LAIS, Lecture #21

"Low-rank matrix completion"

$\mho_\Omega : M(D\times N, r) \to \mathbb{R}^\Omega$

$\Omega = \bigcup_{j\in[N]} \omega_j \times \{j\} = \bigcup_{i\in[D]} \{i\}\times \psi_i$

$\omega_j \subseteq [D]$, $\psi_i \subseteq [N]$

Prp 1 $S^* = \mathcal{B}(\mathcal{X}^*) \subseteq \mathbb{R}^D$, $\dim \mho_{\omega_j}(S^*) = r \ \forall j$

Then $\mathcal{X}^*$ is uniquely determined from $\mho_\Omega(\mathcal{X}^*)$, S^*.

(Similarly for the row-space working with $\mathcal{X}^{*T}$) ⊡

Dfn 2 [subspace formulation of LRMC] $S^* \subseteq \mathbb{R}^D$

$x_j^* \in S^*$, $j\in[N]$

$\omega_j \subset [D]$, $\mho_{\omega_j} : \mathbb{R}^D \to \mathbb{R}^{\#\omega_j}$

given: $\mho_{\omega_j}(x_j^*)$, $j\in[N]$

goal: find S^*. ⊡

$\mathcal{X}^* = [x_1^* \dots x_N^*]$

given $\mho_\Omega(\mathcal{X}^*)$

ulation

$\mathbb{R}^D$

$\mathbb{R}^{\#\omega_j}$

$N]$

Prp 3 If $\#\omega_j < r$ for some j, then $\upsilon_\Omega^{-1}(\upsilon_\Omega(\mathcal{X}^*))$ is infinite.

Prf $\mathcal{X}^* = [x_1^* \cdots x_j^* \cdots x_N^*]$

$B = [b_1 \cdots b_r] \in \mathbb{R}^{D \times r}$ basis for S^*, $\upsilon_{\omega_j}(x_j^*) \in \upsilon_{\omega_j}(S^*)$

$\Rightarrow$ the linear system

$$\underbrace{\upsilon_{\omega_j}(B)}_{\#\omega_j \times r}\ \underbrace{c_j}_{r} = \underbrace{\upsilon_{\omega_j}(x_j^*)}_{\#\omega_j} \quad (A)$$

is consistent

Since $\upsilon_{\omega_j}(B^*)$ has more columns than rows

$\dim \mathcal{N}(\upsilon_{\omega_j}(B^*)) \geq 1$

so there are infinitely many solutions c_{jk} to (A). For every such solution c_{jk} define

$y_{jk} = B c_{jk} \in S^*$.

Since B has full column rank,

$c_{jk} \neq c_{ji} \Rightarrow y_{jk} \neq y_{ji}$

Also $\upsilon_{\omega_j}(y_{jk}) = \upsilon_{\omega_j}(B) c_{jk} = \upsilon_{\omega_j}(x_j^*)$.

Define

$\mathcal{X}_k = [x_1^* \cdots x_{j-1}^*\ y_{jk}\ x_{j+1}^* \cdots x_N^*]$

$\operatorname{rank}(\mathcal{X}_k) \leq r$

$\upsilon_\Omega(\mathcal{X}^*) = \upsilon_\Omega(\mathcal{X}_k)$. $\blacksquare$

Cor 4 $\#\omega_j, \#\psi_i \ge r \ \forall i,j$ is necessary for unique completion. ∎

Que 5 What is the minimal $\#\Omega$ for unique completion?

A $\#\Omega \ge \dim M(D\times N, r) = r(D+N-r)$. ∎

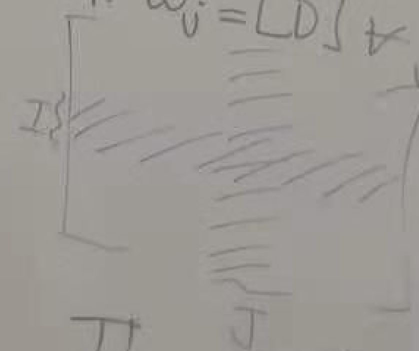
intuitive explanation

$\begin{bmatrix} I_m \\ \\ \end{bmatrix}_{D\times r} \quad \begin{bmatrix} & & \end{bmatrix}_{r\times N}$

$Dr + Nr - r^2$

Que 6 Where should we put those observed entries?
(active matrix completion) ∎
(algebraic combinatorics)

Thm 7 Suppose $\exists\ I \subset [D]$ with $\psi_i = [N]$ $\forall i \in I$ and $\#I = r$, and $\exists\ J \subset [N]$, $\#J = r$ s.t. $\omega_j = [D]$ $\forall j \in J$.



Suppose $\det(X^*_{I,J}) \neq 0$.

Then $\mathcal{P}_\Omega^{-1}(\mathcal{P}_\Omega(X^*)) = \{X^*\}$. ∎

LA1S, Lecture #21

"Spectral Clustering"

* simple principle
* flexible
* linear algebra

Suppose we have a population of N subjects/objects and a measure of similarity $w_{ji} = \underset{\sim}{w_{ij}} \geq 0 \; \forall \, i,j \in [N]$

The higher w_{ij} is, the more similar objects i and j are.

$W \in \mathbb{R}^{N \times N}_{\geq 0}$ similarity/affinity matrix.

Ex 1 [affinity matrix]

Suppose we have a set of measurements $x_j \in \mathbb{R}^D$ for any object $j \in [N]$.

i) $w_{ij} = |x_j^T x_i| \,/\, \|x_i\|_2 \|x_j\|_2$

ii) Gaussian Kernel

$$w_{ij} = \exp\left(- \frac{\|x_i - x_j\|_2^2}{2\sigma}\right)$$

$W \longrightarrow$ chop W with "nearest neighbors"

optional

Dfn 2 $G=(V,E)$ weighted graph

$V=[N]$ vertices

$$E=\left\{(i,j) : \begin{array}{l} i,j\in[N] \\ W_{ij}>0 \end{array}\right\}$$

We assume G has c connected components with vertices $G_1,\dots,G_c$.

So as a set

$$V=[N]=\bigcup_{i\in[c]} G_i$$

"The similarity between i,j is $\neq 0$ if and only if i,j are in the same class ($\exists\, k\in[c]$ s.t. $i,j\in G_k$) and each class is connected."

Dfn 3 Degree matrix

$$D_{ij}=\begin{cases} 0, & i\neq j \\ \sum_{k\in[N]} W_{ik}, & i=j \end{cases}$$

Dfn 4 Laplacian matrix

$$L=D-W.$$

Prp 5 $L\geq 0$

Prf $\alpha\in\mathbb{R}^N$

$$\alpha^T L\alpha=\alpha^T D\alpha-\alpha^T W\alpha$$

$$=\sum_{i\in[N]}\alpha_i^2\Big(\sum_j W_{ij}\Big)-\sum_i\alpha_i\Big(\sum_j W_{ij}\alpha_j\Big)$$

$\alpha_k^2 W_{kl}$

$\alpha_l^2 W_{lk}$

$-\alpha_k W_{kl}\alpha_l$

$-\alpha_l W_{lk}\alpha_k$

$=$

$$= \sum_{i<j} w_{ij}(\alpha_i - \alpha_j)^2 \geq 0 \quad \boxminus$$

$\boxminus$

atrix

Dfn 6 $i \in [c]$

e_{G_i} : indicator vector of class i. That is

$$e_{G_i}(\ell) = \begin{cases} 1, & \text{if } \ell \in G_i \\ 0, & \text{if } \ell \notin G_i \end{cases} \quad \boxminus$$

Thm 7 $\dim \mathcal{N}(L) = c$ and $e_{G_1}, \dots, e_{G_c}$ is a basis for $\mathcal{N}(L)$.

$A \geq 0 \quad A\xi = 0$

$\Leftarrow \Rightarrow \xi^T A \xi = 0 \quad A = B^T B$

$\Rightarrow B\xi = 0 \Rightarrow B^T B \xi = 0$

$W\alpha$

$\left(\sum_j w_{ij}\alpha_j\right)$

α_ℓ

α_k

Prf We first prove $e_{G_k} \in \mathcal{N}(L)$.

$$e_{G_k}^T L e_{G_k} \overset{Prp\,5}{=} \sum_{i<j} w_{ij}\left(e_{G_k}(i) - e_{G_k}(j)\right)^2 = $$

$$= \sum_{\substack{i,j \in G_\ell \\ \text{for some } \ell \in [c]}} \underbrace{w_{ij}}_{>0}\left(e_{G_k}(i) - e_{G_k}(j)\right)^2 = 0$$

$\Rightarrow e_{G_k} \in \mathcal{N}(L) \ \forall k \in [c]$

Clearly the e_{G_k}'s are linearly independent.

Let us prove that the e_{G_k}'s are a spanning set for $\mathcal{N}(L)$.

Take

$\overset{Prp\,5}{\Rightarrow}$

$\Rightarrow$

fo

$\Rightarrow$

$\Rightarrow$

... first prove

$(\mathcal{L})$.

Prp5

$_k =$

$\left(e_{G_k}(i) - e_{G_k}(j)\right)^2 =$

$\left(e_{G_k}(i) - e_{G_k}(j)\right)^2 = 0$

0

]

$\in \mathcal{N}(\mathcal{L}) \ \forall k \in [c]$

the e_{G_k}'s are

independent.

prove that the

are a span...

Take $\xi \in \mathcal{N}(\mathcal{L}) \Rightarrow \xi^T \mathcal{L} \xi = 0$

$$\overset{\text{Prp}5}{\Longrightarrow} \sum_{i<j} w_{ij} \left(\xi(i) - \xi(j)\right)^2 = 0$$

$$\Rightarrow \sum_{\substack{i,j \in G_k \\ \text{for some} \\ k \in [c]}} \underbrace{w_{ij}}_{>0} \left(\xi(i) - \xi(j)\right)^2 = 0$$

here we used connectedness

$\Rightarrow \xi(i) = \xi(j)$ whenever $i, j \in G_k$ for some k

$\Rightarrow \xi = \alpha_1 e_{G_1} + \dots + \alpha_c e_{G_c}$ ∎

LAIS, Lecture #21

Thm 8 Let $\xi_1, \dots, \xi_c$ be a basis for $\mathcal{N}(\mathcal{L})$.
Form the matrix

$$Y = \begin{bmatrix} \xi_1^T \\ \vdots \\ \xi_c^T \end{bmatrix} \in \mathbb{R}^{c \times N}$$

$$= [y_1 \cdots y_N]$$

Then $y_i = y_j \Leftrightarrow$ $i, j \in G_k$ for some $k \in [c]$

Prf $[\xi_1 \cdots \xi_c] = [e_{G_1} \cdots e_{G_c}] A$ for some $A \in \mathbb{R}^{c \times c}$ invertible

$$\Rightarrow Y = A^T \begin{bmatrix} e_{G_1}^T \\ \vdots \\ e_{G_c}^T \end{bmatrix} \quad \boxed{=}$$

mapping

$$[N] \longrightarrow \mathbb{R}^c$$

"k-means"