Løsningsforslag

eksamen i IMAA2150,
IMAG2150 og IMAT2150
Høst 2023

NTNU
Institutt for matematiske fag

-

$\boxed{1}$

a) La $A = \begin{bmatrix} -1 & 0 & 1 \\ 1 & 2 & 3 \\ 1 & 0 & 1 \end{bmatrix}$. Finn en øvre triangulær matrise U og en nedre triangulær matrise L på formen $L = \begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{bmatrix}$, slik at $A = LU$.

Løsning: LU-faktorisering gjøres ved å gausseliminere A for å få U, og lagre radoperasjonene vi gjør i matrisen L, slik at $LU = A$:

$$\begin{bmatrix} -1 & 0 & 1 \\ 1 & 2 & 3 \\ 1 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} -1 & 0 & 1 \\ -1 & 2 & 4 \\ -1 & 0 & 2 \end{bmatrix}$$

L, og U blir da:

$$L = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}$$

$$U = \begin{bmatrix} -1 & 0 & 1 \\ 0 & 2 & 4 \\ 0 & 0 & 2 \end{bmatrix}$$

b) Løs systemet $\begin{matrix} -x & + & 0y & + & z & = & 1 \\ x & + & 2y & + & 3z & = & 1 \\ x & + & 0y & + & z & = & 1 \end{matrix}$ ved hjelp av LU-faktoriseringen til matrisen i punkt a).

Løsning: Systemet kan beskrives som $Ax = b$, hvor A er som i a) og:

$$b = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Fordi vi har LU-faktoriseringen av A, kan vi løse systemet ved først å løse $Lc = b$ og deretter $Ux = c$.

$$Lc = b$$

$$\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$c = \begin{bmatrix} 1 \\ 1 + 1 \cdot 1 = 2 \\ 1 + 1 \cdot 1 = 2 \end{bmatrix}$$

Deretter:

$$Ux = c$$

$$\begin{bmatrix} -1 & 0 & 1 \\ 0 & 2 & 4 \\ 0 & 0 & 2 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$$

Vi løser da likningssystemet:

$$\begin{aligned} 2 \cdot x_3 = 2 &\implies x_3 = 1 \\ 2x_2 + 4x_3 = 2 &\implies x_2 = \frac{2-4}{2} = -1 \\ -1x_1 + 0x_2 + x_3 = 1 &\implies x_1 = \frac{1-1}{-1} = 0 \end{aligned}$$

Løsning:

$$x = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$$

2 En funksjon er definert ved

$$\begin{aligned} S_1(x) &= x^3 + 3x, \quad 0 \le x \le 1 \\ S_2(x) &= -x^3 + 6x^2 - 3x + 2, \quad 1 \le x \le 2 \end{aligned}$$

a) Vis at denne funksjonen er en normal kubisk spline.

Løsning: For at funksjonen skal være en kubisk spline skal funksjonsverdiene samt de første to deriverte være identisk ved overgangen mellom S_1 og S_2. Videre, for at funksjonen skal være en *normal* kubisk spline, skal den andrederiverte i endepunktene være lik 0.
Den første deriverte av funksjonen er:

$$\begin{aligned} \frac{dS_1(x)}{dx} &= 3x^2 + 3, \quad 0 \le x \le 1 \\ \frac{dS_2(x)}{dx} &= -3x^2 + 12x - 3, \quad 1 \le x \le 2 \end{aligned}$$

Den andre deriverte av funksjonen er:

$$\begin{aligned}\frac{d^2S_1(x)}{dx^2} &= 6x, \quad 0 \le x \le 1\\ \frac{d^2S_2(x)}{dx^2} &= -6x+12, \quad 1 \le x \le 2\end{aligned}$$

Dermed finner vi:

$$\begin{aligned}S_1(1) &= 4 = & S_2(1)\\ S_1'(1) &= 6 = & S_2'(1)\\ S_1''(1) &= 6 = & S_2''(1)\end{aligned}$$

Funsjonsverdien samt de første to deriverte er dermed like ved overgangene mellom S_1 og S_2, ved $x = 1$. Funksjonen er altså en kubisk spline. Vi regner til slutt ut $S_1''(0) = 0$ og $S_2''(2) = 0$ som betyr at funsjonen er en normal kubisk spline.

b) Tilpass funksjonsdefinisjonen i intervallet $[1, 2]$ slik at funksjonen fortsatt er en kubisk spline, men med et nytt endepunkt $(x, y) = (2, 11)$.

Løsning: For at funksjonen fortsatt skal være en kubisk spline skal S_2 være et tredjegrads polynom slik at fortsatt

$$\begin{aligned}S_2(1) &= 4\\ S_2'(1) &= 6\\ S_2''(1) &= 6\end{aligned}$$

Dette kan vi for eksempel få til slik:

$$\begin{aligned}S_2(x) &= 4+6(x-1)+\frac{1}{2}\cdot 6(x-1)^2+K(x-1)^3 =, \quad 1 \le x \le 2\\ &= 4+6(x-1)+3(x-1)^2+K(x-1)^3, \quad 1 \le x \le 2\end{aligned}$$

Til slutt skal vi velge K slik at

$$S_2(2) = 4+6(2-1)+3(2-1)^2+K(2-1)^3 = 11$$

Vi finner $K = -2$ slik at

$$S_2(x) = 4+6(x-1)+3(x-1)^2-2(x-1)^3$$

Om vi ønsker det kan vi skrive dette om til f.eks.

$$S_2(x) = -2x^3+9x^2-6x+3$$

3 Gram-Schmidt-ortogonaliseringsalgoritmen for å danne tre ortogonale vektorer v_1, v_2 og v_3 ut av vilkårlige lineært uavhengige vektorer x_1, x_2 og x_3 er gitt ved

$$\begin{aligned} v_1 &= x_1 \\ v_2 &= x_2 - \frac{x_2 \cdot v_1}{v_1 \cdot v_1} v_1 \\ v_3 &= x_3 - \frac{x_3 \cdot v_1}{v_1 \cdot v_1} v_1 - \frac{x_3 \cdot v_2}{v_2 \cdot v_2} v_2 \end{aligned}$$

a) Vis at vektoren v_1 er ortogonal på vektoren v_2.

Løsning: Gitt $v_1 = x_1$
og $v_2 = x_2 - \frac{x_2 \cdot v_1}{v_1 \cdot v_1} v_1$
Vi beviser ortogonaliteten ved å vise at prikkproduktet er lik null:
$v_1 \cdot v_2 = v_1 \cdot (x_2 - \frac{x_2 \cdot v_1}{v_1 \cdot v_1} v_1) =$
$v_1 \cdot x_2 - v_1 \cdot \frac{x_2 \cdot v_1}{v_1 \cdot v_1} v_1 =$
$v_1 \cdot x_2 - \frac{x_2 \cdot v_1}{v_1 \cdot v_1} v_1 \cdot v_1 =$
$v_1 \cdot x_2 - x_2 \cdot v_1 = 0$

b) Vis at vektoren v_1 er ortogonal på vektoren v_3.

Løsning:
Løsningsmetode 1 for oppgave a og b: La generelt $\vec{P}_j^i$ være projeksjonen av x_i ned på v_j, og la $\vec{N}_j^i$ være normalen ned på v_j fra x_i. (Tegn en figur!)

Altså kan vi skrive $v_2 = x_2 - \vec{P}_1^2 = \vec{N}_1^2$, som gir at $v_1 \cdot v_2 = 0$, dvs. $v_1 \perp v_2$.

Tilsvarende kan vi skrive $v_3 = \vec{N}_1^3 - \vec{P}_2^3$, som gir at $v_1 \cdot v_3 = 0$ da $\vec{N}_1^3$ er normal på v_1 og $\vec{P}_2^3$ er parallell med v_2 (og $v_1 \cdot v_2 = 0$). Så, $v_1 \perp v_3$.

Og, selv om vi ikke spurte om det:
At også $v_2 \cdot v_3 = 0$ følger da vi kan skrive $v_3 = \vec{N}_2^3 - \vec{P}_1^3$. M.a.o. er også $v_2 \perp v_3$.
Løsningsmetode 2: Gitt at $v_1 = x_1$
og
$v_3 = x_3 - x_3 \cdot v_1/(v_1 \cdot v_1)\, v_1 - x_3 \cdot v_2/(v_2 \cdot v_2)\, v_2$
Vi beviser ortogonaliteten ved å vise at prikkproduktet er lik null:
$v_1 \cdot v_3 = v_1 \cdot (x_3 - x_3 \cdot v_1/(v_1 \cdot v_1)\, v_1 - x_3 \cdot v_2/(v_2 \cdot v_2)\, v_2) =$
$v_1 \cdot x_3 - x_3 \cdot v_1/(v_1 \cdot v_1)\, v_1 \cdot v_1 - x_3 \cdot v_2/(v_2 \cdot v_2)\, v_1 \cdot v_2 =$
$v_1 \cdot x_3 - x_3 \cdot v_1 - x_3 \cdot v_2/(v_2 \cdot v_2)\, v_1 \cdot v_2 =$
$-x_3 \cdot v_2/(v_2 \cdot v_2)\, v_1 \cdot v_2$
men vi har allerede sett at $v_1 \cdot v_2 = 0$ slik at $v_1 \cdot v_3 = 0$.

c) La $V = \{[1,\ 1,\ 1]\}$, utvid V til en ortogonal basis for $\mathbb{R}^3$.

Løsning: La $V = \{x_1, x_2, x_3\}$ med $x_1 = (1,1,1)$ gitt. Vi velger
$x_2 = (1,0,0)$ og
$x_3 = (0,1,0)$.
Det er lett å se at dette gir at disse tre vektorene er lineært uavhengige i $\mathbb{R}^3$ enten ved bruk av 3×3 determinant-testen eller fra selve definisjonen på begrepet lineær uavhengighet. Merk at vektorene er valgt slik at $x_1 \cdot x_2 = 1$ og $x_1 \cdot x_3 = 1$. Vi får:
$v_1 = x_1 = (1,1,1)$,
$v_2 = x_2 - x_2 \cdot v_1/(v_1 \cdot v_1)\, v_1 = x_2 - (x_2 \cdot x_1)/(v_1 \cdot v_1)\, v_1 = (2/3, -1/3, -1/3)$ og
$v_3 = x_3 - (x_3 \cdot v_1)/(v_1 \cdot v_1)\, v_1 - (x_3 \cdot v_2)/(v_2 \cdot v_2)\, v_2 = x_3 = (0,1,0) - (1/3,1/3,1/3) + 1/3/(2/3) \cdot (2/3,-1/3,-1/3) = (0,1/2,-1/2)$. Vi har altså $v_1 = (1,1,1)$, $v_2 = (2/3,-1/3,-1/3)$ og $v_3 = (0,1/2,-1/2)$. Vi ser at $v_1 \cdot v_2 = 0$, $v_1 \cdot v_3 = 0$ og $v_2 \cdot v_3 = 0$, som det skal.

4 Gitt kurvene $y = x^3$ og $x^2 + y^2 = 1$ i planet.

a) Skisser grafene til de to kurvene på en figur og argumentér geometrisk for at disse kurvene har to skjæringspunkter: et skjæringspunkt i første kvadrant og et i tredje kvadrant.

Løsning: Dette er eksempel 2.32, s. 132 fra SAUER.

b) Hvis et av skjæringspunktene er (u, v), hva er det andre skjæringspunktet?

Løsning: Fra symmetri ser vi at det andre skjæringspunktet er $(-u, -v)$. Funksjonen $y = x^3$ er odde. Derfor gjelder $y(x) = -y(-x)$. Om punkt (u, v) skjærer sirkelen om origo må da også punkt $(-u, -v)$ skjære samme sirkel.

c) Bruk Newtons iterasjonsmetode én gang med startverdi $(1, 2)$ til å bestemme en tilnærmelse til ett av skjæringspunktene.

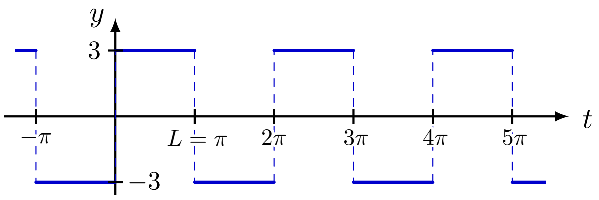
Løsning: Skjæringspunktene er (u, v) og $(-u, -v)$ for $u = 0,826031\cdots$ og $v = 0,563624\cdots$ etter at Newton iterasjonene har stabilisert seg.

5 **a)** Skisser den periodiske funksjonen

$$f(t) = \begin{cases} 3 & \text{for } 0 < t < \pi, \\ -3 & \text{for } \pi < t < 2\pi, \\ 0 & \text{for } t = 0 \text{ og } t = \pi \end{cases}$$

der $f(t + 2\pi) = f(t)$.

Løsning:



b) Er funksjonen i a-oppgaven jamn, odde, eller ingen av delene? Begrunn svaret.

Løsning: Funksjonen er odde siden $f(-t) = -f(t)$, som vi kan se av grafen.

c) Beregn de to første leddene ulik null i fourierrekka for funksjonen i a).

Løsning: $a_i = 0$ for alle verdier av i for en odde funksjon.
Videre er $b_n = \frac{2}{L}\int_0^L f(x)\sin(\frac{n\pi x}{L})dx$
der $L = \pi$ siden L er halvparten av perioden til funksjonen. Derfor:
$b_n = \frac{2}{\pi}\int_0^\pi 3\sin(\frac{n\pi x}{\pi})dx$
siden $f(t) = 3$ i intervallet $[0, \pi]$.
Vi finner $b_n = 3\frac{2}{\pi}\int_0^\pi \sin(nx)dx =$
$\frac{6}{\pi}[-\frac{1}{n}\cos(nx)]_0^\pi =$
$\frac{-6}{n\pi}[\cos(n\pi) - \cos(0)] =$

$$\begin{cases} \frac{-6}{n\pi}(-1-1) = \frac{12}{n\pi} & \text{for } n \text{ oddetall,} \\ \frac{-6}{n\pi}(1-1) = 0 & \text{for } n \text{ partall,} \end{cases}$$

De første to koeffisientene ulik null blir derfor $b_1 = \frac{12}{\pi}$ og $b_3 = \frac{4}{\pi}$.
Sinusrekken er

$\sum_{n=1}^{\infty}(b_n \sin(\pi nt/L)) =$
$\sum_{n \text{ pos. oddetall}}(\frac{12}{n\pi}\sin(nt))$
Summen av de første to leddene ulik null blir: $\frac{12}{\pi}\sin(t) + \frac{4}{\pi}\sin(3t)$

6 Vi skal i denne oppgaven bruke IEEE 754-formatet for DP (*double precision*) for å lagre flyttall i datamaskina. Dette formatet anvender totalt 64 biter der den første biten brukes for å lagre fortegnet fulgt av de neste 11 bitene for å lagre eksponenten (såkalt *biased exponent* BE). BE-en beregnes ved å legge til 1023 til flyttallseksponenten. Det hele avsluttes med å lagre mantissa ved bruk av de siste 52 bitene. For IEEE 754-representasjonen av et reelt tall n bruker vi her $fl(n)$.

a) Bestem den binære representasjonen for $fl(n)$ for $n = 19,59375$. Du kan bruke at $n = 19+19/32$ for å forenkle beregningene. Glem ikke avrunding hvis nødvendig.

Løsning: Vi har at $n = 19 + 19/32$, og da $19 = 1 + 2 + 16 = 2^4 + 2^1 + 1$ er $19/32 = 19/2^5 = 1/2 + 1/2^4 + 1/2^5$, dvs. $n = 10011.10011$. (Med minimal regning!)

Dette gir at eksponenten er $e = 4$ med biased exponent"$BE = 2^{10} - 1 + 4 = 2^{10} + 2^1 + 1$.

Dette gir at IEEE 754 representasjonen for n er

$\boxed{|0|100|0000|0011|0011|1001|1000|0000|0000|0000|0000|0000|000|0000|\cdots|0000|}$
(avsluttes med 43 0-er).

Binærrepresentasjonen er endelig. (Forkortet brøk for n har nevner som er lik 2^5, en potens av 2.) Mantissa utgjør 9 biter, så representasjonen er eksakt.[5]
For $x = 9,4$ er $fl(x) = 4022CCCCCCCCCCCD$ i 16-tallssystemet. Spesialsymbolene betyr: $A = 1010$, $B = 1011$, $C = 1100$, $D = 1101$, $E = 1110$ og $F = 1111$.

b) Bruk dette til å angi $fl(2x)$ og $fl(-x/2)$ i samme format. Glem ikke avrunding hvis nødvendig. Begrunn fremgangsmåten.

Løsning: For $x = 9,4$ er $fl(x) = 4022CCCCCCCCCCCD$ oppgitt.
Binær representasjonen til x er uendelig periodisk så representasjonen er ikke eksakt. (Det har vært avrunding oppover.)

Vi skal bestemme representasjonen for $fl(2x)$ og $fl(-x/2)$ ut fra $fl(x)$.

Det gjelder å innse at det er kun eksponentene som skiller disse flyttallene, samt fortegnet til det andre tallet.

Vi har for $fl(x)$ at 402 betyr $\boxed{0|100|0000|0010}$, som utgjør fortegnet og BE for $fl(x)$.

Nå er $BE_{2x} = BE_x + 1$ og $BE_{x/2} = BE_x - 1$. Til slutt, for å endre fortegnet til det andre tallet endrer vi den aller første biten, som her korresponderer til å legge til 8 til det første sifferet slik at det blir $4 + 8 = C$

Dette gir at $fl(2x) = 4032CCCCCCCCCCCD$ og $fl(-x/2) = C012CCCCCCCCCCCD$.

7 Anta at vi har gjort følgende målinger av konsentrasjonen y av en medisin i blodet til en pasient t timer etter inntak.

Tilpass en modell $y = c_1 e^{c_2 \cdot t}$, til dataene ved hjelp av minste kvadraters metode, og gi RMSE til tilpasningen.

[5] HEX representasjonen for n er 4033980000000000. (Dette var ikke påkrevd.)

t	y
0	1000
6	550
12	270
18	150

Løsning: For å tilpasse modellen $y = c_1 e^{c_2 t}$ må vi få det over på form $b = k_1 + k_2 t$ ved å ta den naturlige logaritmen av y, for å bruke minste kvadraters metode.

$$\begin{aligned} &y = c_1 e^{c_2 t} \\ &ln(y) = ln(c_1 e^{c_2 t}) \\ &ln(y) = ln(c_1) + c_2 t \end{aligned}$$

Når systemet er på denne formen kan man bruke normalsystemet fra formelsamlingen $A^T Ax = A^T b$, der A er en matrise satt opp av koeffisientene som er multiplisert med de ukjente vi er ute etter, c_1 og c_2. Da blir A-matrisen og b-vektoren som følger:

$$A = \begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ 1 & t_3 \\ 1 & t_4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 6 \\ 1 & 12 \\ 1 & 18 \end{bmatrix}$$

$$b = \begin{bmatrix} ln(y_1) \\ ln(y_2) \\ ln(y_3) \\ ln(y_4) \end{bmatrix} = \begin{bmatrix} ln(1000) \\ ln(550) \\ ln(270) \\ ln(150) \end{bmatrix} = \begin{bmatrix} 6,9078 \\ 6,3099 \\ 5,5984 \\ 5,0106 \end{bmatrix}$$

Og normalsystemet vil være et likningssystem med:

$$A^T A = \begin{bmatrix} 4 & 36 \\ 36 & 504 \end{bmatrix}$$

$$A^T b = \begin{bmatrix} 23,8267 \\ 195,232 \end{bmatrix}$$

altså

$$\begin{bmatrix} 4 & 36 \\ 36 & 504 \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} = \begin{bmatrix} 23,8267 \\ 195,232 \end{bmatrix}$$

som gir at $ln(c_1) = 6,9171$ og $c_2 = -0.1067$. Innsatt i modellen blir det

$$y = e^{6,9171} e^{-0.1067t} = 1009,39 e^{-0.1067t}$$

For å finne RMSE, må man regne ut residualene først, der formlene for RMSE og residualene er som følger:

$$\begin{aligned} \text{RMSE} &= \sqrt{\frac{r_1^2 + r_2^2 + ... + r_m^2}{m}} \\ r &= b - A\bar{x} \end{aligned}$$

her er b lik y-verdiene gitt i oppgaveteksten $b = \begin{bmatrix} 1000 \\ 550 \\ 270 \\ 150 \end{bmatrix}$, og

$$A\bar{x} = \begin{bmatrix} 1009,39 \cdot e^{-0.1067\cdot 0} \\ 1009,39 \cdot e^{-0.1067\cdot 6} \\ 1009,39 \cdot e^{-0.1067\cdot 12} \\ 1009,39 \cdot e^{-0.1067\cdot 18} \end{bmatrix}$$

$$r = b - \ A\bar{x} = \begin{bmatrix} 1000 - 1009,39 \\ 550 - 532,14 \\ 270 - 280,54 \\ 150 - 147.90 \end{bmatrix} = \begin{bmatrix} -9,39 \\ 17,86 \\ -10.54 \\ 2.10 \end{bmatrix}$$

$$\text{RMSE} = \sqrt{\frac{(-9,39)^2 + (17,86)^2 + (-10,54)^2 + (2,10)^2}{4}} = 11.43$$