Lineær algebra
Lektion 6

En $m \times n$ matrix A er kvadratisk hvis $m = n$.

En kvadratisk matrix A er symmetrisk hvis $A^T = A$.

Idet $(BC)^T = C^T B^T$ er enhver matrix på formen $A^T A$ symmetrisk.

En vektor $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} \neq \mathbf{0}$ er egenvektor for en $n \times n$ matrix A hvis der findes et tal (en egenværdi) λ så $A\mathbf{x} = \lambda\mathbf{x}$.

En matrix A er diagonaliserbar hvis der findes en basis $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ for $\mathbb{R}^n$, hvor hver vektor er en egenvektor $A\mathbf{x}_i = \lambda_i \mathbf{x}_i$. Dette er ækvivalent med at der findes en matrix P og en diagonalmatrix D så $P^{-1}AP = D$.

Enhver symmetrisk matrix er diagonaliserbar.

Enhver matrix på formen $A^T A$ er altså diagonaliserbar.

Hvis λ er en egenværdi for $A^T A$ med en tilhørende egenvektor $\mathbf{x}$, $A^T A\mathbf{x} = \lambda\mathbf{x}$, så er

$$\lambda||\mathbf{x}||^2 = \lambda\mathbf{x}\cdot\mathbf{x} = A^T A\mathbf{x}\cdot\mathbf{x} = (A^T A\mathbf{x})^T\mathbf{x} =$$

$$\mathbf{x}^T A^T A\mathbf{x} = (A\mathbf{x})^T A\mathbf{x} = ||A\mathbf{x}||^2 \geq 0.$$

Altså $\lambda \geq 0$.

Hvis søjlerne i A er lineært uafhængige så gælder der faktisk $\lambda > 0$.

En symmetrisk matrix hvis egenværdier alle er > 0 siges at være positiv definit.

En kvadratisk matrix U siges at være en øvre (upper) triangulær matrix hvis der står 0 på alle pladser under diagonalen.

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

En kvadratisk matrix L siges at være en nedre (lower) triangulær matrix hvis der står 0 på alle pladser over diagonalen.

$$L = \begin{bmatrix} \ell_{11} & 0 & 0 & \cdots & 0 \\ \ell_{21} & \ell_{22} & 0 & \cdots & 0 \\ \ell_{31} & \ell_{32} & \ell_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \ell_{n1} & \ell_{n2} & \ell_{n3} & \cdots & \ell_{nn} \end{bmatrix}.$$

U er en øvre triangulær matrix hvis og kun hvis U^T er nedre triangulær matrix.

En matrix D er en diagonalmatrix hvis og kun hvis den er både øvre triangulær og nedre triangulær.

Cholesky dekomposition

Vi ønsker nu at skrive en $n \times n$ matrix A som et produkt

$$A = U^T U,$$

hvor U er en øvre triangulær matrix.
Dette kaldes Cholesky dekompositionen af A.

I litteraturen skriver man ofte Cholesky dekompositionen som

$$A = LL^T,$$

hvor $L = U^T$ er en nedre triangulær matrix.

MATLAB: chol(A) beregner Cholesky dekomposition U
Maple: LUDecomposition(A,method=Cholesky) beregner $L = U^T$

Enhver matrix på formen $U^T U$ er symmetrisk og hvis tallene på diagonalen i U er $\neq 0$ så er søjlerne lineært uafhængige og dermed er $U^T U$ positiv definit.

Vi vil derfor antage at A er symmetrisk, positiv definit matrix.

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

$$a_{11} = u_{11}^2$$
$$a_{12} = u_{11}u_{12}$$
$$a_{13} = u_{11}u_{13}$$
$$a_{22} = u_{12}^2 + u_{22}^2$$
$$a_{23} = u_{12}u_{13} + u_{22}u_{23}$$
$$a_{33} = u_{13}^2 + u_{23}^2 + u_{33}^2.$$

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

$$u_{11} = \sqrt{a_{11}} \quad u_{12} = \frac{a_{12}}{u_{11}} \quad u_{13} = \frac{a_{13}}{u_{11}}$$
$$u_{22} = \sqrt{a_{22} - u_{12}^2} \quad u_{23} = \frac{a_{23} - u_{12}u_{13}}{u_{22}}$$
$$u_{33} = \sqrt{a_{33} - u_{13}^2 - u_{23}^2}$$

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

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

Vi betragter derfor normalligningen

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

Matricen $A^T A$ er symmetrisk og da søjlerne i A sædvanligvis er lineært uafhængige er $A^T A$ også positiv definit.

Vi kan derfor finde en Cholesky dekomposition

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

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

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

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

Vi kan løse denne ligning 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

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

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

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

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

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

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

ifølge normalligningen.
Dermed er

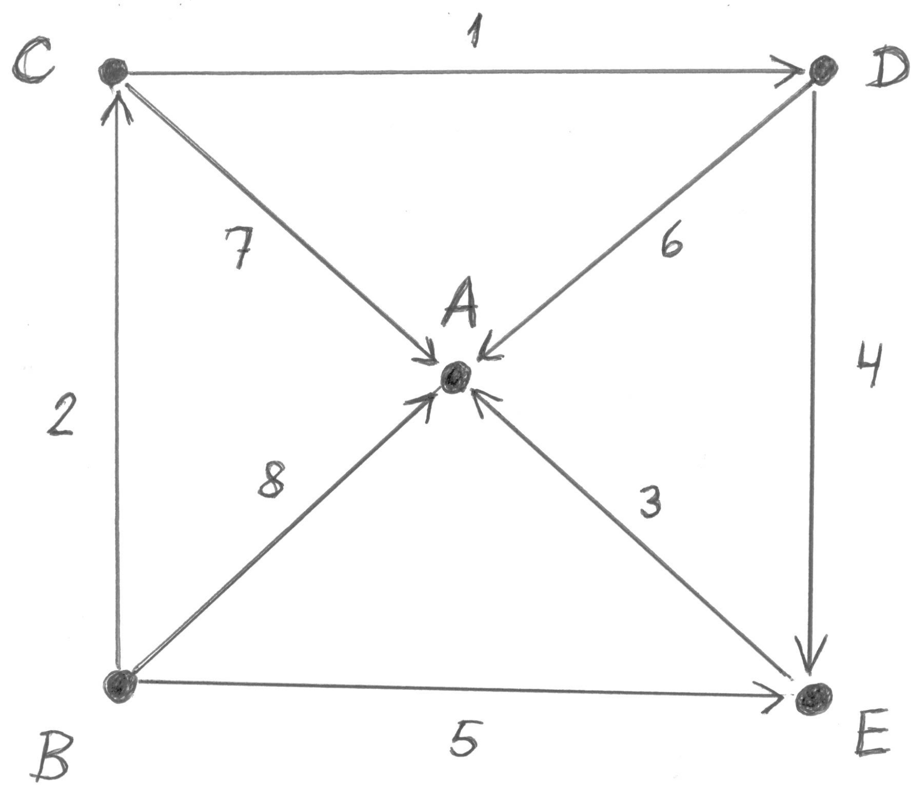
$$\mathbf{z}^T\mathbf{z} + s^2 = \mathbf{b}^T A\hat{\mathbf{x}} + \mathbf{b}^T\mathbf{b} - \mathbf{b}^T A\hat{\mathbf{x}} = \mathbf{b}^T\mathbf{b}.$$

$U_2^T U_2$ 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.



C
1
D
7
6
A
2
4
8
3
B
5
E

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

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

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

og

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

Vi udregner Cholesky dekoposition af A^TA

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

Problem: Matricen A^TA er ikke positiv definit.
Men udregning af Cholesky dekompositionen går godt fordi kun det sidste diagonalelement er 0.
Hvis der ikke er frie variable i løsningen så er der pivotposition i alle søjler i A. Søjlerne er dermed lineært uafhængige og A^TA 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.