Lineær algebra
Kursusgang 5

Ved bestemmelse af mindste kvadraters løsning til (store) ligningssystemer vil man gerne anvende en metode der opfylder to krav:

- antallet af regneoperationer er “lille”
- metoden er stabil over for numeriske fejl (afrundinger).

Mulige metoder (ud over Gauss elimenation):

- QR-faktorisering (pensum på 1. semester)
- Cholesky dekomposition (vises her).

I det følgende vises hvordan man kan faktorisere matricen og derefter løse ligningssystemer ved hjælp af faktoriseringen.

I praksis vil man nok bruge andre metoder til at faktorisere.

Ligningssystem: $A\mathbf{x} = \mathbf{b}$, hvor søjlerne i A er lineært uafhængige. A er en $m \times n$ matrix.

QR-faktorisering

Skriv $A = QR$, hvor Q er en $m \times n$ matrix med ortogonale vektorer med længde 1, altså $Q^T Q = I_n$ og R er en $n \times n$ øvre triangulær matrix. Tallene på diagonalen af R er $\neq 0$, så R har en invers matrix.
Søjlerne i Q *kan* bestemmes ved Gram-Schmidt ortogonalisering af søjlerne i A.

$$A^T\mathbf{x} = A^T\mathbf{b} \Leftrightarrow (QR)^T(QR)\mathbf{x} = (QR)^T\mathbf{b} \Leftrightarrow R^TQ^TQR\mathbf{x} = R^TQ^T\mathbf{b}$$

Da R^T har invers kan dette forkortes $Q^TQR\mathbf{x} = Q^T\mathbf{b}$.
Da $Q^TQ = I$ er kan ligningen skrives

$$R\mathbf{x} = Q^T\mathbf{b}$$

Eksempel, QR.

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 5.2 \\ 8.9 \\ 3.4 \end{bmatrix}$$

Gram-Schmidt

$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$.

$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$, $\mathbf{v}_2 = \mathbf{u}_2 - \frac{\mathbf{u}_2 \cdot \mathbf{v}_1}{||\mathbf{v}_1||^2} \mathbf{v}_1 = \begin{bmatrix} \frac{1}{2} \\ 1 \\ \frac{1}{2} \end{bmatrix}$.

$\mathbf{w}_1 = \frac{\mathbf{v}_1}{||\mathbf{v}_1||} = \begin{bmatrix} \sqrt{\frac{1}{2}} \\ 0 \\ -\sqrt{\frac{1}{2}} \end{bmatrix}$, $\mathbf{w}_2 = \frac{\mathbf{v}_2}{||\mathbf{v}_2||} = \begin{bmatrix} \sqrt{\frac{1}{6}} \\ \sqrt{\frac{2}{3}} \\ \sqrt{\frac{1}{6}} \end{bmatrix}$.

$$Q = \begin{bmatrix} \sqrt{\frac{1}{2}} & \sqrt{\frac{1}{6}} \\ 0 & \sqrt{\frac{2}{3}} \\ -\sqrt{\frac{1}{2}} & \sqrt{\frac{1}{6}} \end{bmatrix}, \quad R = \begin{bmatrix} \sqrt{2} & -\sqrt{\frac{1}{2}} \\ 0 & \sqrt{\frac{3}{2}} \end{bmatrix}.$$

$$Q^T \begin{bmatrix} 5.2 \\ 8.9 \\ 3.4 \end{bmatrix} = \begin{bmatrix} 1.272 \\ 10.777 \end{bmatrix}$$

Vi skal nu løse

$$\sqrt{2}x_1 - \sqrt{\frac{1}{2}}x_2 = 1.272$$

$$\sqrt{\frac{3}{2}}x_2 = 10.777$$

Vi får

$$x_2 = 8.8, \quad x_1 = 5.3$$

Samme eksempel, Cholesky.

Normalligning

$$\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 1.8 \\ 12.3 \end{bmatrix}$$

$$\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} \sqrt{2} & 0 \\ -\sqrt{\frac{1}{2}} & \sqrt{\frac{2}{3}} \end{bmatrix} \begin{bmatrix} \sqrt{2} & -\sqrt{\frac{1}{2}} \\ 0 & \sqrt{\frac{2}{3}} \end{bmatrix}.$$

Sæt

$$\mathbf{z} = \begin{bmatrix} \sqrt{2} & -\sqrt{\frac{1}{2}} \\ 0 & \sqrt{\frac{2}{3}} \end{bmatrix} \mathbf{x}.$$

Så er

$$\begin{bmatrix} \sqrt{2} & 0 \\ -\sqrt{\frac{1}{2}} & \sqrt{\frac{2}{3}} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} 1.8 \\ 12.3 \end{bmatrix}$$

Løsning $z_1 = 1.272$ og $z_2 = 10.777$.

Vi skal nu løse

$$\sqrt{2}x_1 - \sqrt{\frac{1}{2}}x_2 = 1.272$$

$$\sqrt{\frac{3}{2}}x_2 = 10.777$$

Vi får

$$x_2 = 8.8, \quad x_1 = 5.3$$

Mindste kvadraters metode og Cholesky dekomposition

Vi ønsker at finde en mindste kvadraters løsning til det (inkonsistente) ligningssystem

$$A\mathbf{x} = \mathbf{b},$$

hvor A er en $m \times n$ matrix.

Vi betragter derfor normalligningen

$$A^T A\mathbf{x} = A^T\mathbf{b}.$$

Matricen $A^T A$ er en symmetrisk $n \times n$ matrix som er altså er diagonaliserbar med egenværdier ≥ 0.

Hvis normalligningen ikke har frie variable så er alle $A^T A$'s egenværdier > 0.

Vi betragter derfor en matrix A som er

- kvadratisk ($n \times n$),
- symmetrisk ($A^T = A$) og
- positiv definit (egenværdierne er > 0).

Cholesky dekomposition

Vi ønsker nu at skrive matricen

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

som et produkt

$$A = U^T U,$$

hvor

$$U = \begin{bmatrix} u_{11} & u_{12} & u_{13} & \cdots & u_{1n} \\ 0 & u_{22} & u_{23} & \cdots & u_{2n} \\ 0 & 0 & u_{33} & \cdots & u_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & u_{nn} \end{bmatrix}.$$

MATLAB:

chol(A)

beregner Cholesky dekomposition U

Betragt først 3×3 matricer.

$$\text{Lad } A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{bmatrix} \quad \text{og } U = \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix},$$

og antag at $A = U^T U$.

Så er

$$\begin{aligned} a_{11} &= u_{11}^2 \\ a_{12} &= u_{11}u_{12} \\ a_{13} &= u_{11}u_{13} \\ a_{22} &= u_{12}^2 + u_{22}^2 \\ a_{23} &= u_{12}u_{13} + u_{22}u_{23} \\ a_{33} &= u_{13}^2 + u_{23}^2 + u_{33}^2. \end{aligned}$$

Da vi ønsker at u_{11}, u_{22}, u_{33} skal være positive får vi

$$\begin{array}{lll} u_{11} = \sqrt{a_{11}} & u_{12} = \frac{a_{12}}{u_{11}} & u_{13} = \frac{a_{13}}{u_{11}} \\ & u_{22} = \sqrt{a_{22} - u_{12}^2} & u_{23} = \frac{a_{23} - u_{12}u_{13}}{u_{22}} \\ & & u_{33} = \sqrt{a_{33} - u_{13}^2 - u_{23}^2} \end{array}$$

Man kan bevise at når A er positiv definit så er alle udtryk, der tages kvadratrod af, positive.

For en generel $n \times n$ matrix A beregnes U på følgende måde:
Vi beregner én række ad gangen: række 1, række 2, ...

I første række er $u_{11} = \sqrt{a_{11}}$ og $u_{1\ell} = \frac{a_{1\ell}}{u_{11}}$, for $\ell \geq 2$.

I anden række er $u_{22} = \sqrt{a_{22} - u_{12}^2}$ og $u_{2\ell} = \frac{a_{2\ell} - u_{12}u_{1\ell}}{u_{22}}$, for $\ell \geq 3$.

Når rækkerne $1, \ldots, k-1$ så beregnes række k på følgende måde:

$$u_{kk} = \sqrt{a_{kk} - u_{1k}^2 - u_{2k}^2 - \ldots - u_{k-1,k}^2}$$

og

$$u_{k\ell} = \frac{a_{k\ell} - u_{1k}u_{1\ell} - \ldots - u_{k-1,k}u_{k-1,\ell}}{u_{kk}},$$

for $\ell \geq k+1$.

Cholesky dekomposition og mindste kvadraters metode

Vi vender tilbage til normalligningen

$$A^T A\mathbf{x} = A^T\mathbf{b}.$$

Vi finder en Cholesky dekomposition

$$A^T A = U^T U,$$

hvor U er en øvre triangulær matrix og kan dermed skrive normalligningen som

$$U^T U\mathbf{x} = \mathbf{y},$$

hvor $\mathbf{y} = A^T\mathbf{b}$.

Vi kan løse ligningen

$$U^T(U\mathbf{x}) = \mathbf{y}$$

ved at sætte $\mathbf{z} = U\mathbf{x}$ og løse
først

$$U^T\mathbf{z} = \mathbf{y}$$

og derefter

$$U\mathbf{x} = \mathbf{z}.$$

Disse ligninger kan let løses uden brug af rækkeoperationer, da U er triangulær.

Lad $\hat{\mathbf{x}}$ være en mindste kvadraters løsning til ligningssystemet $A\mathbf{x} = \mathbf{b}$ og lad

$$\hat{\mathbf{r}} = A\hat{\mathbf{x}} - \mathbf{b}$$

være residual vektoren.

Vi beregner kvadratsummen af residualerne (det er denne størrelse der minimeres ved mindste kvadraters metode):

$$\varphi = \hat{\mathbf{r}}^T\hat{\mathbf{r}} = (A\hat{\mathbf{x}} - \mathbf{b})^T(A\hat{\mathbf{x}} - \mathbf{b}) = \hat{\mathbf{x}}^T(A^TA\hat{\mathbf{x}} - A^T\mathbf{b}) - \mathbf{b}^TA\hat{\mathbf{x}} + \mathbf{b}^T\mathbf{b}.$$

Da $\hat{\mathbf{x}}$ er løsning til normalligningen $A^TA\hat{\mathbf{x}} - A^T\mathbf{b} = \mathbf{0}$ er

$$\varphi = \hat{\mathbf{r}}^T\hat{\mathbf{r}} = \mathbf{b}^T\mathbf{b} - \mathbf{b}^TA\hat{\mathbf{x}}.$$

Dette tal er positivt. Sæt

$$s = \sqrt{\varphi}.$$

(Dette er længden af vektoren $\hat{\mathbf{r}}$.)

Lad nu den øvre triangulær matrix U og vektoren $\mathbf{z}$ opfylde

$$A^T A = U^T U \quad \text{og} \quad U^T \mathbf{z} = A^T \mathbf{b}.$$

Sæt

$$U_2 = \begin{bmatrix} U & \mathbf{z} \\ \mathbf{0} & s \end{bmatrix}.$$

Så er

$$U_2^T U_2 = \begin{bmatrix} U^T & \mathbf{0} \\ \mathbf{z}^T & s \end{bmatrix} \begin{bmatrix} U & \mathbf{z} \\ \mathbf{0} & s \end{bmatrix} = \begin{bmatrix} U^T U & U^T \mathbf{z} \\ (U^T \mathbf{z})^T & \mathbf{z}^T \mathbf{z} + s^2 \end{bmatrix} = \begin{bmatrix} A^T A & A^T \mathbf{b} \\ (A^T \mathbf{b})^T & \mathbf{b}^T \mathbf{b} \end{bmatrix}.$$

Da $\mathbf{z} = (U^T)^{-1} A^T \mathbf{b}$ er

$$\mathbf{z}^T \mathbf{z} = \mathbf{b}^T A U^{-1} (U^T)^{-1} A^T \mathbf{b} = \mathbf{b}^T A (U^T U)^{-1} A^T \mathbf{b} =$$

$$\mathbf{b}^T A (A^T A)^{-1} A^T \mathbf{b} = \mathbf{b}^T A \hat{\mathbf{x}},$$

ifølge normalligningen.

Dermed er

$$\mathbf{z}^T\mathbf{z} + s^2 = \mathbf{b}^T A\hat{\mathbf{x}} + \mathbf{b}^T\mathbf{b} - \mathbf{b}^T A\hat{\mathbf{x}} = \mathbf{b}^T\mathbf{b}.$$

$U_2^T U_2$ er altså Cholesky dekomposition af

$$\begin{bmatrix} A^T A & A^T\mathbf{b} \\ (A^T\mathbf{b})^T & \mathbf{b}^T\mathbf{b} \end{bmatrix}$$

og U_2 kan derfor bestemmes ved hjælp af algoritmen til beregning af Cholesky dekomposition.

Tallet s på den sidste diagonalplads i U_2 opfylder altså at s^2 er kvadratsummen af residualerne.

Mellem fem punkter A, B, C, D, E er der målt følgende højdeforskelle:

Fra punkt	Til punkt	Højdeforskel
B	A	8
B	C	2
B	E	5
C	A	7
C	D	1
D	A	6
D	E	4
E	A	3

Sæt

$$A = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 \\ 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 & 1 \\ 1 & 0 & 0 & 0 & -1 \end{bmatrix} \quad \text{og} \quad \mathbf{b} = \begin{bmatrix} 8 \\ 2 \\ 5 \\ 7 \\ 1 \\ 6 \\ 4 \\ 3 \end{bmatrix},$$

og find mindste kvadraters løsning til

$$A\mathbf{x} = \mathbf{b}.$$

Vi skal bruge

$$A^TA = \begin{bmatrix} 4 & -1 & -1 & -1 & -1 \\ -1 & 3 & -1 & 0 & -1 \\ -1 & -1 & 3 & -1 & 0 \\ -1 & 0 & -1 & 3 & -1 \\ -1 & -1 & 0 & -1 & 3 \end{bmatrix},$$

$$A^T\mathbf{b} = \begin{bmatrix} 24 \\ -15 \\ -6 \\ -9 \\ 6 \end{bmatrix}$$

og

$$\mathbf{b}^T\mathbf{b} = 204$$

Vi udregner Cholesky dekoposition af $A^T A$

$$U = \begin{bmatrix} 2.0 & -0.500 & -0.500 & -0.500 & -0.500 \\ 0.0 & 1.66 & -0.754 & -0.151 & -0.754 \\ 0.0 & 0.0 & 1.48 & -0.926 & -0.554 \\ 0.0 & 0.0 & 0.0 & 1.37 & -1.37 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \end{bmatrix}.$$

Problem: Matricen $A^T A$ er ikke positiv definit.
Men udregning af Cholesky dekompositionen går godt fordi det kun er det sidste diagonalelement der er 0.
Hvis der ikke er frie variable i løsningen så er $A^T A$ er positiv definit.

Vi ser også på følgende matrix

$$\begin{bmatrix} A^TA & A^T\mathbf{b} \\ (A^T\mathbf{b})^T & \mathbf{b}^T\mathbf{b} \end{bmatrix} = \begin{bmatrix} 4 & -1 & -1 & -1 & -1 & 24 \\ -1 & 3 & -1 & 0 & -1 & -15 \\ -1 & -1 & 3 & -1 & 0 & -6 \\ -1 & 0 & -1 & 3 & -1 & -9 \\ -1 & -1 & 0 & -1 & 3 & 6 \\ 24 & -15 & -6 & -9 & 6 & 204 \end{bmatrix}.$$

Stort problem: Forsøg på at udregne Cholesky dekomposition af denne matrix fører til division med 0, idet $u_{55} = 0$.

Vi får nu oplyst at E er et fikspunkt med højde 10.

Vi sætter så

$$A = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 0 & 0 & -1 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \end{bmatrix} \quad \text{og} \quad \mathbf{b} = \begin{bmatrix} 8 \\ 2 \\ -5 \\ 7 \\ 1 \\ 6 \\ -6 \\ 13 \end{bmatrix},$$

og betragter ligningen

$$A\mathbf{x} = \mathbf{b}.$$

Vi udregner

$$\begin{bmatrix} A^TA & A^T\mathbf{b} \\ (A^T\mathbf{b})^T & \mathbf{b}^T\mathbf{b} \end{bmatrix} = \begin{bmatrix} 4 & -1 & -1 & -1 & 34 \\ -1 & 3 & -1 & 0 & -5 \\ -1 & -1 & 3 & -1 & -6 \\ -1 & 0 & -1 & 3 & 1 \\ 34 & -5 & -6 & 1 & 384 \end{bmatrix}.$$

Vi finder Cholesky demposition af denne matrix

$$U_2 = \begin{bmatrix} 2.0 & -0.500 & -0.500 & -0.500 & 17.0 \\ 0.0 & 1.66 & -0.754 & -0.151 & 2.11 \\ 0.0 & 0.0 & 1.48 & -0.926 & 2.77 \\ 0.0 & 0.0 & 0.0 & 1.37 & 9.04 \\ 0.0 & 0.0 & 0.0 & 0.0 & 1.10 \end{bmatrix}.$$

Vi ser at Cholesky dekopositionen af $A^T A$ er

$$U = \begin{bmatrix} 2.0 & -0.500 & -0.500 & -0.500 \\ 0.0 & 1.66 & -0.754 & -0.151 \\ 0.0 & 0.0 & 1.48 & -0.926 \\ 0.0 & 0.0 & 0.0 & 1.37 \end{bmatrix},$$

og at

$$\mathbf{z} = \begin{bmatrix} 17.0 \\ 2.11 \\ 2.77 \\ 9.04 \end{bmatrix}$$

er løsningen til ligningen $U^T \mathbf{z} = A^T \mathbf{b}$.

Den mindste kvadraters løsning til $A\mathbf{x} = \mathbf{b}$ er løsning $\hat{\mathbf{x}}$ til ligningen $U\mathbf{x} = \mathbf{z}$.

Vi får

$$\hat{\mathbf{x}} = \begin{bmatrix} 12.80 \\ 4.60 \\ 6.00 \\ 6.60 \end{bmatrix}.$$

Desuden ser vi at kvadratsummen af residualerne er $1.10^2 = 1.20$.

Vægtet mindste kvadraters metode

Find vægtet mindste kvadraters løsning til

$$A\mathbf{x} = \mathbf{b}$$

med vægtmatrix C.

Normalligning

$$A^T C A\mathbf{x} = A^T C\mathbf{b}.$$

Vi ser på matricen

$$\begin{bmatrix} A^T C A & A^T C\mathbf{b} \\ (A^T C\mathbf{b})^T & \mathbf{b}^T C\mathbf{b} \end{bmatrix}$$

og dennes Cholesky dekomposition

$$U_2 = \begin{bmatrix} U & \mathbf{z} \\ \mathbf{0} & s \end{bmatrix}.$$

$\hat{\mathbf{x}}$ er løsning til $U\mathbf{x} = \mathbf{z}$
og s^2 er den vægtede kvadratsum af residualer.

Hvis A er en 4×4 positiv definit matrix så beregnes U ved:

$$
\begin{array}{llll}
u_{11} = \sqrt{a_{11}} & u_{12} = \frac{a_{12}}{u_{11}} & u_{13} = \frac{a_{13}}{u_{11}} & u_{14} = \frac{a_{14}}{u_{11}} \\
 & u_{22} = \sqrt{a_{22} - u_{12}^2} & u_{23} = \frac{a_{23} - u_{12}u_{13}}{u_{22}} & u_{24} = \frac{a_{24} - u_{12}u_{14}}{u_{22}} \\
 & & u_{33} = \sqrt{a_{33} - u_{13}^2 - u_{23}^2} & u_{34} = \frac{a_{34} - u_{13}u_{14} - u_{23}u_{24}}{u_{33}} \\
 & & & u_{44} = \sqrt{a_{44} - u_{14}^2 - u_{24}^2 - u_{34}^2}
\end{array}
$$