75032 Matrisemetoder

Løsningsforslag til oppgaver gitt til onsdag 5/4–95

Påminnelse: En SVD $A = Q_1 \Sigma Q_2^{\mathrm{T}}$ kan man finne ved å merke seg at vi skal ha $A^{\mathrm{T}}A = Q_2\Sigma^{\mathrm{T}}\Sigma Q_2^{\mathrm{T}}$, så man finner egenverdiene (kvadratrøttene blir singulærverdiene) til $A^{\mathrm{T}}A$, og lar egenvektorene bli søyelene i Q_2. Deretter finner man de første søylene (like mange som det er singulærverdier $\neq 0$) i Q_1 ut fra $AQ_2 = Q_1\Sigma$, og resten ved å utvide til ortonormal basis.

Men man kan også skrive $AA^{\mathrm{T}} = Q_1\Sigma\Sigma^{\mathrm{T}}Q_1^{\mathrm{T}}$, så man diagonaliserer AA^{T} og utnytter tilslutt $A^{\mathrm{T}}Q_1 = Q_2\Sigma^{\mathrm{T}}$ for å finne de første søyelene i Q_2.

A.2 side 451: Raskest er å skrive $AA^{\mathrm{T}} = [4]$ som har egenverdien 4 med normalisert egenvektor $[1]$, slik at $\sigma_1 = \sqrt{4} = 2$, og vi får $Q_1 = [1]$ og $\Sigma = [2, 0, 0, 0]$. Første søyle i Q_2 må bli $2^{-1}A^{\mathrm{T}}[1] = (1/2)[1, 1, 1, 1]^{\mathrm{T}}$, og resten finner man ved å utvide til en ortonormal basis. For eksempel

$$Q_2 = \begin{bmatrix} 1/2 & 1/2 & 1/\sqrt{2} & 0 \\ 1/2 & 1/2 & -1/\sqrt{2} & 0 \\ 1/2 & -1/2 & 0 & 1/\sqrt{2} \\ 1/2 & -1/2 & 0 & -1/\sqrt{2} \end{bmatrix}$$

som jeg fant uten noe videre regning (men generelt må man dra igang hele maskineriet med Gram-Schmidt eller lignende).

BB^{T} blir en 2×2 identitetsmatrise, så singulærverdiene blir $\sigma_1 = \sigma_2 = 1$. Jeg fikk

$$B = Q_1\Sigma Q_2^{\mathrm{T}} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Det er jo søylene i Q_1 som velges først (egenvektorene for BB^{T}), og opprinnelig hadde jeg valgt dem i omvendt rekkefølge. Men så oppdaget jeg at om jeg gjorde som over ble $Q_1\Sigma = B$, og dermed var det bare å sette $Q_2 = I$ uten mer regning.

Jeg fikk

$$C^{\mathrm{T}}C = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad \sigma_1^2 = 2, \sigma_2^2 = 0, \quad \Sigma = \begin{bmatrix} \sqrt{2} & \\ & 0 \end{bmatrix}, \quad Q_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \quad Q_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

A.4 side 451: Vi diagonaliserer $A^{\mathrm{T}}A$:

$$A^{\mathrm{T}}A = \begin{bmatrix} 10 & 6 \\ 6 & 10 \end{bmatrix} = Q_2\Lambda Q_2^{\mathrm{T}}, \quad \Lambda = \begin{bmatrix} 16 & \\ & 4 \end{bmatrix}, \quad Q_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

Den søkte kvadratroten er

$$S = Q_2\Sigma Q_2^{\mathrm{T}} = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \qquad \text{der } \Sigma = \Lambda^{1/2} = \begin{bmatrix} 4 & \\ & 2 \end{bmatrix}.$$

(b) Vi kan nå finne polardekomposisjonen $A = QS$ ut fra SVD-en $A = Q_1\Sigma Q_2^{\mathrm{T}} = Q_1 Q_2^{\mathrm{T}} S$, men da må vi finne Q_1. Ikke at det er så vanskelig, men vi kan like gjerne sette

$$Q = AS^{-1} = \frac{1}{\sqrt{10}} \begin{bmatrix} 10 & 6 \\ 0 & 8 \end{bmatrix} \frac{1}{8} \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix} = \frac{1}{\sqrt{10}} \begin{bmatrix} 3 & 1 \\ -1 & 3 \end{bmatrix}.$$

Til slutt finner vi den "omvendte" polardekomposisjonen: $A = QS = QSQ^{\mathrm{T}}Q = S'Q$ der altså

$$S' = QSQ^{\mathrm{T}} = \frac{1}{10} \begin{bmatrix} 36 & 8 \\ 8 & 24 \end{bmatrix}.$$

A.6 side 452:

(a) Her har jeg ikke mindre enn tre løsninger! For det første, når A har l.u. søyler slik at $A^{\mathrm{T}}A$ er invertibel, har normalligningen $A^{\mathrm{T}}Ax = A^{\mathrm{T}}b$ entydig løsning som jo er gitt ved $x = (A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}b$, og dermed må den pseudoinverse være matrisen $(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}$.

Den andre løsningen er å bruke en SVD: $A = Q_1\Sigma Q_2^{\mathrm{T}}$ gir

$$A^{\mathrm{T}}A = Q_2\Sigma^{\mathrm{T}}\Sigma Q_2^{\mathrm{T}}, \quad \Sigma^{\mathrm{T}}\Sigma = \begin{bmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_n^2 \end{bmatrix}$$

der antagelsene forteller oss at A har n singulærverdier $\neq 0$, så $\Sigma^{\mathrm{T}}\Sigma$ er invertibel, og

$$(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}} = Q_2(\Sigma^{\mathrm{T}}\Sigma)^{-1}Q_2^{\mathrm{T}}Q_2\Sigma^{\mathrm{T}}Q_1^{\mathrm{T}} = Q_2\underbrace{(\Sigma^{\mathrm{T}}\Sigma)^{-1}\Sigma^{\mathrm{T}}}_{\Sigma^+}Q_1^{\mathrm{T}} = A^+.$$

Den tredje løsningen består i å forstå hvordan A^+ virker på $\mathcal{R}(A)$ og på $\mathcal{N}(A^{\mathrm{T}})$:

1. Om $b \in \mathcal{N}(A^{\mathrm{T}})$ er $A^+b = 0$,
2. Om $b \in \mathcal{R}(A)$ kan vi skrive $b = Ax$ for en $x \in \mathcal{N}(A)^{\perp} = \mathcal{R}(A^{\mathrm{T}})$, og da er $A^+b = x$.

og siden $\mathcal{R}(A)$ og $\mathcal{N}(A^{\mathrm{T}})$ tilsammen utspenner $\mathbf{R}^m$ vil enhver matrise som oppfyller de samme kravene være lik A^+. At $(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}$ oppfyller betingelsene er en triviell utregning.

(b) Den første løsningen fra (a) har kanskje ingen enkel versjon i dette tilfellet, men den andre løsningen kan enkelt tilpasses. Jeg viser ikke det her, men tar den tredje varianten.

Når A har l.u. rekker så er $\mathcal{N}(A^{\mathrm{T}}) = \{0\}$, så om $b \in \mathcal{N}(A^{\mathrm{T}})$ er trivielt $A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}b = 0$.

Og om $b \in \mathcal{R}(A)$ er skrevet som $b = Ax$ for en $x \in \mathcal{R}(A^{\mathrm{T}})$ kan vi videre skrive $x = A^{\mathrm{T}}c$, så $A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}b = A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}AA^{\mathrm{T}}c = A^{\mathrm{T}}c = x$.

A.10 side 452: En geometrisk måte å tolke det på er slik: For en $m \times n$-matrise A splitter vi $\mathbf{R}^n$ opp i ortogonale komponenter $\mathcal{R}(A^{\mathrm{T}})$ og $\mathcal{N}(A)$, og likedan $\mathbf{R}^m$ i ortogonale komponenter $\mathcal{R}(A)$ og $\mathcal{N}(A^{\mathrm{T}})$. A avbilder $\mathcal{R}(A^{\mathrm{T}})$ en-til-en på $\mathcal{R}(A)$ med invers A^+, mens A og A^+ er null på henhodsvis $\mathcal{N}(A)$ og $\mathcal{N}(A^{\mathrm{T}})$.

Dermed er $AA^+x = 0$ når $x \in \mathcal{N}(A^{\mathrm{T}})$ og $AA^+x = x$ når $x \in \mathcal{R}(A)$, så AA^+ er ortogonalprojeksjonen på $\mathcal{R}(A)$.

Likedan er $A^+Ax = 0$ når $x \in \mathcal{N}(A)$ og $A^+Ax = x$ når $x \in \mathcal{R}(A^{\mathrm{T}})$, så A^+A er ortogonalprojeksjonen på $\mathcal{R}(A^{\mathrm{T}})$.

Alternativt kan man bruke SVD:

$$AA^+ = Q_1\Sigma Q_2Q_2^{\mathrm{T}}\Sigma^+Q_1^{\mathrm{T}} = Q_1\Sigma\Sigma^+Q_1^{\mathrm{T}} = Q_1\begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & 0 & \\ & & & & \ddots \end{bmatrix}Q_1^{\mathrm{T}}$$

og matrisen i midten (med r enere på diagonalen) er en prosjeksjon, dermed blir AA^+ det. Vi ser også at $\mathcal{R}(AA^+)$ blir utspent av de første r søylene i Q_1, som vi allerede vet utspenner $\mathcal{R}(A)$. A^+A kan behandles på samme måten.

Eksamen mai 1989: 1: *Denne oppgaven skulle ikke vært gitt. Jeg skrev oppgave 4 på tavla, men 1 i WWW-siden. I tilfelle noen regnet oppgave 1 så gir jeg en løsning her.*

(a) Besvarelsen i oppgavesamlingen pivoterer litt, men det har vel ikke så mye hensikt her. Uten pivotering får jeg

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

(b) $\mathcal{N}(A) = \mathcal{N}(U)$ er utspent av $[0, -2, 1]^{\mathrm{T}}$. Basis for $\mathcal{R}(A)$ er første og annen søyle i A eller i L (siden pivotene i U finnes i de to søylene). Men vi ser at nullvektoren jeg fant ovenfor er -2 ganger annen søyle i L, så tar vi den vektoren og første søyle i L har vi en basis for $\mathcal{R}(A)$.

(c) $\mathcal{N}(B) \subseteq \mathcal{R}(B)$ gir oss $3 - r \leq$ for rangen r til B. Dermed må $r \geq 2$. Enhver 3×3-matrise B med rang 3 er jo invertibel, så $\dim \mathcal{N}(B) = 0$. Og vi har nettopp sett et eksempel på at rangen kan være 2.

Eksamen mai 1989: 4:

(a) Singulærverdidekomposisjonen for A_k er jo $A_k = Q_1 \Sigma_k Q_2^{\mathrm{T}}$, så rangen er k fordi vi har k singulærverdier på diagonalen i Σ_k. Mer detaljert: Σ_k har rang k fordi den er på trappeform med k pivoter. Multiplikasjon med invertible matriser på den ene eller andre siden endrer ikke rangen.

Vi følger hintet, og merker oss at $(\Sigma - \Sigma_k)x = [0, \ldots, 0, \sigma_{k+1} x_{k+1}, \ldots, \sigma_r x_r, 0, \ldots, 0]^{\mathrm{T}}$ slik at

$$\|(\Sigma - \Sigma_k)x\|^2 = \sum_{j=k+1}^{r} \sigma_j^2 x_j^2 \leq \sum_{j=k+1}^{r} \sigma_{k+1}^2 x_j^2 \leq \sigma_{k+1}^2 \sum_{j=1}^{n} x_j^2 = \sigma_{k+1}^2 \|x\|^2 .$$

Der den første ulikheten skyldes at singulærverdiene er sortert. Dette benytter vi så slik:

$$\|(A - A_k)x\| = \left\| Q_1 (\Sigma - \Sigma_k) Q_2^{\mathrm{T}} x \right\| = \left\| (\Sigma - \Sigma_k) Q_2^{\mathrm{T}} x \right\| \leq \sigma_{k+1} \left\| Q_2^{\mathrm{T}} x \right\| = \sigma_{k+1} \|x\| .$$

(b) La $u_1, \ldots, u_k$ være en basis for U og $v_1, \ldots, v_\ell$ en basis for V. Da er alle disse vektorene lineært uavhengige, for om

$$c_1 u_1 + \cdots + c_k u_k + d_1 v_1 + \cdots + d_\ell v_\ell = 0$$

er

$$c_1 u_1 + \cdots + c_k u_k = -(d_1 v_1 + \cdots + d_\ell v_\ell) \in U \cap V = \{0\}$$

fordi venstresiden i likheten tilhører U og høyresiden tilhører V. Vi har altså $c_1 u_1 + \cdots + c_k u_k = 0$ så $c_1 = \cdots = c_k = 0$ fordi *u*ene er lineært uavhengige. På samme måte finner vi $d_1 = \cdots = d_\ell = 0$.

Vi har nå $k + \ell$ lineært uavhengige vektorer i $\mathbf{R}^n$, og da er $k + \ell \leq n$ ved plassmangelteoremet.

(c) Vi tar hintet igjen, og betrakter $x = [x_1, \ldots, x_{k+1}, 0, \ldots, 0]^{\mathrm{T}} \neq 0$. Hintet anbefaler oss å vise at $BQ_2 x \neq 0$, så vi antar det motsatte, altså $BQ_2 x = 0$, og forsøker å utlede en motsigelse. At fra antagelsen har vi $\|(A - B)Q_2 x\| \leq C \|Q_2 x\|$, som i vårt tilfelle blir

$$\|AQ_2 x\| \leq C \|x\| \tag{1}$$

På den annen side blir

$$\begin{aligned} \|AQ_2 x\| = \|Q_1 \Sigma x\| = \|\Sigma x\| &= \left\| [\sigma_1 x_1, \ldots, \sigma_{k+1} x_{k+1}, 0, \ldots, 0]^{\mathrm{T}} \right\| \\ &= \sqrt{\sum_{j=1}^{k+1} \sigma_j^2 x_j^2} \geq \sqrt{\sum_{j=1}^{k+1} \sigma_{k+1}^2 x_j^2} = \sigma_{k+1} \|x\| \end{aligned}$$

som sammen med (1) gir $C \geq \sigma_{k+1}$ i motstrid med antagelsen. Altså er $BQ_2 x \neq 0$.

Hva skal vi så bruke dette til? Vi skal vise at rangen til B er stor, mens vi formodentlig skal bruke resultatet fra (b) som bare kan brukes til å vise at dimensjonen til et underrom er lite. Men den ønskede ulikhet er ekvivalent med $n - \dim \mathcal{N}(B) > k$, altså $k + 1 + \dim \mathcal{N}(B) \leq n$. Men ovenfor fant vi jo nettopp et rom av dimensjon $k + 1$, nemlig alle vektorer på formen $Q_2[x_1, \ldots, x_{k+1}, 0, \ldots, 0]^{\mathrm{T}}$, som har trivielt snitt med $\mathcal{N}(B)$. Dermed vil resultatet fra (b) gi oss svaret vi trenger.

(Dette punktet var det nesten ingen som klarte på eksamen. Jeg skal forsøke å ikke ha et like vrient punkt på årets eksamen, men en liten "nøtt" synes jeg må til.)