CALCUL NUMERIC
Valori Singulare

Paul Irofti
Cristian Rusu
Andrei Pătraşcu

Departament Informatică
Universitatea din Bucureşti

Cuprins

- **Valori singulare**
- Descompunerea Valorilor Singulare
- Rotaţie subspaţii
- Calcul DVS

Teoremă

Fie matricea $A \in \mathbb{R}^{m\times n}$ cu rangul $r \in [0, \min(m, n)]$, atunci există matricile ortonormale $U \in \mathbb{R}^{m\times m}$, $V \in \mathbb{R}^{n\times n}$ astfel încât:

$$U^T AV = \Sigma = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix}$$

unde

$$\Sigma_1 = \mathrm{diag}\left(\sigma_1, \cdots, \sigma_r\right), \quad \sigma_1 \geq \cdots \geq \sigma_r > 0.$$

Expresia $A = U\Sigma V^T$ se numeşte Descompunerea Valorilor Singulare a matricei A.

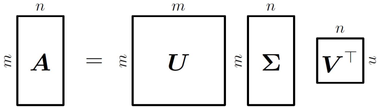
- σ_i se numesc valori singulare
- $U = \begin{bmatrix} u_1 & \cdots & u_m \end{bmatrix}$ se numesc vectori singulari la stânga
- $V = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}$ se numesc vectori singulari la dreapta

Nr. de valori singulare pozitive ale unei matrice este egal cu rangul acesteia

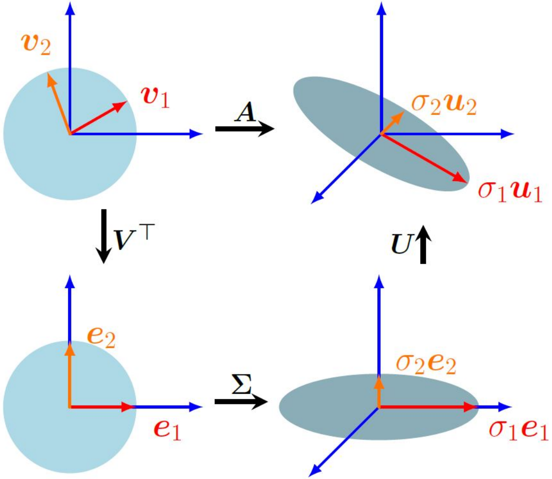
Exemplu: Fie matricea $A = \begin{bmatrix} 1 & -0.8 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$ cu DVS:

$$\begin{bmatrix} 1 & -0.8 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} = \underbrace{\begin{bmatrix} -0.79 & 0 & -0.62 \\ 0.38 & -0.78 & -0.49 \\ -0.48 & -0.62 & -0.62 \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} 1.62 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}}_{\Sigma} \underbrace{\begin{bmatrix} -0.78 & 0.62 \\ -0.62 & -0.78 \end{bmatrix}}_{V^T}$$

$\sigma_1 = 1.62, \sigma_2 = 1$, deci A are rang 2 (echivalent, cele 2 coloane sunt liniar independente)



- Σ is unique and has the same form as A
- if $m < n$ then $\Sigma = \begin{bmatrix} \sigma_1 & 0 \cdots & 0 & 0 \cdots & 0 \\ 0 & \sigma_2 \cdots & 0 & 0 \cdots & 0 \\ \vdots & \vdots & \vdots & 0 \cdots & 0 \\ 0 & 0 \cdots & \sigma_m & 0 \cdots & 0 \end{bmatrix}$
- if $m > n$ then $\Sigma = \begin{bmatrix} \sigma_1 & 0 \cdots & 0 \\ 0 & \sigma_2 \cdots & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 \cdots & \sigma_m \\ 0 & \cdots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \cdots & 0 \end{bmatrix}$



v_2
v_1
A
$\sigma_2 u_2$
$\sigma_1 u_1$
$V^\top$
U
e_2
e_1
Σ
$\sigma_2 e_2$
$\sigma_1 e_1$

Proprietăţi ale DVS:

- $Av_i = \sigma_i u_i$ şi $A^T u_i = \sigma_i v_i$ pentru $i = 1 : n$
- $\sigma_{\min} \|x\|_2 \leq \|Ax\|_2 \leq \sigma_{\max} \|x\|_2$
- Forma redusă: $A = \sum_{i=1}^{r} \sigma_i u_i v_i^T = U \Sigma V^T$

Problema Procrust Ortogonală: Fie matricile $A, B \in \mathbb{R}^{m \times n}$, este posibil ca matricea A să fie obţinută prin rotaţia lui B?

$$\min_{Q} \ \|A - BQ\|_F$$
$$\text{s.l.} \ \ Q^T Q = I_n$$

- Se caută matricea ortogonală optimă Q^* care minimizează distanţa de la A la transformarea BQ^*
- Soluţia Q^* se calculează folosind descompunerea DVS!

$$
\begin{aligned}
\|A - BQ\|_F^2 &= \sum_{i=1}^{n} \|A_i - BQ_i\|_2^2 \\
&= \sum_{i=1}^{n} \|A_i\|_2^2 + \|BQ_i\|_2^2 - 2\langle A_i, BQ_i\rangle \\
&= \|A\|_2^2 + \|BQ\|_2^2 - 2\sum_{i=1}^{n} [A^T BQ]_{ii} \\
&= \|A\|_2^2 + \|B\|_2^2 - 2Tr(A^T BQ)
\end{aligned}
$$

Primii doi termeni $\|A\|_2^2$ şi $\|B\|_2^2$ sunt constante în raport cu Q

Problema Procust originală se reduce la:

$$\max_{Q} \; Tr(Q^T B^T A)$$
$$\text{s.l.} \; Q^T Q = I_n$$

DVS: $U^T(B^T A)V = \Sigma$
Notăm: $Z := V^T Q^T U$

$$Tr(Q^T B^T A) = Tr(Q^T U \Sigma V^T) = Tr(V^T Q^T U \Sigma)$$
$$= Tr(Z\Sigma) = \sum_{i=1}^{n} Z_{ii}\sigma_i \leq \sum_{i=1}^{n} \sigma_i$$

Soluţia presupune atingerea marginii superioare $\sum_{i=1}^{n} \sigma_i$ prin găsirea unui Q astfel încât $Z = V^T Q^T U = I_n$

Observăm că alegând $Q = UV^T$ rezultă $Z = V^T(UV^T)^T U = V^T V U^T U = I_n$

Problema Procust originală se reduce la:

$$\max_{Q} \quad Tr(QB^TA)$$
$$\text{s.l.} \quad Q^TQ = I_n$$

Algorithm 1: Procust

Data: $A, B \in \mathbb{R}^{m\times n}$

1 $C = B^TA$

2 Calculează DVS: $U^TCV = \Sigma$

3 $Q = UV^T$

Se observă uşor din relaţiile specifice vectorilor singulari:

$$\begin{cases} Av_i = \sigma_i u_i \\ A^T u_i = \sigma_i v_i \end{cases}$$

prin înmulţirea de ambele părţi cu A^T (respectiv, a doua relaţie, cu A) se obţine

$$\begin{cases} A^T A v_i = \sigma_i^2 v_i \\ A A^T u_i = \sigma_i^2 u_i \end{cases}$$

ce arată că v_i sunt vectori proprii ai gramianului $A^T A$, iar u_i ai lui AA^T.

Pentru orice $A \in \mathbb{R}^{m \times n}$, gramianul $A^T A \in \mathbb{R}^{n \times n}$ este o matrice simetrică pozitiv semidefinită, i.e.

$$A^T A = P \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} P^T \qquad \text{unde } P^T P = I_n$$

Spectrul $\{\lambda_i\}_{1 \le i \le n}$ reprezintă valorile proprii ale matricii A.

Dacă presupunem disponibilă DVS a matricii A, atunci

$$A^T A = (U\Sigma V^T)^T (U\Sigma V^T) = V\Sigma \underbrace{U^T U}_{I_n} \Sigma V^T$$
$$= V \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \sigma_n^2 \end{bmatrix} V^T$$

Se identifică:

$$V^T = P^T$$
$$\sigma_i^2 = \lambda_i$$

Vectorii proprii ai lui $A^T A$ compun vectorii singulari la dreapta a lui A!

Aplicăm o procedură similară pentru vectorii singulari la stânga:

$$AA^T = (U\Sigma V^T)(U\Sigma V^T)^T = U\Sigma \underbrace{V^T V}_{I_n} \Sigma U^T$$

$$= U \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \sigma_m^2 \end{bmatrix} U^T$$

Matricea AA^T este diagonalizabilă, deci este posibil calculul vectorilor proprii $AA^T = QSQ^T$

Se identifică:

$$U = Q$$

Vectorii proprii ai lui AA^T compun vectorii singulari la stânga ai lui A!

- Step 1: Formează $C = A^T A$
- Step 2: Aplică algoritmul QR simetric pentru calculul V astfel: $V^T C V = \text{diag}\left(\sigma_1^2, \cdots, \sigma_n^2\right)$
- Step 3: Calculează $u_i = \frac{1}{\sigma_i^2} A v_i$ pentru $i = 1 : n$
- Step 3': Calculează folosing fact. QR $AV = \underbrace{U}_{Q} \underbrace{\Sigma}_{R}$

- Step 1: Formează $C = A^T A$
- Step 2: Aplică algoritmul QR simetric pentru calculul V astfel: $V^T C V = \text{diag}\left(\sigma_1^2, \cdots, \sigma_n^2\right)$
- Step 3: Calculează $u_i = \frac{1}{\sigma_i^2} A v_i$ pentru $i = 1 : n$
- Step 3': Calculează folosing fact. QR $AV = \underbrace{U}_{Q}\underbrace{\Sigma}_{R}$

În practică formarea $A^T A$ sau AA^T este costisitoare!

- Step 1: Formează $C = A^T A$
- Step 2: Aplică algoritmul QR simetric pentru calculul V astfel: $V^T C V = \text{diag}\left(\sigma_1^2, \cdots, \sigma_n^2\right)$
- Step 3: Calculează $u_i = \frac{1}{\sigma_i^2} A v_i$ pentru $i = 1 : n$
- Step 3': Calculează folosing fact. QR $AV = \underbrace{U}_{Q} \underbrace{\Sigma}_{R}$

În practică formarea $A^T A$ sau AA^T este costisitoare! Exemplu î n carte!

- Van Loan, Charles F., and G. Golub. ”Matrix computations (Johns Hopkins studies in mathematical sciences).” (1996).