Linjär Algebra Föreläsning 5

Förra vecka

- skalärprodukt: För $\vec{u}, \vec{v} \in \mathbb{R}^n$ är $\vec{u} \cdot \vec{v} = u_1 v_1 + \dots + u_n v_n$
- $\vec{u}, \vec{v}$ ortogonala om $\vec{u} \cdot \vec{v} = 0$
- $\{\vec{u}_1, \dots, \vec{u}_m\}$ ortogonal mängd om $\vec{u}_i \cdot \vec{u}_j = 0$ för alla $i \neq j$
  - om $\vec{u}_1, \dots, \vec{u}_m$ är nollskilda utgör de en bas till rummet de spänner. (ortogonalbas)
  - lätt att beräkna koordinater relativt en ortogonalbas:

$$\vec{v} = c_1 \vec{u}_1 + \dots + c_m \vec{u}_m \qquad \text{där} \quad c_j = \frac{\vec{v} \cdot \vec{u}_j}{|\vec{u}_j|^2} \qquad (1 \le j \le m)$$

- Om $W \subseteq \mathbb{R}^n$ är ett delrum med ortogonalbas $\vec{u}_1, \dots, \vec{u}_m$:
  $\vec{v} \in \mathbb{R}^n$ kan skrivas på ett unikt sätt som $\vec{v} = \hat{\vec{v}} + \vec{v}^{\perp}$, där

$$\hat{\vec{v}} = \sum_{i=1}^{m} \frac{\vec{u}_i \cdot \vec{v}}{|\vec{u}_i|^2} \vec{u}_i \in W, \qquad \vec{v}^{\perp} = \vec{v} - \hat{\vec{v}} \in W^{\perp}.$$

Idag Hur får man fram en ortogonalbas? – Användningar: Minsta kvadratlösningar

Betrakta $W = \text{span}\{\vec{u}_1, \vec{u}_2\}$, $\dim W = 2$. Vi ska finna en ortogonalbas $\{\vec{v}_1, \vec{v}_2\}$ till W.

Plan: Ta $\vec{v}_1 = \vec{u}_1$. Sen ersätta $\vec{u}_2$ med en vektor $\vec{v}_2 \in W$ som är ortogonal med $\vec{v}_1 = \vec{u}_1$.

Vi vet: $\vec{u}_2$ kan skrivas på ett unikt sätt som $\vec{u}_2 = \hat{\vec{u}}_2 + \vec{u}_2^{\perp}$ där $\hat{\vec{u}}_2 \in \text{span}\{\vec{u}_1\}$ och $\vec{u}_2^{\perp}$ är ortogonal med $\vec{v}_1$. Vidare vet vi att $\hat{\vec{u}}_2 = \frac{\vec{u}_2 \cdot \vec{v}_1}{|\vec{v}_1|^2} \vec{v}_1$.

Vi kan alltså sätta

$$\vec{v}_2 = \vec{u}_2^{\perp} = \vec{u}_2 - \hat{\vec{u}}_2 = \vec{u}_2 - \frac{\vec{u}_2 \cdot \vec{v}_1}{|\vec{v}_1|^2} \vec{v}_1 .$$

Då utgör $\{\vec{v}_1, \vec{v}_2\}$ en ortogonalbas till W.

Ex $\vec{u}_1 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$, $\vec{u}_2 = \begin{pmatrix} -2 \\ -1 \\ 1 \end{pmatrix}$, $W = \text{span}\{\vec{u}_1, \vec{u}_2\}$. Hitta ortogonalbas!

Svar Vi sätter $\vec{v}_1 = \vec{u}_1 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$ och

$$\vec{v}_2 = \vec{u}_2 - \frac{\vec{u}_2 \cdot \vec{v}_1}{|\vec{v}_1|^2} \vec{v}_1 = \begin{pmatrix} -2 \\ -1 \\ 1 \end{pmatrix} - \frac{(-2)\cdot 1 + (-1)\cdot 2 + 0}{1^2 + 2^2 + 0^2} \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} -2 \\ -1 \\ 1 \end{pmatrix} + \frac{4}{5} \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} -\overset{4}{\cancel{8}}/5 \\ 3\cancel{6}/5 \\ 1 \end{pmatrix}.$$

$\{\vec{v}_1, \vec{v}_2\}$ är då en ortogonalbas.

Strategin kan generaliseras:

Sats Låt $\{\vec{u}_1, \ldots, \vec{u}_m\}$ vara en bas till ett delrum $W \subseteq \mathbb{R}^n$, och definiera

$$\vec{v}_1 = \vec{u}_1$$

$$\vec{v}_2 = \vec{u}_2 - \frac{\vec{u}_2 \cdot \vec{v}_1}{|\vec{v}_1|^2}\vec{v}_1$$

$$\vec{v}_3 = \vec{u}_3 - \frac{\vec{u}_3 \cdot \vec{v}_1}{|\vec{v}_1|^2}\vec{v}_1 - \frac{\vec{u}_3 \cdot \vec{v}_2}{|\vec{v}_2|^2}\vec{v}_2$$

$$\vdots$$

$$\vec{v}_m = \vec{u}_m - \frac{\vec{u}_m \cdot \vec{v}_1}{|\vec{v}_1|^2}\vec{v}_1 - \ldots - \frac{\vec{u}_m \cdot \vec{v}_{m-1}}{|\vec{v}_{m-1}|^2}\vec{v}_{m-1}.$$

Då är $\{\vec{v}_1, \ldots, \vec{v}_m\}$ en ortogonalbas till W.

Bevisidé

Låt $W_p = \operatorname{span}\{\vec{u}_1, \ldots, \vec{u}_p\} \subseteq W$ för $1 \le m \le p$ vara delrummet som spanns av de första p basvektorerna.

I det ~~förs~~ k-te steget anta att $\{\vec{v}_1, \ldots \vec{v}_{k-1}\}$ är en ortogonalbas till W_{k-1}.

Eftersom $W_{k-1} \subseteq W_k$ är delrum kan vi skriva $\vec{u}_k = \hat{\vec{u}}_k + \vec{u}_k^{\perp}$, där

$$\hat{\vec{u}}_k = \frac{\vec{u}_k \cdot \vec{v}_1}{|\vec{v}_1|^2}\vec{v}_1 + \ldots + \frac{\vec{u}_k \cdot \vec{v}_{k-1}}{|\vec{v}_{k-1}|^2}\vec{v}_{k-1} \in W_{k-1}, \quad \vec{u}_k^{\perp} = \vec{u}_k - \hat{\vec{u}}_k.$$

Vektorn $\vec{v}_k := \vec{u}_k - \hat{\vec{u}}_k$ är alltså ortogonal mot $\vec{v}_1, \ldots, \vec{v}_{k-1}$, och $\{\vec{v}_1, \ldots, \vec{v}_k\}$ är en ortogonalbas till W_k.

Nu iterera: kolla på $k+1$, $k+2$, etc.

Ex Hitta en ortogonalbas till $\mathbb{R}^3$ som innehåller vektorn $\vec{v}_1 = \begin{pmatrix}1\\2\\3\end{pmatrix}$.

Svar $\vec{v}_1, \vec{e}_2, \vec{e}_3$ är linjärt oberoende och därför en bas till $\mathbb{R}^3$.

Vi tar

$$\vec{v}_2 = \vec{e}_2 - \frac{\vec{v}_1 \cdot \vec{e}_2}{|\vec{v}_1|^2}\vec{v}_1 = \begin{pmatrix}0\\1\\0\end{pmatrix} - \frac{2}{14}\begin{pmatrix}1\\2\\3\end{pmatrix} = \begin{pmatrix}0\\1\\0\end{pmatrix} - \frac{1}{7}\begin{pmatrix}1\\2\\3\end{pmatrix} = \frac{1}{7}\begin{pmatrix}-1\\5\\-3\end{pmatrix}$$

och

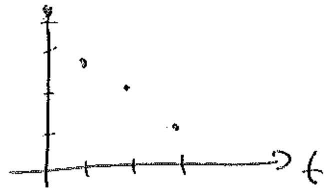
$$\vec{v}_3 = e_3 - \frac{\vec{v}_1 \cdot \vec{e}_3}{|\vec{v}_1|^2}\vec{v}_1 - \frac{\vec{v}_2 \cdot \vec{e}_3}{|\vec{v}_2|^2} = \begin{pmatrix}0\\0\\1\end{pmatrix} - \frac{3}{14}\begin{pmatrix}1\\2\\3\end{pmatrix} - \frac{-3/7}{\frac{1}{49}+\frac{25}{49}+\frac{9}{49}}\begin{pmatrix}-1\\5\\-3\end{pmatrix}$$

$$= \begin{pmatrix}0\\0\\1\end{pmatrix} - \frac{3}{14}\begin{pmatrix}1\\2\\3\end{pmatrix} + \frac{3}{5}\begin{pmatrix}-1\\5\\-3\end{pmatrix} = \frac{1}{70}\begin{pmatrix}-57\\180\\-101\end{pmatrix}$$

Minsta-kvadrat-metoden

Antag att man vid ett experiment får mätvärdena

t	0	1	2	3
b	4,1	2,8	2,1	1,0



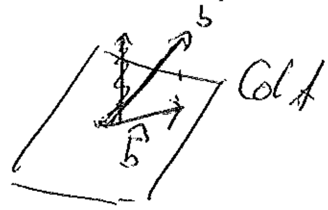
Vi vill hitta linjen $y = u + vt$ som bäst approximerar mätvärdena. Detta motsvarar systemet

$$A\begin{pmatrix} u \\ v \end{pmatrix} = \vec{b}, \quad A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} 4,1 \\ 2,8 \\ 2,1 \\ 1,0 \end{pmatrix}$$

Systemet har ingen lösning (punkterna ligger inte på en linje). Men hur hittar man bästa approximationen till en lösning?

Def Om A är en $m \times n$ matris och $\vec{b} \in \mathbb{R}^n$, så är $\vec{x} \in \mathbb{R}^n$ en minstakvadratlösning till systemet $A\vec{x} = \vec{b}$ om $|\vec{b} - A\vec{x}| \leq |\vec{b} - A\vec{y}|$ för alla $\vec{y} \in \mathbb{R}^n$.

Det följer från definitionen av kolonnrummet att, om $\vec{y}$ varierar, så tar $A\vec{y}$ alla värden i $\operatorname{Col} A$. Vi ska hitta den vektor $\hat{\vec{b}} \in \operatorname{Col} A$ som ligger närmast $\vec{b}$. Med andra ord så ska vi bestämma ortogonalprojektionen $\hat{\vec{b}}$ av $\vec{b}$ på $\operatorname{Col} A$.



Sats Minstakvadratlösningen till $A\vec{x} = \vec{b}$ är ~~den~~ alla vektorer i den icke-tomma mängden av lösningar till

$$A^T A\,\vec{x} = A^T\vec{b}.$$

Notis A är en $m \times n$ matris där m oftast är mycket större än n, vilket betyder att systemet $A\vec{x} = \vec{b}$ bestående av m ekvationer i n variabler vanligtvis har inga lösningar. Matrisen $A^T A$ har formatet $n \times n$ (en $n \times m$ gånger $m \times n$ matris), och systemet $A^T A\vec{x} = A^T\vec{b}$ är alltså lösbart (i de flesta fallen redan därför att $A^T A$ är kvadratiskt, och här speciellt enligt konstruktionen).

Bevisidé Vi skriver $\vec{b} = \hat{\vec{b}} + \vec{b}^\perp$ där $\hat{\vec{b}} \in \operatorname{Col} A$ och $\vec{b}^\perp \cdot \vec{a}_i = 0$ för alla kolonnvektorer $\vec{a}_i$ i A. Eftersom $\hat{\vec{b}} \in \operatorname{Col} A$ finns det $\vec{x} \in \mathbb{R}^n$ s.a. $A\vec{x} = \hat{\vec{b}}$. Vidare är $\vec{a}_i \cdot \vec{b}^\perp = \vec{a}_i^T\,\vec{b}^\perp = (a_{i1}, \dots, a_{in})\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ som matrisprodukt. Vi får alltså $0 = \vec{a}_i^T\vec{b}^\perp = \vec{a}_i^T(\vec{b} - \hat{\vec{b}}) = \vec{a}_i^T(\vec{b} - A\vec{x})$, och genom att kombinera radvektorerna $\vec{a}_i^T$, även $A^T(\vec{b} - A\vec{x}) = 0$.

Ex Hitta minstakvadratlösningen i inledande exemplet ovan!

Svar

$$A^TA = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 4 & 6 \\ 6 & 14 \end{pmatrix},$$

$$A^T\vec{b} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 4,1 \\ 2,8 \\ 2,1 \\ 1,0 \end{pmatrix} = \begin{pmatrix} 10 \\ 10 \end{pmatrix}$$

[redacted] Vi får alltså systemet $A^TA\vec{x} = A^T\vec{b}$, vilket blir

$$\begin{pmatrix} 4 & 6 \\ 6 & 14 \end{pmatrix} \vec{x} = \begin{pmatrix} 10 \\ 10 \end{pmatrix}.$$ [redacted]

Genom att betrakta det utökade systemet: $\begin{pmatrix} 4 & 6 & 10 \\ 6 & 14 & 10 \end{pmatrix} \sim \ldots \sim \begin{pmatrix} 1 & 0 & 4 \\ 0 & 1 & -1 \end{pmatrix}$,

alltså $x_1 = 4$ och $x_2 = -1$. Linjen som passar bäst är alltså $y = 4 - x$.

Obs Svaret är inte alltid entydigt!

Ex Bestäm minstakvadratlösningen till $A\vec{x} = \vec{b}$ där

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 3 \\ 3 & 0 & 3 \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix}.$$

Svar $A^TA = \ldots = \begin{pmatrix} 14 & 2 & 16 \\ 2 & 1 & 3 \\ 16 & 3 & 19 \end{pmatrix}$, $\vec{A}^T\vec{b} = \begin{pmatrix} 6 \\ -2 \\ 4 \end{pmatrix}$,

och vi får det utökade systemet

$$\left(A^TA \mid A^T\vec{b}\right) = \left(\begin{array}{ccc|c} 14 & 2 & 16 & 6 \\ 2 & 1 & 3 & -2 \\ 16 & 3 & 19 & 4 \end{array}\right) \sim \ldots \sim \left(\begin{array}{ccc|c} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & -4 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

fri variabel!

$\Rightarrow A^TA\vec{x} = A^T\vec{b}$ har den allmänna lösningen $\vec{x} = \begin{pmatrix} 1 \\ -4 \\ 0 \end{pmatrix} + x_3\begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}$,

vilket är då den generella minstakvadratlösningen till $A\vec{x} = \vec{b}$.

I vilka fall händer det?

Sats Låt A vara en $m \times n$ matris. De följande påståenden är ekvivalenta:

(a) $A\vec{x} = \vec{b}$ har en unik minstakvadratlösning för alla $\vec{b}$

(b) kolonnerna i A är linjärt oberoende

(c) A^TA är inverterbar.