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 $\mathcal{B}=\{u_1,\dots,u_p\}$. Kom ihåg att den ortogonala projektionen av $x\in\mathbb{R}^n$ på W ges av

$$\mathrm{proj}_W\, x = \frac{x\cdot u_1}{u_1\cdot u_1}u_1 + \dots + \frac{x\cdot u_p}{u_p\cdot u_p}u_p .$$

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

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

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

$$\|x-\mathrm{proj}_W\, x\| \le \|x-y\| \quad \forall y\in W,$$

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

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

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

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

Pythagoras sats:

$$\underbrace{\|(x-\mathrm{proj}_W\, x)+(\mathrm{proj}_W\, x-y)\|^2}_{=\|x-y\|^2} = \|x-\mathrm{proj}_W\, x\|^2 + \underbrace{\|\mathrm{proj}_W\, x-y\|^2}_{\ge 0}$$

dvs. $\|x-y\|^2 \ge \|x-\mathrm{proj}_W\, x\|^2$ med likhet endast om $y=\mathrm{proj}_W\, x$.

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 $U U^T = I_n$. $\square$

$\longrightarrow$

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

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 $U^T U = I = UU^T$, 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

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

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. $\square$

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

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

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.

Lag antar att vi redan har en (icke ortogonal) bas, då Sats 4.3.5 säger att vi alltid kan hitta en sådan genom att ta bort lin.ber. vektorer.

Detta kan dock inkluderas i algoritmen. Idén är att

> 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 y_2-bidraget från $y_3, \dots, y_p$, osv.
> Får vi en nollvektor har vi hittat en lin.ber. x_j

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. 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 från x_2]

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

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

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.

Ex. forts.

III Låt $v_3 = x_3 - \text{proj}_{W_2} x_3$ [Ta bort x_1- och x_2-bidragen från 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 lin.ober., och eftersom $\dim W = 3$ säger bassatsen att W_3 är en ortogonal bas för W.

Sats 6.4.11+

Låt $W = \text{Span}\{x_1, \dots, x_p\} \neq \{0\}$. Då ges en ortogonal bas för W av $V_p = \{v_j \mid v_j \neq 0,\ j = 1, \dots, p\}$, där $v_1 = x_1$ och $v_{k+1} = x_{k+1} - \text{proj}_{V_k} x_{k+1}$, $k = 1, \dots, p-1$.

Bevis: Sätt $W_k = \text{Span}\{x_1, \dots, x_k\}$, dvs. $W_p = W$.

Då är $v_1 = x_1$ en ortogonal bas för W_1.

Resten är induktion:

Antag att $V_k := \{v_j \mid v_j \neq 0,\ j = 1, \dots, k\}$ är en ortogonal bas för W_k. Per def. är $v_{k+1} \in W_k^\perp$.

$\Rightarrow$ V_{k+1} är en ortogonal mängd.

$\Rightarrow$ V_{k+1} är en lin.ober. mängd

Spänner den upp W_{k+1}?

$\longrightarrow$

Två fall:

- $x_{k+1} \in W_k$, dvs. x_{k+1} lin. ber. av $x_1, \dots, x_k$

  $\Rightarrow$ v_{k+1} behövs inte för att spänna upp W_{k+1},

  det räcker med $V_k \subset V_{k+1}$. $\Rightarrow \operatorname{Span}(V_{k+1}) = W_{k+1}$.

  (I det här fallet blir $v_{k+1} = 0$ och $V_{k+1} = V_k$.)

- $x_{k+1} \notin W_k \Rightarrow v_{k+1} \neq 0$, då vi tar bort

  $\operatorname{proj}_{W_k} x_{k+1} \in W_k$.

  $\Rightarrow$ V_{k+1} är en lin. ober. mängd av icke-noll element,

  där $\#\{V_{k+1}\} = \#\{V_k\} + 1$ (# = antal element)

Enligt bassatsen så är V_{k+1} en bas för ett rum av dimension $\dim V_k + 1$, vilket måste vara W_{k+1}. (Ty $W_k \subset V_{k+1} \subset W_{k+1}$ och $\dim W_{k+1} = \dim W_k + 1$.)

Via induktion får vi ortogonala baser $V_1, V_2, \dots$ för $W_1, W_2, \dots$. Till slut har vi den ortogonala basen V_p för $W_p = W$. $\square$

QR-faktorisering

En LU-faktorisering sparar operationerna (L) och resultatet (U) av en radreducering av $A = LU$.

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

$$A = \underset{\text{ortogonal}}{Q}\ \underset{\text{övertriangulär}}{R}$$

Detta görs i princip via Gram-Schmidt. Läs Sats 6.4.12 i Lay själva! (Fungerar även då A singulär, men med R singulär.)