Singulärvärden och tillämpningar

$A \quad m \times n$ - matris $\Rightarrow$ $B = A^t A \quad n \times n$ - matris

$B^t = (A^t A)^t = A^t (A^t)^t = A^t A = B \Rightarrow B$ symmetrisk

Spektralsatsen $\Rightarrow$ finns ON-bas $v_1, \dots, v_n$ för $\mathbb{R}^n$ av egenvektorer till B.
Vi tänker på vektorerna som (koordinat-)kolonner så $Bv_j = \lambda_j v_j$.

Sätt $w_j = Av_j \in \mathbb{R}^m$. $v_j^t v_j = |v_j|^2 = 1 \Rightarrow \lambda_j = \lambda_j v_j^t v_j = v_j^t(\lambda_j v_j) =$
$v_j^t(Bv_j) = v_j^t(A^t A v_j) = (v_j^t A^t)(Av_j) = (Av_j)^t(Av_j) = w_j^t w_j = |w_j|^2 \geq 0$

Alla egenvärden $\lambda_j \geq 0$ alltså. Ordna dem så att $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$
och sätt $\sigma_j = \sqrt{\lambda_j}$.

$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$ kallas A:s singulärvärden.

$(w_j|w_k) = w_j^t w_k = (Av_j)^t(Av_k) = (v_j^t A^t)(Av_k) = v_j^t(A^t A v_k) = v_j^t B v_k =$
$v_j^t \lambda_k v_k = \lambda_k v_j^t v_k = \lambda_k (v_j|v_k) = \lambda_k \cdot 0 = 0$ om $j \neq k$.

$w_1, \dots, w_n$ är ortogonala och $|w_j|^2 = (w_j|w_j) = \lambda_j = \sigma_j^2$.

Normera $u_j = w_j/|w_j| = w_j/\sigma_j \Rightarrow u_1, \dots, u_n$ ON-mängd i $\mathbb{R}^m$ om $m \geq n$.
Utvidga vid behov med fler u_j till ON-bas för $\mathbb{R}^m$ (görs också om några $\sigma_j = 0$ så att u_j inte kan definieras från w_j).

Låt U och V vara matriserna med u_j resp. v_j som kolonner. U är $m \times m$ och V är $n \times n$. Båda är ortogonala: $U^{-1} = U^t, V^{-1} = V^t$

Vi har $Av_j = \sigma_j u_j$ vilket kan skrivas $AV = U\Sigma$ där Σ är $m \times n$ - matrisen

$$\Sigma = \begin{pmatrix} \sigma_1 & 0 & 0 & \dots \\ 0 & \sigma_2 & 0 & \dots \\ 0 & 0 & \sigma_3 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

$V^{-1} = V^t$ ger nu $A = U\Sigma V^t$.

Om $n > m$ kommer kan man göra ett liknande resonemang, t.ex. för A^t.

Slutsats: Singulärvärdesfaktorisering/uppdelning av en matris (SVD). Varje matris A kan skrivas $A = U\Sigma V^t$. Om A är $m \times n$ är U $m \times m$, V $n \times n$ och Σ $m \times n$, U och V ortogonala samt Σ "diagonal" med A:s singulärvärden på diagonalen.

Man kan visa att $\text{rang}(A) = \text{rang}(A^t A)$. Om $r = \text{rang}(A)$ betyder detta att $\lambda_j = 0$ och $\sigma_j = 0$ för $j > r$.

$A = U\Sigma V^t$ kan då skrivas (jämför spektraluppdelning)

$A = \sigma_1 u_1 v_1^t + \cdots + \sigma_r u_r v_r^t$

Varje term i summan är en $m \times n$ - matris av rang 1.

$A = U\Sigma V^t = \sigma_1 u_1 v_1^t + \cdots + \sigma_r u_r v_r^t$

SVD används ofta p.g.a. följande sats:

Låt A_s vara bästa rang-s approximationen av A (d.v.s. A_s har rang $\leq s$ och minimerar $|A - A_s|$, som definieras som roten av summan av kvadraterna på alla element i $A - A_s$). Då är

$A_s = \sigma_1 u_1 v_1^t + \cdots + \sigma_s u_s v_s^t$ (de s första termerna i SVD för A).

En 1000×1000 - matris har 10^6 element. En rang-20 approximation behöver bara $u_1, \ldots, u_{20}, v_1, \ldots, v_{20}$, d.v.s. 40.000 element, en stor minnesbesparing! Att approximera en matris med en annan av lägre rang är en standardteknik inom många ingenjörsområden, t.ex. signalbehandling och bildkompression. Det kan användas för att filtrera bort brus. En annan användare är Netflix vars algoritm när de rekommenderar filmer till kunder bygger på SVD.

Man behöver bara räkna ut de s egenvektorer som ska användas i approximationen, inte alla n egenvektorerna.

SVD kan också bidra med geometriska tolkningar av den linjära avbildning som en matris representerar.

Exempel:

$$A = \begin{pmatrix} 3 & 1 & -1 \\ 3 & 3 & 3 \\ -1 & 1 & 1 \\ 2 & 1 & 1 \\ -3 & 1 & -1 \end{pmatrix} \Rightarrow A^t A = \begin{pmatrix} 32 & 10 & 10 \\ 10 & 13 & 9 \\ 10 & 9 & 13 \end{pmatrix}$$

Egenvärden till $A^t A$ är $\lambda_1 = 42$, $\lambda_2 = 12$, $\lambda_3 = 4$ med egenvektorer

$$v_1 = \frac{1}{\sqrt{6}} \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix},\ v_2 = \frac{1}{\sqrt{3}} \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} \text{ resp. } v_3 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \quad \text{(bildar ON-bas för } \mathbb{R}^3\text{).}$$

Singulärvärden till A är $\sigma_1 = \sqrt{42}$, $\sigma_2 = \sqrt{12}$, $\sigma_3 = 2$ och

$$Av_1 = \sqrt{42}\frac{1}{\sqrt{7}} \begin{pmatrix} 1 \\ 2 \\ 0 \\ 1 \\ -1 \end{pmatrix} = \sigma_1 u_1,\ Av_2 = \sqrt{12}\frac{1}{2} \begin{pmatrix} -1 \\ 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} = \sigma_2 u_2,\ Av_3 = 2\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \sigma_3 u_3$$

u_1, u_2, u_3 kan utvidgas med u_4, u_5 till ON-bas för $\mathbb{R}^5$ men behövs ej här.

SVD av A: $A = \begin{pmatrix} 3 & 1 & -1 \\ 3 & 3 & 3 \\ -1 & 1 & 1 \\ 2 & 1 & 1 \\ -3 & 1 & -1 \end{pmatrix} = U\Sigma V^t = \sigma_1 u_1 v_1^t + \sigma_2 u_2 v_2^t + \sigma_3 u_3 v_3^t =$

$$\begin{pmatrix} \frac{1}{\sqrt{7}} & \frac{-1}{2} & \frac{1}{\sqrt{2}} & | & | \\ \frac{2}{\sqrt{7}} & \frac{1}{2} & 0 & | & | \\ 0 & \frac{1}{2} & 0 & u_4 & u_5 \\ \frac{1}{\sqrt{7}} & 0 & 0 & | & | \\ \frac{-1}{\sqrt{7}} & \frac{1}{2} & \frac{1}{\sqrt{2}} & | & | \end{pmatrix} \begin{pmatrix} \sqrt{42} & 0 & 0 \\ 0 & \sqrt{12} & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ \frac{-1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} =$$

$$\sqrt{42} \begin{pmatrix} \frac{1}{\sqrt{7}} \\ \frac{2}{\sqrt{7}} \\ 0 \\ \frac{1}{\sqrt{7}} \\ \frac{-1}{\sqrt{7}} \end{pmatrix} \begin{pmatrix} \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{pmatrix} + \sqrt{12} \begin{pmatrix} \frac{-1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ 0 \\ \frac{1}{2} \end{pmatrix} \begin{pmatrix} \frac{-1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{pmatrix} + 2 \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ 0 \\ 0 \\ \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 0 & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

Bästa rang-1 och rang-2 approximation av A:

$$A_1 = \sigma_1 u_1 v_1^t = \begin{pmatrix} 2 & 1 & 1 \\ 4 & 2 & 2 \\ 0 & 0 & 0 \\ 2 & 1 & 1 \\ -2 & -1 & -1 \end{pmatrix}, \quad A_2 = \sigma_1 u_1 v_1^t + \sigma_2 u_2 v_2^t = \begin{pmatrix} 3 & 0 & 0 \\ 3 & 3 & 3 \\ -1 & 1 & 1 \\ 2 & 1 & 1 \\ -3 & 0 & 0 \end{pmatrix}$$