Přednáška 25 – SVD & vlastní čísla

Opáčko: - viděli jsme užitečnost SVD matice dat $\to$ komprese, analýza

- jak spočítat & co stabilita & podmíněnost?

Linalgebra 2: $A = U\Sigma V^T \Rightarrow A^TA = V\Sigma^T U^T U \Sigma V^T = V\Sigma^T\Sigma V$, V unitární, $\Sigma^T\Sigma$ diagonální

$\rightsquigarrow A^TA = V\Sigma^T\Sigma V$ je spektrální rozklad (= diagonalizace) symetrické $\underbrace{\text{pozitivně semi-definitní}}$ matice A^TA

= všechna vlastní čísla A^TA (tj. prvky na diagonále $\Sigma^T\Sigma$, tj. $\sigma_1^2, \dots, \sigma_n^2$) jsou nezáporná

$\Rightarrow$ výpočet SVD se standardně převádí na výpočet <u>spektrálního rozkladu</u>, (= diagonalizace = EVD = eigenvalue decomposition)

tj. pro danou matici M nalezení jejích vlastních čísel a vlastních vektorů (protože A^TA je symetrická $\Rightarrow$ její EVD vždy existuje)

Motivace č. 2 pro výpočet vl. čísel

Algoritmus vyhledávače google $\to$ PageRank (vymyslel Larry Page & Sergey Brin):

Krok 1: „sestrojíme" si seznam všech webpages $\rightsquigarrow$ délka $n \in \mathbb{N}$

Krok 2: „sestrojíme" si matici $A \in \mathbb{R}^{n\times n}$ jako
$$A_{ij} = \begin{cases} 0 & \text{pokud se webpage } L(i) \text{ nezmiňuje o webpage } L(j) \\ \frac{1}{k} & \text{pokud se webpage } L(i) \text{ zmiňuje o webpage } L(j) \text{ \& } k = \#\text{ všech stránek, které webpage } L(i) \text{ zmiňuje} \end{cases}$$

$\rightsquigarrow$ proč takto? Odpovídá modelu „sdílení důležitosti/popularity", viz googlesheet na cviku.

Krok 3: spočteme dominantní vlastní vektor matice $A \to \vec{v} \in \mathbb{R}^n$.
Lze ukázat, že všechny jeho složky jsou „≥ 0" (= mají stejné znaménko) $\rightsquigarrow$ lze interpretovat jako „ekvilibrium popularity stránek v L", tj. $(\vec{v})_i = v_i > v_j \iff$ stránka $L(i)$ bude podle tohoto modelu populárnější než $L(j)$

Krok 4: Na outputu seřadíme webpages z listu L sestupně podle velikosti složek vektoru $\vec{v} \in \mathbb{R}^n$.

Výpočet dominantního vlastního páru $(\lambda, \vec{v})$

Předpokládejme, že máme $A = S \Lambda S^{-1}$ … S … matice vl. vektorů, $\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ … matice vl. čísel

$$A^k = S \Lambda^k S^{-1} = S \begin{bmatrix} \lambda_1^k & & \\ & \ddots & \\ & & \lambda_n^k \end{bmatrix} S^{-1} \rightsquigarrow \text{a tedy}$$

$$|\lambda_1| > |\lambda_i| \ \forall i > 1 \quad \Downarrow \quad |\lambda_1^k| \gg |\lambda_i^k| \ \forall i > 1$$

$\rightsquigarrow$ a tedy dominantní vl. pár je „ještě více dominantní"

Vezměme si vektor $\vec{w} = c_1 \cdot \vec{s}_1 + c_2 \cdot \vec{s}_2 + \dots + c_n \cdot \vec{s}_n = S \cdot \vec{c}$ (ve vl. bázi A je $\vec{w}$ pouze $\vec{c}$). Pak:

$$A^k \vec{w} = S \Lambda^k S^{-1} \cdot S \vec{c} = S \Lambda^k \vec{c} = S \cdot \begin{bmatrix} \lambda_1^k \cdot c_1 \\ \lambda_2^k \cdot c_2 \\ \vdots \\ \lambda_n^k \cdot c_n \end{bmatrix} = \lambda_1^k c_1 \vec{s}_1 + \dots + \lambda_n^k c_n \cdot \vec{s}_n$$

$$= \lambda_1^k \cdot \left[c_1 \vec{s}_1 + c_2 \left(\frac{\lambda_2}{\lambda_1}\right)^k \vec{s}_2 + \dots + c_n \left(\frac{\lambda_n}{\lambda_1}\right)^k \vec{s}_n \right]$$

$\to 0$ pro $k \to +\infty$ pokud $|\lambda_1| > |\lambda_n|$

$|\lambda_1| > |\lambda_2| \Rightarrow \left(\frac{\lambda_2}{\lambda_1}\right)^k \to 0$ pro $k \to +\infty$

$\Downarrow$

pokud $c_1 \neq 0$ & $|\lambda_1| > |\lambda_i| \ \forall i > 1$

pak $A^k \vec{w} \longrightarrow \lambda_1^k \cdot c_1 \cdot \vec{s}_1$ ⟵ vektor ve směru dominantního vl. vektoru A

$\to$ toto je základní myšlenka tzv. mocninné metody (≡ power method)

V praxi je zapotřebí normalizace

$\to$ tj. nepočítáme $A \cdot A \cdots A \vec{w}$, ale

$$\boxed{\vec{v}_0 \to \vec{w}_1 = A\vec{v}_0,\ \vec{v}_1 = \frac{\vec{w}_1}{\|\vec{w}_1\|} \to \vec{w}_2 = A\vec{v}_1,\ \vec{v}_2 = \frac{\vec{w}_2}{\|\vec{w}_2\|} \to \dots \to \vec{w}_k = A\vec{v}_{k-1},\ \vec{v}_k = \frac{\vec{w}_k}{\|\vec{w}_k\|}}$$

$\rightsquigarrow$ pokud $\vec{v}_k \approx \vec{s}_1$ pak $A\vec{v}_k \approx \lambda_1 \vec{v}_k$

a tedy $\vec{v}_k^T A \vec{v}_k \approx \lambda_1 \vec{v}_k^T \vec{v}_k$ … neboli $\boxed{\lambda_1 \approx \mu_k := \frac{\vec{v}_k^T A \vec{v}_k}{\vec{v}_k^T \vec{v}_k}}$ tzv. Rayleighův podíl

Zobecnění I:

- co když chci nejmenší vl. pár?

  $\rightsquigarrow$ λ_{min} matice A je λ_{max} matice A^{-1} protože

  $|\lambda_n| < |\lambda_i| \ \forall i < n$ & $\lambda_i \neq 0 \implies \left|\frac{1}{\lambda_n}\right| > \left|\frac{1}{\lambda_i}\right| \ \forall i < n$

$$\Rightarrow \boxed{\vec{v}_0 \to \vec{w}_1 = A^{-1}\vec{v}_0,\ \vec{v}_1 = \frac{\vec{w}_1}{\|\vec{w}_1\|} \to \vec{w}_2 = A^{-1}\vec{v}_1,\ \vec{v}_2 = \frac{\vec{w}_2}{\|\vec{w}_2\|} \to \dots \to \vec{w}_k = A^{-1}\vec{v}_{k-1},\ \vec{v}_k = \frac{\vec{w}_k}{\|\vec{w}_k\|}}$$

& jako vždy a všude místo výpočtu $A^{-1}\vec{v}_{i-1}$ budeme řešit $A\vec{w}_i = \vec{v}_{i-1}$

- co když nechci „největší" ani „nejmenší", ale nejbližší číslu $\alpha \in \mathbb{C}$?

  $\rightsquigarrow$ stejná úvaha, místo násobení A budu „násobit" $(A - \alpha I)^{-1}$

Zobecnění II: • co když nechci 1 vl. pár, ale všechny?

$\vec{v}_0 \to \vec{w}_1 = A\vec{v}_0,\ \vec{v}_1 = \frac{\vec{w}_1}{\|\vec{w}_1\|} \to \vec{w}_2 = A\vec{v}_1,\ \vec{v}_2 = \frac{\vec{w}_2}{\|\vec{w}_2\|} \to$
$\to \dots \to \vec{w}_k = A\vec{v}_{k-1},\ \vec{v}_k = \frac{\vec{w}_k}{\|\vec{w}_k\|}$

$\rightsquigarrow$ vybereme si rutinu na QR-faktorizaci, $qr: M \longmapsto Q_M, R_M$ t. že $M = Q_M R_M$

Chceme všechny vl. páry $\Rightarrow$ musíme iterovat s „n vektory" $\rightsquigarrow$ s maticemi

$A \equiv A_0 \to Q_0, R_0 := qr(A_0)$, $A_1 := R_0 Q_0$. Pak jistě $A_0 = Q_0 R_0 = Q_0 A_1 Q_0^T \to$
$\{vl. č. A_0\} = \{vl. č. A_1\}$

$\to Q_1, R_1 := qr(A_1)$, $A_2 := R_1 Q_1$ a opět $\{vl. č. A_1\} = \{vl. č. A_2\} \to$

$\to \dots \to Q_k, R_k := qr(A_k)$, $A_{k+1} := R_k Q_k$

tzv. QR-algoritmus

Podmíněnost

$\to$ odpověď na otázku: jak moc se liší vl. páry matice $A+E$ od vl. párů matice A pro $\|E\|$ malé

- A je symetrická $\to$ $|\lambda_i(A) - \lambda_i(A+E)| \leq \|E\|_2$
- A je obecná $\to$ nelze nic zaručit. Matematicky toto odpovídá podmíněnosti kořenů polynomu v závislosti na změně koeficientů v „neortogonální" bázi polynomů $\rightsquigarrow$ viz. Wilkinson's polynom na začátku semestru

($\Rightarrow$ problém nalezení SVD matice je dobře podmíněný)

Stabilita

- QR-algoritmus je stabilní $\to$ používáme unit. transf.
- mocninná metoda & variace $\to$ záleží na řešiči