Lineær algebra

Kursusgang 5

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 (matricen $A^T A$ på forrige side) 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

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

$$U^T U = \begin{bmatrix} u_{11} & 0 & 0 \\ u_{12} & u_{22} & 0 \\ u_{13} & u_{23} & u_{33} \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix}$$

$$= \begin{bmatrix} u_{11}^2 & u_{11} u_{12} & u_{11} u_{13} \\ \cdot & u_{12}^2 + u_{22}^2 & u_{12} u_{13} + u_{22} u_{23} \\ \cdot & \cdot & u_{13}^2 + u_{23}^2 + u_{33}^2 \end{bmatrix}$$

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

$$
\begin{aligned}
u_{11} &= \sqrt{a_{11}} \quad & u_{12} &= \tfrac{a_{12}}{u_{11}} & u_{13} &= \tfrac{a_{13}}{u_{11}} \\
& & u_{22} &= \sqrt{a_{22} - u_{12}^2} & u_{23} &= \tfrac{a_{23} - u_{12}u_{13}}{u_{22}} \\
& & & & u_{33} &= \sqrt{a_{33} - u_{13}^2 - u_{23}^2}
\end{aligned}
$$

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

$A\vec{x} = \lambda\vec{x}$ $\qquad$ λ egenværdi

$\vec{x}$ egenvektor

$\cancel{A^TA}$ $A = \underline{B^T B} \Rightarrow \lambda \geq 0$

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

$u_{11} = \sqrt{a_{11}}$ $\quad u_{12} = \frac{a_{12}}{u_{11}}$ $\quad u_{13} = \frac{a_{13}}{u_{11}}$ $\quad u_{14} = \frac{a_{14}}{u_{11}}$

$u_{22} = \sqrt{a_{22} - u_{12}^2}$ $\quad u_{23} = \frac{a_{23} - u_{12}u_{13}}{u_{22}}$ $\quad u_{24} = \frac{a_{24} - u_{12}u_{14}}{u_{22}}$

$u_{33} = \sqrt{a_{33} - u_{13}^2 - u_{23}^2}$ $\quad u_{34} = \frac{a_{34} - u_{13}u_{14} - u_{23}u_{24}}{u_{33}}$

$u_{44} = \sqrt{a_{44} - u_{14}^2 - u_{24}^2 - u_{34}^2}$

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

$$A = \begin{bmatrix} 1 & 2 & 0 & -1 \\ 2 & 5 & 1 & -2 \\ 0 & 1 & 5 & 2 \\ -1 & -2 & 2 & 3 \end{bmatrix}$$

$u_{11} = \sqrt{1} = 1, \quad u_{12} = 2, \quad u_{13} = 0$

$u_{14} = -1$

$u_{22} = \sqrt{5 - 2^2} = 1 \qquad u_{23} = \frac{1 - 2 \cdot 0}{1} = 1$

$u_{24} = \frac{-2 - 2 \cdot (-1)}{1} = 0$

$$u_{33} = \sqrt{5 - 0^2 - 1^2} = 2, \quad u_{34} = \frac{2 - 0 - 0}{2} = 1, \quad u_{44} = \sqrt{3 - (-1)^2 - 0^2 - 1^2} = 1$$

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

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

Vi finder en Cholesky dekomposition

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

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

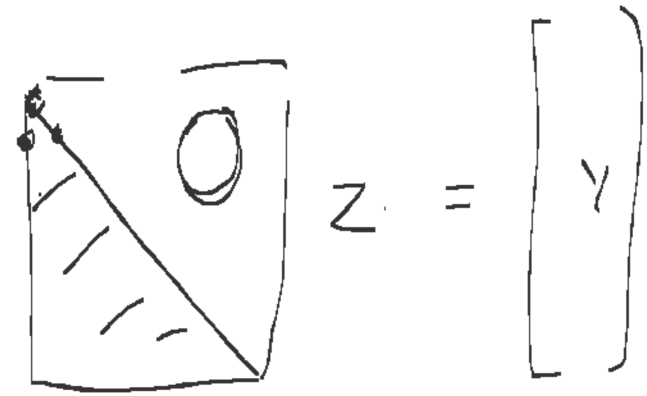
$$U^T U\mathbf{x} = \mathbf{y},$$

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 $\hat{\mathbf{r}}$.)

$$\phi = (A\hat{x} - b)^T(A\hat{x} - b) = (\hat{x}^TA^T - b^T)(A\hat{x} - b) =$$

$\hat{x}^T A^T (A\hat{x} - b) - b(A\hat{x} - b) = \hat{x}^T \underbrace{(A^T A\hat{x} - A^T b)}_{=0} - b(A\hat{x} - b)$

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

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

øvre triangulær

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

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

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

ifølge normalligningen.

$A^T A\hat{x} = A^T b \Rightarrow \hat{x} = (A^T A)^{-1}A^T b \qquad (U^T)^{-1} = (U^{-1})^T$

$$s^2 = b^T b - b^T A \hat{x}$$

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

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.

Eksempel

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 $A^T A$

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

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

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

Dekomposition af A^TA

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

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

$1.37\, x_4 = 9{,}04$

$\Rightarrow x_4 = 6{,}6$

$1{,}48\, x_3 - 0{,}926 \cdot 6{,}6 = 2{,}77$

$\Rightarrow x_3 = 6{,}0$

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

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

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

$$C = \begin{bmatrix} c_1 & & & 0 \\ & c_2 & & \\ & & \ddots & \\ 0 & & & c_n \end{bmatrix}$$

$$W = \begin{bmatrix} \sqrt{c_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{c_n} \end{bmatrix}$$

$$C = W^T W$$

$$A^T C A = A^T W^T W A$$

$$= (WA)^T WA$$

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