Recitation #8

Irina Espejo (iem244@nyu.edu)

Center for Data Science

DS-GA 1014: Optimization and Computational Linear Algebra for Data Science

recall: Singular Value Decomposition (SVD)

SVD

Theorem: Let $A \in \mathbb{R}^{n \times m}$ then there exists two orthogonal matrices $U \in \mathbb{R}^{n \times n}$ and $V \in \mathbb{R}^{m \times m}$ and a matrix $\Sigma \in \mathbb{R}^{n \times m}$ such that

$$\boxed{A = U \Sigma V^T}$$

with $\Sigma_{11} \geq \Sigma_{22} \geq \ldots \geq 0$ and $\Sigma_{ij} = 0$ for $i \neq j$

recall: SVD

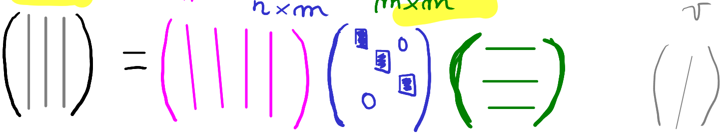
$$A = U \, \Sigma \, V^T$$

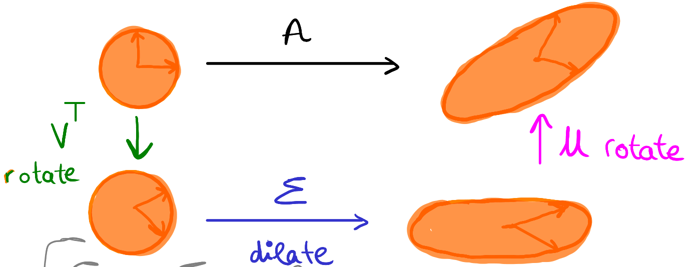
left

singular values

right

$n \times m$ $\quad$ $n \times n$ $\quad$ $n \times m$ $\quad$ $m \times m$



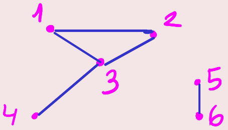


$$\boxed{\Sigma v = \sum \sigma_i v_i}$$

Adjacency matrix of a graph

Let G be a graph with n nodes, its adjacency matrix $A \in \ltimes \times \ltimes$ if defined by

$$A_{ij} = 1 \text{ if } i \sim j$$
$$A_{ij} = 0 \text{ else}$$



Laplacian of a graph

The Laplacian matrix $L \in \mathbb{R}^{n \times n}$ of a graph is

$$\boxed{L = D - A}$$

where $D \in \mathbb{R}^{n \times n}$ is the degree matrix $D = diag(d(1)...d(n))$.
Some properties:

- L symmetric and PSD (all eigenvalues are $\lambda \geq 0$)
- smallest eigenvalue $\lambda = 0$ with eigenvector $v = (1...1)$

Exercise 1

Handshaking lemma: let G be a graph with n nodes and m edges. Show that

$$\sum_{i=1}^{n} deg(\text{node}_i) = 2m$$

(if there is a party with n attendees then an even number of people shakes an odd number of other people's hands)

$$\text{Neigh}(i) = \{ j \in \{1 \ldots n\} \mid i \sim j \}$$

$A =$

If $j \in \text{Neigh}(i) \Rightarrow i \in \text{Neigh}(j)$

$$\sum_{i=1}^{n} deg(i) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} = 2 \underbrace{\sum_{i=1}^{n} \sum_{j=1}^{i} a_{ij}}_{m} \quad \square$$

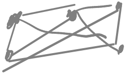
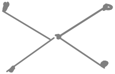
Exercise 2

Let G be a graph with n nodes and let λ_2 (Fiedler) be the smallest non-zero eigenvalue of the Laplacian L. Show that the value of λ_2 increases when one adds more edges to G for the same number of nodes n.

L : λ_1 ... λ_n // A : λ_1 ... λ_n

λ_1: $\downarrow\downarrow$, 0; λ_n: $\uparrow\uparrow$ // λ_1: $\uparrow\uparrow$; λ_n: $\downarrow\downarrow$

Goal $\lambda_2 \uparrow\uparrow$ when $m \uparrow\uparrow$ n fixed

$\lambda_2 \uparrow\uparrow$  $\lambda_2 \downarrow\downarrow$ 

Exercise 2

Recall for any $x \in \mathbb{R}^n$ $\quad x^T L x = \sum_{i \sim j} (x_i - x_j)^2$

$L v = \lambda_2 v \quad$ v eigenvector of λ_2

$$v^T \underbrace{L v} = \sum_{i \sim j} (v_i - v_j)^2$$

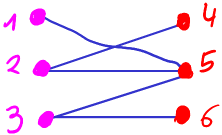
$$\lambda_2 \cancelto{1}{\|v\|^2} = \sum_{i \sim j} \underbrace{(v_i - v_j)^2}_{\geq 0}$$

if n increases

this sum has more terms and eventually it will increase. $\square$

Exercise 3

Let G be a connected graph with n nodes and let L be its Laplacian. Let λ_n be the highest eigenvalue of the Laplacian L. Show that $\lambda_n = 2$ if and only if the graph G is bipartite.
(*A bipartite graph has its set of nodes divided into two disjoint subgroups such that all the edges go from one group to the other but never between nodes of the same group*)

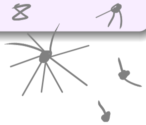


No need to do it, just know what a bipartite graph is

Exercise 4

Let G be a connected graph with n nodes and let A be its adjacency matrix. Show that the highest valued eigenvalue λ_1 is bounded by the maximum degree, that is

$$\lambda_1 \leq \max_{i \in \{1..n\}} deg(i)$$



let $v \in \mathbb{R}^n$ $\quad Av = \lambda_1 v$

$|\lambda_1| \geq |\lambda_i| \quad i \in 1,..,n$

without loss of generality

rescale $v \to \frac{v}{\|v\|}$

$|v_i| = 1$

Exercise 4

$$|\lambda_1| = |\lambda_1| \cdot |\cancel{v_i}^{\,1}|$$

$$= \left|\sum_{j=1}^{n} a_{ij} v_i\right|$$

triangular inequality

$$\leq \sum_{j=1} |a_{ij} v_i|$$

$$= \sum_{j=} |a_{ij}| \, |\cancel{v_i}^{\,1}|$$

$d(i)$

$$\leq \max_{k \in \{1,\dots,n\}} d(k)$$

$\square$

$$Av = \lambda_1 v$$

$$\sum_j a_{ij} v_i = \underline{\underline{\lambda_1 v_i}}$$

recall: Spectral clustering in graphs

the method

graph Laplacian L, number of clusters k

1. Compute the first k orthonormal vectors $v_1, .., v_k$ of the Laplacian L
2. Associate nodes to vectors in the following way: node i to vector $x_i = (v_2(i), .., v_k(i))$
3. Cluster the points $x_1, .., x_n$ with k-means
4. Deduce clustering nodes of the graph

extra SVD problems

Exercise 5

Let $M \in \mathbb{R}^{n \times m}$ have full rank and let $n \geq m$. ΣV^T.

1. Show that $M^T M$ is invertible
2. Which vnectors span the $Im(M)$? Write the matrix of orthogonal projection onto $Im(M)$ and give basis transformation for that matrix.
3. Let $w \in \mathbb{R}^{n}$ and let u be the orthogonal projection of w onto $Im(M)$. Show that $M^T u = M^T w$.
4. Show that $M(MM^T)^{-1}M^T$ is the matrix of an orthogonal projection onto $Im(M)$

Exercise 5

① Goal: $M^T M$ is invertible

SVD: $M = \mathcal{U}\Sigma V^T$ — M has full rank so $\mathcal{U}$ is invertible and $\Sigma_{ii} \neq 0 \ \forall i \in 1 \dots n$

Now, $M^T M = (\mathcal{U}\Sigma V^T)^T (\mathcal{U}\Sigma V^T)$

$= V\Sigma^T \cancel{\mathcal{U}^T \mathcal{U}} \Sigma V^T$ ($\mathcal{U}$ is orthogonal)

$\mathrm{rank}(\Sigma^T\Sigma) = n - \underbrace{\#((\Sigma^T\Sigma)_{ii} = 0)}_{\text{number of zeros in diagonal}}$ — But $\Sigma_{ii} \neq 0$ and $\Sigma^T_{ii} \neq 0$

Therefore $\mathrm{rank}(\Sigma^T\Sigma) = n$ and $M^T M$ is invertible □

(2) $\underset{n\times m}{M} = U \Sigma V^T$ $\quad M : \mathbb{R}^m \to \mathbb{R}^n$

- $\mathrm{Im}(M) = \{ y \in \mathbb{R}^n \mid \text{exists } x \in \mathbb{R}^m \text{ such that } Mx = y \}$

$$Mx = y \to U \Sigma \underbrace{V^T x}_{x' \in \mathbb{R}^m} = y \to U \Sigma x' = y$$

- Let $x = \sum_i^m \alpha_i v_i$ where v_i are the ~~ro~~ columns of V

$$x' = \sum_{i=1}^m v_i^T \alpha_i v_i = \sum_{i=1}^m \alpha_i e_i$$

$$\Sigma x' = \sum_{i=1}^{\min(n,m)} \alpha_i e_i \sigma_i$$

$$U \Sigma x' = \sum \alpha_i \sigma_i u_i$$

therefore

$\mathrm{Im}(M) = \mathrm{Span}(u_1, \ldots, u_K)$ columns of U

where K is the last index such that $\sigma_K \neq 0$.

2) matrix of the projection onto $\mathrm{Im}(M)$

$$P_{\mathrm{Im}(M)} = U_K U_K^T$$

where
$$U_K = \begin{pmatrix} | & | & & | \\ u_1 & u_2 & \cdots & u_K \\ | & | & & | \end{pmatrix}$$

where K is the last index such that $\sigma_i \neq 0$

singular values Σ

3

$$u = P_{Im(M)}(w) \qquad M^T = V \Sigma^T U^T$$

$$u = U_K U_K^T w$$

- First: let $u = \sum_{i=1}^{m} \alpha_i u_i$ so $M^T u = V \Sigma^T U^T u = \dots = \sum_i^{m} \sigma_i \alpha_i v_i$

- Second: let $w = \sum_{i=1}^{n} \alpha_i u_i$

$$M^T w = \sum_i^{\min(n,m)} \sigma_i \alpha_i v_i$$

Because $n > m$ they are identical

Exercise 5

4

$$P = M(M^TM)^{-1}M^T \qquad \begin{aligned} M &= U\Sigma V^T \\ M^T &= V\Sigma^T U^T \end{aligned}$$

Lets $x \in \mathbb{R}^m$

$$P = U\Sigma V^T(V\Sigma^T \cancel{U^T}\cancel{U}\Sigma V^T)^{-1} V\Sigma^T U^T$$

$$P = U\Sigma \cancel{V^T}(\cancel{V}(\Sigma^T\Sigma)^{-1}\cancel{V^T})\cancel{V}\Sigma^T U^T$$

$$P = U\underbrace{\Sigma(\Sigma^T\Sigma)^{-1}\Sigma^T}_{= \; n\times n} U^T = U_K U_K^T \quad \square$$

$$\begin{pmatrix} 1 & & & & 0 \\ & 1 & & & \\ & & 1 & & \\ & & & 0 & \\ 0 & & & & 0 \end{pmatrix} \quad m$$

identity matrix but only m ones

$K = m$

Questions