IDAG: Gram-Schmidts metod

Mål:
- Visa fler egenskaper av ortogonala projektioner
- Konstruera ortonormala baser

Låt W vara ett underrum av $\mathbb{R}^n$ med en ortogonal bas $B = \{u_1, \ldots, u_p\}$. Kom ihåg att den ortogonala projektionen av $x \in \mathbb{R}^n$ på W ges av

$$\text{proj}_W x = \frac{x \cdot u_1}{u_1 \cdot u_1} u_1 + \cdots + \frac{x \cdot u_p}{u_p \cdot u_p} u_p .$$

$\text{proj}_W x$ är den närmsta punkten till x i W:

(Detta tas ibland som definitionen av $\text{proj}_W x$ och sen härleds formeln från detta.)

Sats 6.3.9 För alla $x \in \mathbb{R}^n$ har vi att

$$\|x - \text{proj}_W x\| \leq \|x - y\| \quad \forall y \in W,$$

med likhet endast för $y = \text{proj}_W x$.

Dvs. $\text{dist}(x, \text{proj}_W x) \leq \text{dist}(x, y) \quad \forall y \in W$.

Bevis: $y \in W \Rightarrow \text{proj}_W x - y \in W$

Vi vet att $x - \text{proj}_W x$ är ortogonal mot W, så $x - \text{proj}_W x$ och $\text{proj}_W x - y$ är ortogonala

Pythagoras sats:

$$\underbrace{\|(x - \text{proj}_W x) + (\text{proj}_W x - y)\|^2}_{= \|x-y\|^2} = \|x - \text{proj}_W x\|^2 + \underbrace{\|\text{proj}_W x - y\|^2}_{\geq 0}$$

dvs. $\|x - y\|^2 \geq \|x - \text{proj}_W x\|^2$ med likhet endast om $y = \text{proj}_W x$.

Nu går vi över till matriser igen.

Def. En matris $U \in \mathbb{R}^{n \times n}$ är ortogonal om dess kolonner är ortonormala.

(Olyckligt ordval som använts för länge för att ändras nu.)

Sats 6.2.6 En matris $U \in \mathbb{R}^{m \times n}$ har ortonormala kolonner omm $U^T U = I_n$. Om $U \in \mathbb{R}^{n \times n}$ är ortogonal har vi dessutom att $U^T U = I_n = U U^T$, dvs. U är inverterbar och $U^{-1} = U^T$.

(Mycket användbart! Behöver inte lösa L.E.S., transponera bara U och multiplicera.)

Bevis: Låt $U \in \mathbb{R}^{m \times n}$ och $U = [u_1 \cdots u_n]$, $u_k \in \mathbb{R}^m$.

Vi får $U^T = \begin{bmatrix} u_1^T \\ \vdots \\ u_n^T \end{bmatrix}$ så

$$U^T U = \begin{bmatrix} u_1^T u_1 & u_1^T u_2 & \cdots & u_1^T u_n \\ u_2^T u_1 & u_2^T u_2 & & \\ \vdots & & \ddots & \\ u_n^T u_1 & & & u_n^T u_n \end{bmatrix}$$

Om $\{u_k\}$ är ortonormala så har vi $u_i^T u_j = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}$

Dvs. $U^T U = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & \\ \vdots & & \ddots & \\ 0 & & & 1 \end{bmatrix} = I_n$

Är $U \in \mathbb{R}^{n \times n}$ så ger Satsen om inverterbara matriser att $U^T = U^{-1}$ och alltså $U U^T = I_n$. □

Sats 6.2.7 Låt $U \in \mathbb{R}^{m \times n}$ ha ortonormala kolonner och låt $x, y \in \mathbb{R}^n$.

Då gäller
(I) $\|Ux\| = \|x\|$
(II) $(Ux) \cdot (Uy) = x \cdot y$.

Bevis: (II): $(Ux) \cdot (Uy) = (Ux)^T Uy = x^T \underbrace{U^T U}_{= I_n} y$

$$= x^T y = x \cdot y$$

(I): $\|Ux\|^2 = (Ux) \cdot (Ux) \overset{\text{(II)}}{=} x \cdot x = \|x\|^2$. □

Sats Om $U \in \mathbb{R}^{n \times n}$ är ortogonal så är U:s rader också ortonormala.

Bevis: U^T är också ortogonal då $U^T U = I = U U^T$, och U:s rader är U^T:s kolonner.

Med hjälp av detta kan vi skriva projektionsformeln på ett enklare sätt:

Sats 6.3.10 Om $U = [u_1 \cdots u_p]$ och dess kolonner är en ortonormal bas för underrummet $W \subset \mathbb{R}^n$ så har vi för $x \in \mathbb{R}^n$ att

$$\text{proj}_W x = U U^T x = (x \cdot u_1) u_1 + \dots + (x \cdot u_p) u_p .$$

Bevis: $U^T x = \begin{bmatrix} u_1^T x \\ \vdots \\ u_p^T x \end{bmatrix} = \begin{bmatrix} x \cdot u_1 \\ \vdots \\ x \cdot u_p \end{bmatrix}$ så sista likheten håller.

Resten följer av att $u_i \cdot u_i = 1$ då mängden är ortonormal. □

Notera att om $U \in \mathbb{R}^{m \times n}$ (inte kvadratisk) så är

$U^T U = I_n$ men $U U^T \neq I_n$ och inte heller $U U^T \neq I_m$.

Gram-Schmidt

Antag att $W = \text{Span}\{x_1, \dots, x_p\} \subset \mathbb{R}^n$.

Vi vill hitta en ortogonal bas $\{v_1, \dots, v_r\}$ för W.

Vi vet redan att vi kan ta bort lin.ber. vektorer v_k och få en bas (Sats 4.3.5), så vi antar att $\{x_1, \dots, x_p\}$ är en bas för W. (Ej ortogonal.)

Idén bakom Gram-Schmidts metod är liknande:

> Ta bort bidraget i x_1-riktningen från $x_2, \dots, x_p$, kalla resultatet $x_1, y_2, y_3, \dots, y_p$. Ta sen bort bidraget i y_2-riktningen från $y_3, \dots, y_p$, osv.

Ex.

$$W = \text{Span}\left\{ \underbrace{\begin{bmatrix} -1 \\ 2 \\ 0 \\ 1 \end{bmatrix}}_{x_1}, \underbrace{\begin{bmatrix} 3 \\ -1 \\ 4 \\ 2 \end{bmatrix}}_{x_2}, \underbrace{\begin{bmatrix} 1 \\ 0 \\ 2 \\ 3 \end{bmatrix}}_{x_3} \right\}.$$

($\{x_1, x_2, x_3\}$ lin.ober. $\Rightarrow$ bas, men t.ex. $x_1 \cdot x_2 \neq 0$.)

I. Låt $v_1 = x_1$ och $W_1 = \{v_1\}$. Då är W_1 ortogonal och $\text{Span}\{W_1\} = \text{Span}\{x_1\}$.

II. Låt $v_2 = x_2 - \text{proj}_{W_1} x_2$ [Ta bort x_1-bidraget]

$$= x_2 - \frac{x_2 \cdot v_1}{v_1 \cdot v_1} v_1$$

$$= \begin{bmatrix} 3 \\ -1 \\ 4 \\ 2 \end{bmatrix} - \frac{-3}{6} \begin{bmatrix} -1 \\ 2 \\ 0 \\ 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 5 \\ 0 \\ 8 \\ 5 \end{bmatrix}$$

Ta $v_2' = 2v_2$ (för att underlätta räkningarna) och låt $W_2 = \{v_1, v_2'\}$. Då $v_2' \in W_1^\perp$ så är W_2 ortogonal och $W_2 = \text{Span}\{x_1, x_2\}$ då x_1, x_2 lin.ober.

$\longrightarrow$

Ex. forts.

(III) Låt $v_3 = x_3 - \text{proj}_{W_2} x_3$

$$= x_3 - \frac{x_3 \cdot v_1}{v_1 \cdot v_1} v_1 - \frac{x_3 \cdot v_2'}{v_2' \cdot v_2'} v_2'$$

$$= \begin{bmatrix} 1 \\ 0 \\ 2 \\ 3 \end{bmatrix} - \frac{2}{6} \begin{bmatrix} -1 \\ 2 \\ 0 \\ 1 \end{bmatrix} - \frac{36}{114} \begin{bmatrix} 5 \\ 0 \\ 8 \\ 5 \end{bmatrix}$$

$$= \dots = \frac{1}{57} \begin{bmatrix} -14 \\ -38 \\ -30 \\ 62 \end{bmatrix}$$

$v_3 \in W_2^\perp$ så $W_3 = \{v_1, v_2', v_3\} = W_2 \cup \{v_3\}$ är ortogonal. Då $\{x_1, x_2, x_3\}$ är lin. ober. får vi också att $v_3 \neq 0$. En ortogonal mängd är linjärt oberoende, så enligt bassatsen har vi att W_3 är en (ortogonal) bas för W.

(Då $\dim W = 3$.)

Sats 6.4.11

Låt $\{x_1, \dots, x_p\}$ vara en bas för $W \neq \{0\}$.

Då ges en ortogonal bas för W av $\{v_1, \dots, v_p\}$

där $v_1 = x_1$ och $v_{k+1} = x_{k+1} - \frac{x_{k+1} \cdot v_1}{v_1 \cdot v_1} v_1 - \dots - \frac{x_{k+1} \cdot v_k}{v_k \cdot v_k} v_k$

för $k = 1, \dots, p-1$.

Bevis: Sätt $W_k = \text{Span}\{x_1, \dots, x_k\}$. Då är $\text{Span}\{v_1\} = W_1$ så vi har en ortogonal bas för W_1.

Antag att $\{v_1, \dots, v_k\}$ är en ortogonal bas för W_k.

Eftersom $v_{k+1} = x_{k+1} - \text{proj}_{W_k} x_{k+1}$ så är $v_{k+1} \in W_k^\perp$.

$v_{k+1} \neq 0$ då $x_{k+1} \notin W_k$ och $\{x_1, \dots, x_k\}$ är lin. ober.

$\Rightarrow \{v_1, \dots, v_{k+1}\} \subset W_{k+1}$ är en ortogonal mängd av $k+1$ icke-noll vektorer. Ortogonal $\Rightarrow$ lin. ober. så bassatsen ger att de är en ortogonal bas för W_{k+1}. ($\dim W_{k+1} = k+1$). $\longrightarrow$

Bevis forts.

Via induktion får vi ortogonala baser för

$W_1, W_2, \dots, W_k, \dots$

Till slut har vi en ortogonal bas för $W_p = W$. □

QR-faktorisering

I början av kursen såg vi att en LU-faktorisering av A består av att spara både operationerna (L) och resultatet (U) av en radreducering av A.

En QR-faktorisering sparar i princip istället operationerna (R) och resultatet (Q) av en ortogonalisering av A:s kolonner.

Sats 6.4.12 Låt $A = [a_1 \cdots a_n] \in \mathbb{R}^{m\times n}$ ha lin. ober. kolonner $\{a_j\}$. Då finns $Q \in \mathbb{R}^{m\times n}$ med ortonormala kolonner och en övertriangulär (inv.) matris $R \in \mathbb{R}^{n\times n}$ med positiva element på diagonalen så att $A = QR$.

Bevis: Gram-Schmidt ger en ortonormal bas $\{u_1, \dots, u_n\}$ för Col A sådan att $\operatorname{Span}\{u_1,\dots,u_k\} = \operatorname{Span}\{a_1,\dots,a_k\}$.

Låt $Q = [u_1 \cdots u_n]$ och observera att $a_k \in$ betyder att

$$a_k = r_{1k}u_1 + r_{2k}u_2 + \dots + r_{kk}u_k + 0u_{k+1} + \dots + 0u_n$$

$$= Q \underbrace{\begin{bmatrix} r_{1k} \\ \vdots \\ r_{kk} \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_{r_k}$$ för någon vektor r_k

Om $r_{kk} < 0$ byter vi u_k mot $-u_k$ $\Rightarrow r_{kk} > 0$.

Men då har vi med $R = [r_1 \cdots r_n]$ att

$$A = [a_1 \cdots a_n] = [Qr_1 \cdots Qr_n] = QR.$$

R övertriangulär med lin.ober. kolonner $\Rightarrow$ inv. och $r_{kk} > 0$.

□