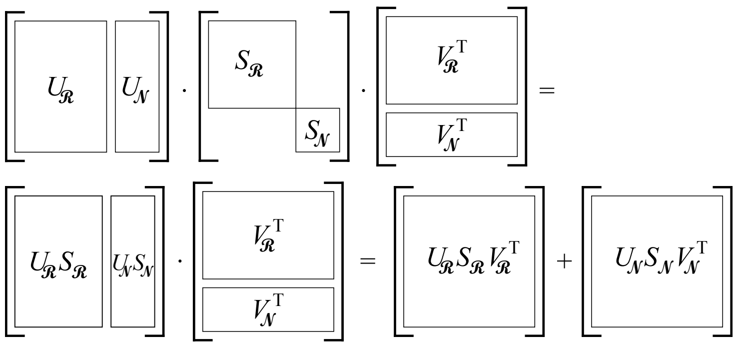
Antag att samtliga nödvändiga egenskaper hos A är uppfyllda för respektive faktorisering ovan. *För var och en* av faktoriseringarna, *beskriv* form, storlek samt övriga specifika egenskaper hos de resulterande matriserna vid en faktorisering av A.

Beskriv hur man med hjälp av *respektive* faktorisering kan lösa ett ekvationssystem.

Uppgift 3 (3 + 3 + 3 p) – *SVD, kondition och stabilitet*

Om vi har en matris som är 3×3 och gör en singulärvärdesfaktorisering (SVD) av denna, $A = U \cdot S \cdot V^T$, så är S en diagonalmatris med de singulära värdena (σ_i) på diagonalen. Vi kan notera att $A = u_1\sigma_1 v_1^T + u_2\sigma_2 v_2^T + u_3\sigma_3 v_3^T$ om u_i och v_i är kolumnerna i U och V.

Om $\sigma_1, \sigma_2 > 0$ och $\sigma_3 = 0$ innebär det att u_1 och u_2 spänner upp värderummet till A (v_3 spänner upp nollrummet). Vi kan då också göra följande omskrivning av A där $U_{\mathcal{R}} = [u_1 u_2]$, $V_{\mathcal{R}} = [v_1 v_2]$, $U_{\mathcal{N}} = u_3$ och $V_{\mathcal{N}} = v_3$.



Detta ger oss att $A = U_{\mathcal{R}} S_{\mathcal{R}} V_{\mathcal{R}}^T + U_{\mathcal{N}} S_{\mathcal{N}} V_{\mathcal{N}}^T$

Istället för att lösa $Ax = b$ som enbart har en lösning om b ligger i värderummet till A, kan vi försöka hitta $\tilde{x}$ så att $A\tilde{x}$ är den punkt som ligger närmast b på värderummet till A. Detta får vi genom att använda den s.k. *pseudo-inversen*, $A^+ = V_{\mathcal{R}} S_{\mathcal{R}}^{-1} U_{\mathcal{R}}^T$. Lösningen $\tilde{x}$ kan då beräknas som $\tilde{x} = A^+ b$.

a)

Verifiera att matrisoperationerna för att beräkna A^+ är giltiga med avseende på matrisdimensionerna. Visa också hur $S_{\mathcal{R}}^{-1}$ ser ut.

b)

Antag nu att σ_3 inte är 0 utan $1 \cdot 10^{-14}$ men vi förändrar inte storleken på $U_{\mathcal{R}}$, $S_{\mathcal{R}}$ och $V_{\mathcal{R}}$.

Hur långt ligger $U_{\mathcal{R}} S_{\mathcal{R}} V_{\mathcal{R}}^T \tilde{x}$ från $A\tilde{x}$ med avseende på 2-normen?

Resonera kring vad detta säger om stabiliteten i att använda pseudo-inversen till att beräkna $\tilde{x}$ när vi har små singulära värden?

c)

Givet ett ekvationssystem $Ax = b$ ger en förändring i b (δb) ger en förändring i x (δx). Alltså gäller att

$$A(x + \delta x) = b + \delta b,$$

d.v.s. att

$$\delta x = A^{-1}\delta b$$

Studera nu ekvationen $u_3\sigma_3 v_3^T \delta x = \delta b$ och *resonera* kring hur ett fel i b fortplantas till x om vi tar med $\sigma_3 = 1 \cdot 10^{-14}$ vid lösningen av ekvationssystemet $Ax = b$.

Uppgift 4 (4 + 4 p) – *Interpolation och approximation*

a)

En planet har en elliptisk bana som kan beskrivas i ett Cartesiskt koordinatsystem (x, y) enligt modellen

$$ax^2 + bxy + cx + dy + e = x^2$$

där a, b, c, d, och e är obekanta modellparametrar som man önskar bestämma för att få en komplett modell. Följande data för planetens position är given:

x	1.02	0.95	0.87	0.77	0.67	0.56	0.44	0.30	0.16	0.01
y	0.39	0.32	0.27	0.22	0.18	0.15	0.13	0.12	0.13	0.15

Ställ upp det linjära ekvationssystem som ska lösas för att bestämma de okända modellparametrarna från de givna mätvärdena. Du ska alltså ange det $\mathbf{A}$, $\mathbf{x}$, och $\mathbf{b}$ du då får, och avgöra huruvida det i denna uppgift är frågan om *interpolation eller approximation. Motivera* utförligt ditt val. Du ska inte explicit räkna ut vektorn $\mathbf{x}$ samt högerledsvektorn $\mathbf{b}$. Observera vidare skillnaden mellan det skalära x-värdet som anger en koordinat och lösningsvektorn $\mathbf{x}$.

b)

Beskriv vad som menas med Runges fenomen och hur detta kan undvikas.

Uppgift 5 (4 + 4 p) – *Richardsonextrapolation*

a)

Antag att vi har följande värden givna

x	1	2	4
y	1	3	-3

Vi vill nu interpolera dessa punkter med ett andragradspolynom $f(x) = c_2x^2 + c_1x + c_0$. *Härled och visa* hur vi med Richardsonextrapolation kan bestämma värdena på de obekanta c_0, c_1 och c_2.

b)

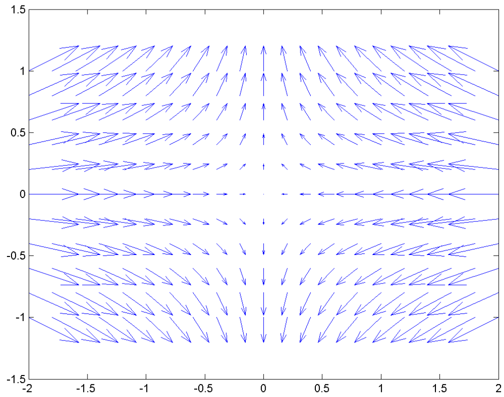
När vi använder trapetsregeln till att beräkna integralen av en funktion inom ett intervall ges denna av

$$T(h) = h(\tfrac{f_0}{2} + f_1 + f_2 + \ldots + f_{n-1} + \tfrac{f_n}{2})$$
$$\int_a^b f(x) = T(h) + c_1h^2 + c_2h^4 + c_3h^6 + c_4h^8 + \ldots$$

där h är steglängden och c_i är konstanter. Med hjälp av Richardsonextrapolation kan vi eliminera en eller flera av feltermerna c_ih^{2i} och därmed få en noggrannare uppskattning av integralen utan att behöva minska steglängden. Resonera kring begreppen trunkeringsfel, tabellfel och maskinnoggrannhet och därmed hur många av feltermerna som det är meningsfullt att eliminera.

Uppgift 6 (4 + 4 p) – *ODE*

a)



Beskriv principerna bakom framåteuler (Explicit Euler) för att lösa ett bygynnelsvärdesproblem. Utgå i din beskrivning från ett riktningsfält.

b)

Givet är följande system av ODE

$$\begin{aligned} \ddot{x} &= (3 - \sin(t))\dot{x} + \tfrac{x}{1+y^2} \\ \dot{y} &= -\cos(t)y - \tfrac{\dot{x}}{1+t^2} \end{aligned}$$

där x och y är kontinuerliga funktioner som beror på tiden t, d.v.s. $x(t)$ och $y(t)$, samt att $\ddot{x} = \frac{d^2x}{dt^2}$ och $\dot{x} = \frac{dx}{dt}$. Vid starten gäller vidare följande:

$$\begin{cases} x(0) = 0 \\ \dot{x}(0) = -1 \\ y(0) = 4 \end{cases}$$

Du ska *inte lösa* differentialekvationssystemet utan bara *ställa upp det på standardformen* $\frac{d\mathrm{y}}{dt} = \mathrm{f}(\mathrm{y}, t)$, $\mathbf{y}(t_0) = \mathrm{y}_0$ där det klart och tydligt ska framgå vilka element vektorerna $\frac{d\mathrm{y}}{dt}$, f, y, och y_0 har.