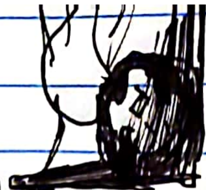
Teorema (Eckart-Young)



A melhor aproximação de posto-r da matriz no contexto de mínimos quadrados, é dada pela decomposição SVD de posto-r e denotada por A_r, ou seja

$$\underset{A_r,\ \text{sujeito}\ \text{posto}(A_r)=r}{\arg\min} \|A - A_r\|_F = \hat{U}\hat{\Sigma}\hat{V}^* \quad \text{ou} \quad U_r \Sigma_r V_r^*$$

Neste caso, $\hat{U}$ e $\hat{V}$ denotam as primeiras colunas principais de U e V, enquanto que $\hat{\Sigma}$ contém a submatriz $r \times r$ principal de Σ

Resultado importante: quantificação do erro da aproximação

$$\|A - A_r\|_F^2 = \sum_{k=r+1}^{n} \sigma_k^2$$

Obs. $A \in \mathbb{C}^{m \times n}$, $A = \begin{bmatrix} | & | & & | \\ a_1 & a_2 & \cdots & a_n \\ | & | & & | \end{bmatrix}$

Em geral, adota-se a normalização do erro

$$\|A - A_r\|_F^2 \big/ \|A\|_F^2$$

$$\|A\|_F = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2} \quad \text{(Frobenius)}$$

Considerando a norma-2, norma espectral, podemos reescrever o teorema como:

$$\underset{\substack{A_r,\ \text{sujeito} \\ \cancel{A_r}\ \text{posto}(A)=r}}{\arg\min} \|A - A_r\|_2 = \hat{U}\hat{\Sigma}^{-1}\hat{V}^*$$

<u>Obs</u>. $$\|A\|_2 = \max_{v \neq 0} \frac{\|Av\|_2}{\|v\|_2}$$

$$A - A_r = \sum_{k=r+1}^{n} \sigma_k u_k v_k^*$$

$$(A - A_r)v = \left(\sum_{k=r+1}^{n} \sigma_k u_k v_k^*\right)v$$, com v_k vetores ortonormais

O máximo de

$$\|(A - A_r)v\|_2 = \left\|\left(\sum \sigma_k u_k v_k^*\right)v\right\|_2 = \sigma_{r+1}$$

quando $v = v_{r+1}$. $\boxed{\text{Ou seja: } \|A - A_r\|_2 = \sigma_{r+1}}$

<u>Obs</u>. $A_k := u_k v_k^*$ (produto externo ou diádico)

Portanto, podemos interpretar a aproximação de posto r obtida pela decomposição SVD como <u>a projeção</u> de uma matriz de posto completo A em um espaço de <u>dimensão reduzida</u> de matrizes com posto no máximo r.

Teorema. A norma espectral de A é o maior valor singular da matriz A,

$$\|A\|_2 = \sigma_1$$

Esse teorema justifica

$$\|A - A_r\|_2 = \sigma_{r+1}$$

Propriedades:

$$A = U \Sigma V^* = [\hat{U} \quad \hat{U}^\perp] \begin{bmatrix} \hat{\Sigma} \\ 0 \end{bmatrix} V^* = \hat{U} \hat{\Sigma} V^*$$

As colunas de $\hat{U}^\perp$ geram um espaço vetorial que é complementar e ortogonal ao espaço gerado por $\hat{U}$.

Ao considerarmos as matrizes de correlação AA^* e A^*A podemos descrever algumas informações:

$$AA^* = U \begin{bmatrix} \hat{\Sigma} \\ 0 \end{bmatrix} V^* V [\hat{\Sigma} \quad 0] U^* = U \begin{bmatrix} \hat{\Sigma}^2 & 0 \\ 0 & 0 \end{bmatrix} U^*$$

$$A^*A = V [\hat{\Sigma} \quad 0] U^* U \begin{bmatrix} \hat{\Sigma} \\ 0 \end{bmatrix} V^* = V \hat{\Sigma}^2 V^*$$

Utilizando o fato que U e V são unitárias, podemos considerar os seguintes problemas de autovalores

$$AA^*U = U \begin{bmatrix} \hat{\Sigma}^2 & 0 \\ 0 & 0 \end{bmatrix} \quad e \quad A^*AV = V \hat{\Sigma}^2$$

Digitalizado com CamScanner

As colunas de U são autovetores da matriz AA^*, e as coluna de V são os autovetores de A^*A

Como os valores singulares são ordenados de forma decrescente em ordem de magnitude ($\sigma_1 \geq \sigma_2 \geq \cdots$) então as colunas de U são ordenadas hierarquicamente de quanto correlação elas capturam nas colunas de A.
Note que semelhantemente V captura correlação nas linhas de A.

Método de captura instantânea

$$\begin{bmatrix} \\ A \\ \\ \end{bmatrix} \quad \begin{bmatrix} & & \\ & AA^* & \\ & & \end{bmatrix} \quad \begin{bmatrix} A^*A \end{bmatrix}$$

Ao resolvermos o problema de autovalores:

$$A^*AV = V\hat{\Sigma}^2$$

obtemos V e $\hat{\Sigma}$. Caso haja zeros em $\hat{\Sigma}$ então podemos apenas considerar os valores não nulos (os primeiros r valores), $\hat{\Sigma}_r$ e as colunas correspondentes de V (ou seja V_r)

Com isso obtemos

$$U_r = A V_r \hat{\Sigma}_r^{-1}$$

Obs.

As colunas da matriz U formam uma base ortonormal para o espaço coluna de A.

As colunas da matriz V formam uma base ortonormal para o espaço linha de A.

Se as colunas de A são medidas espaciais no tempo, então U incorpora (encondes) os padrões espaciais e V os padrões temporais.

SVD e transformações Unitárias

Considere uma transformação unitária à esquerda de A, ou seja

$$Y = CA$$

Calculando a matriz de correlação

$$Y^* Y = A^* C^* C A = A^* A$$

e portanto os dados projetados tem a mesma autodecomposição, resultando que obtemos V_A e Σ_A.

Notação: Dada uma transf. unitária à esquerda a decomposição SVD de A será denotada $U_A \Sigma_A V_A^*$

Dada uma transf. unitária à direita por P^*, $Y = AP^*$, então a decomposição SVD de Y será denotada $U \Sigma V^*$.

Usando o método da captura instantânea para reconstruir U_Y, encontramos:

$$U_Y = Y V_A \Sigma_A^{-1} = C A V_A \Sigma_A^{-1} = C U_A$$

Ou seja,

$$\left.\begin{aligned} U_Y &= C U_A \\ \Sigma_Y &= \Sigma_A \\ V_Y &= V_A \end{aligned}\right\} \quad Y = CA = C U_A \Sigma_A V_A^*$$

Agora, vamos considerar uma transf. unitária à direita, $Y = AP^*$ cuja matriz de correlação

$$Y^* Y = P A^* A P^* = P V_A \Sigma_A^2 V_A^* P^*$$

ou seja

$$Y^* Y P V_A = P V_A \Sigma_A^2$$

Assim,

$$V_Y = P V_A$$

$$\Sigma_Y = \Sigma_A$$

Podemos usar o método da captura para reconstruir U_Y

$$U_Y = Y P V_A \Sigma_A^{-1} = A V_A \Sigma_A^{-1} = U_A$$

Digitalizado com CamScanner

Portanto $U_Y = U_A$, e a decomposição SVD de Y

$$Y = AP^* = U_A \Sigma_A V_A^* P^*$$

Aplicação de SVD : matriz pseudo-inversa (Moore-Penrose)

$$B = \tilde{U}\tilde{\Sigma}\tilde{V}^*$$

$$B^+ \equiv \tilde{V}\tilde{\Sigma}^{-1}\tilde{U}^* \Rightarrow B^+ B = \tilde{V}\tilde{V}^*$$

Vamos considerar o sistema

$$Ax = b \quad \begin{cases} \text{indeterminado se } A \in \mathbb{C}^{m \times n},\ m \ll n \\ \text{sobredeterminado se } m \gg n \end{cases}$$

No caso de sobredeterminado, quando não existe solução do sist., buscamos a solução x que minimiza a quantidade

$$\|b - Ax\|_2^2$$

Esse é a solução de mínimos quadrados (quadrados mínimos)

Já no caso de sist. indeterminado, quando infinitas soluções existem, podemos determinar a solução x que minimiza a quantidade

$$\|x\|_2 \text{ tal que } Ax = b$$

(solução norma mínima)

Vamos utilizar a pseudo inversa de A, da seguinte forma

$$A^{\dagger} A \tilde{x} = A^{\dagger} b \Rightarrow \tilde{x} = \tilde{V}\tilde{\Sigma}^{-1}\tilde{U}^{*} b$$

independentemente se o sistema original é (over ou under) "sobre" ou indeterminado. temos que:

$$A\tilde{x} = \tilde{U}\tilde{\Sigma}\tilde{V}^{*}\tilde{V}\tilde{\Sigma}^{-1}\tilde{U}^{*} b = \tilde{U}\tilde{U}^{*} b$$

Note: Para decomposição exata SVD, $\tilde{U}U^{*} = U^{*}\tilde{U} = I \in \mathbb{R}^{m\times m}$
Isso não é válido para uma versão aproximada ou truncada do SVD.

$\tilde{x}$ é solução exata de $Ax = b$ quando b está no espaço coluna de $\tilde{U}$, ou seja no espaço coluna de A.

Portanto, <u>muito cuidado</u> para não assumir ~~que~~ $\tilde{U}\tilde{U}^{*}$ como a matriz identidade, o que pode ser assumido é $\tilde{U}\tilde{U}^{*} = I \in \mathbb{R}^{r\times r}$, r o posto de A.