HW 2

1) a) WE HAVE $Av = \lambda v$ AND $Av' = \lambda' v'$ WITH $\lambda' \neq \lambda$; v IS ORTHOGONAL TO v' IF $\langle v; v' \rangle = 0$

$$\langle v, v' \rangle = \frac{1}{\lambda} \langle \lambda v, v' \rangle = \frac{1}{\lambda} \langle A v ; v' \rangle = \frac{1}{\lambda} \langle v ; A^t v' \rangle$$

BUT A IS SYMMETRIC $A^t = A$ $\rightarrow$ $\frac{\lambda'}{\lambda} \langle v, v' \rangle = \langle v, v' \rangle$

$\rightarrow$ $\langle v, v' \rangle = 0$ IF $\lambda \neq \lambda'$

1 b) GIVEN A SUBSPACE V OF R^d LET $v_1 \dots v_n$ BE AN ORTHONORMAL BASIS FOR V; NOW BY DEFINITION SPAN $\{v_1 \dots v_n\} = V$ BUT SINCE S SPAN V EVERY v_i CAN BE WRITTEN AS $\sum_i \alpha_i S_i$ WHERE $S_i \in S$ AND $AS_i = \lambda$

FOR THIS REASON $A v_i = \lambda v_i$ i.e THEY ARE EIGENVECTOR OF A WITH EIGENVALUE λ.

1 c) THE SPACE OF DIMENSION d CAN BE WRITTEN AS AN ORTHGONAL SUM OF THE SUBSPACES SPANNED BY THE DIFFERENT EIGENVALUE BECAUSE THEY ARE ORTHOGONAL AND HAVE DIMENSION EQUAL TO THE MOLTEPLICITY OF THE EIGENVALU

SO EVERY $v \in R^d = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \dots$; IN $V_{\lambda_2} \oplus \dots$ SPACE A IS DIAGONAL AND SO $A = U \Lambda U^T = \sum_i \lambda_i v_i v_i^T$

~~SO A CAN BE WRITTEN AS A LINEAR~~

2 a) EVERY VECTOR IN V CAN BE EXPRESS IN THE BASE $P_1 \ldots P_n$ SO $Y = \sum_i (P_i \cdot v) P_i$

NOW

$$\underset{Y \in V}{\| x - y' \|^2} = \| x^2 \| + \| \overbrace{\sum_i (P_i v) \cdot P_i}^{y'} \|^2 - 2 \| \sum (P_i \cdot v)(P_i \cdot x) \|$$

$$\frac{\partial}{\partial v} \| \quad \| = 0 \qquad 2 \sum (P_i \cdot v_i) - 2 \sum P_i (P_i \cdot x) = 0$$

P_i ARE UNIT VECTORS $|P_i| = 1$

$$\longrightarrow \qquad v = \sum (x \cdot P_i) \cdot P_i$$

ANOTHER WAY TO SHOW THAT IS THE FOLLOWING:

NOTE THAT GIVEN

$P_V(x) = \sum (x \cdot P_i) P_i$ WE HAVE $x - P_V(x) \in V^{\perp}$ SINCE

$\langle x - P_W(x), P_i \rangle = 0$; NOW $\| x - y \|^2 = \underbrace{\| x - P_V(x) \|^2}_{\in V^{\perp}} + \underbrace{\| P_V(x) - y \|^2}_{\in V}$

WHERE WE HAVE APPLIED THE PYTHAGOREAN THEOREM;

THIS IS $\| x - y \|^2 \geq \| x - P_V(x) \|^2$ FOR EVERY X AND IT IS EQUAL WHEN $Y = P_V(x) = \sum (x \cdot P_i) P_i$

2 b) $\| x_i - (P_i \cdot x) P_i \|^2 = (x_i - (P \cdot x) P_i)(x_i - (P \cdot x) P)$

$= x_i \cdot x_i - (P_i \cdot x) P_i \cdot x_i - (P \cdot x)(P_i x_i) + (P_i \cdot x)^2 P_i \cdot P$

$= x_i \cdot x_i - 2 (P_i \cdot x)^2 + (P_i \cdot x)^2 \cdot 1$

$= x_i \cdot x_i - (P_i \cdot x)^2$;

THE FIRST TERM DOES NOT DEPEND ON P SO TO MINIMIZE $\| x_i - (P \cdot x) P_i \|$ WE WANT TO MAXIMIZE $\frac{1}{m} \sum_i (P \cdot x_i)$

$$\frac{1}{m} \sum (x_i \cdot P)^t (x_i P) = \frac{1}{m} \sum P^t x_i^t x_i P = P^t \hat{\Sigma} P$$

SO THIS IS EQUIVALENT TO THE RAYLEIGH QUOTIENT PROBLEM; SINCE WE WANT TO MAXIMIZE WE TAKE THE LEADING EIGEN VECTORS

3) THE GRAM MATRIX IS A $m \times m$ MATRIX DEFINED AS $G_{ij} = x_i \cdot x_j$ WHERE $x_i \in \mathbb{R}^d$;

NOW $\text{rank}(A \cdot B) \leq \min(\text{rk}(A)\ \text{rk}(B))$; IN THIS CASE

$\text{rank}(G) = \text{rk}(x_i \cdot x_i) \leq \text{rank}(\bar{x}_i)$, x_i IS A $1 \times m$ COMPOSED OF ELEMENTS OF $\mathbb{R}^d$ FOR THIS REASON $\text{rk}(x) \leq d$ SINCE ONLY d OF THEM CAN BE LINEARLY INDIPENDENT; SO $\text{rk}(G) \leq \text{rk}(x) \leq d$

3b) GIVEN $K \in \mathbb{R}^{m \times m}$ I CAN TAKE ITS EIGEN DECOMPOSITION SINCE IT IS SIMMETRIC AND POSITIVE SEMI-DEFINITE

$K = Q \Lambda Q^T$ WHER $Q \in \mathbb{R}^{m \times d}$ BECAUSE $\text{rk}(K) = d$

DEFINE $x_i = [Q \Lambda^{1/2}]_i$; NOW IT IS EASY TO SEE THAT

THE GRAM MATRIX OF $x_i \ldots x_m$ $G = x_i x_j = [Q\Lambda^{1/2}]_i \cdot [Q\Lambda^{1/2}]_j$

IS EXACTLY K

4) LET'S WRITE $\psi(Y)$ AS $\sum_{ij} M_{ij}(Y_i \cdot Y_j)$ WITH

$M_{ij} = \delta_{ij} - W_{ij} - W_{ji} + \sum_{k} W_{ki} W_{kj} = (\mathbb{1} - W)^t (\mathbb{1} - W)$

WHERE $\mathbb{1}$ IS THE IDENTITY MATRIX;

SO WE NEED TO MINIMIZE $\psi(Y) = Y^T M Y$ WITH THE CONSTRAINT

$\mathcal{L}(Y;\mu) = Y^t M Y - \mu(m^{-1} YY - 1)$

WITH $\frac{\partial \mathcal{L}}{\partial \mu} = 0$ $\quad 2MY - 2\mu m^{-1} Y = 0 \longrightarrow \boxed{MY = \frac{\mu}{m} Y}$

THIS MEANS THAT Y MUST BE AN EIGENVECTOR OF M;

BECAUSE WE ARE TRYINT TO MINIMIZE $Y^t M Y$ WE NEED TO TAKE THE EIGENFUNCTION CORRESPONDING TO THE SMALLEST EIGEN VALUES.

THE MEAN CONSTRAINT IS NOT NEEDED; INDEED IF WE INCLUDE $A = \mathbb{1}\mathbb{1}^t$ THE CONSTRAINT IS $AY=0$

THIS IS A CONSERQUENCE OF TRANSLATIONAL INVARIANCE

5 BEING L THE LAPLACIAN WE HAVE

$$f^T L f = \sum_{i\sim j} w_{ij} \left(f(v_i) - f(v_j)\right)^2$$ BUT BECAUSE IT IS A

BIPARTITION $w_{ij} \neq 0$ IF $v_i \in S$ AND $v_j \in \bar{S}$ SO GIVEN THE DEFINITION OF f

$$= \sum_{\substack{v_i \in S \\ v_j \in \bar{S}}} w_{ij} \left(\frac{|\bar{S}|}{S} + \frac{|S|}{|\bar{S}|} + 2\underbrace{\sqrt{\frac{|S||\bar{S}|}{|S||\bar{S}|}}}_{=2} \right) = \sum \left(\frac{|\bar{S}|}{S} + \frac{|S|}{|\bar{S}|} + \underbrace{\frac{S}{S} + \frac{\bar{S}}{\bar{S}}}_{=2} \right) w_{ij}$$

$$= \sum w_{ij} \left(\frac{1}{|S|} + \frac{1}{|\bar{S}|} \right) (S + \bar{S}) = (S+\bar{S})\, \phi(S;\bar{S}) = n\, \phi(S;\bar{S})$$

SINCE $|S| + |\bar{S}| = n$ TOTAL NUMBER OF VERTICES

5 b) $$\bar{f} = \begin{pmatrix} f_i \\ \vdots \\ f_n \end{pmatrix} = \begin{pmatrix} f(v_i) \\ \vdots \\ f(v_n) \end{pmatrix}$$ SO $$f \cdot \mathbb{1} = \sum_i^n f_i$$

$$= \sum_{v_i \in S} \sqrt{\frac{|\bar{S}|}{|S|}} + \sum_{v_i \in \bar{S}} \left(- \sqrt{\frac{|S|}{|\bar{S}|}} \right) = |S| \sqrt{\frac{|\bar{S}|}{S}} - |\bar{S}| \sqrt{\frac{|S|}{|\bar{S}|}} = 0$$

5 c) $$\| f^2 \| = \sum_i^n |f_i|^2 = \sum_{v_i \in S} \frac{|\bar{S}|}{|S|} + \sum_{v_i \in \bar{S}} \frac{|S|}{|\bar{S}|} = |S| \frac{|\bar{S}|}{|S|} + |\bar{S}| \frac{|S|}{|\bar{S}|} = |S| + |\bar{S}| = n$$

BECAUSE $V = S \cup \bar{S}$ IS A BIPARTITION OF $V = \{v_1 \ldots v_n\}$

—o—

- THIS IS A RELAXATION OF THE RATIO CUT PROBLEM: IT TURNS A DISCRETE NP PROBLEM INTO A PROBLEM OF SPECTRAL CLUSTERING ON A GRAPH SOLVABLE AS AN EIGENVALUE PROBLEM

5 CONTINUED) ...

LAPLACIAN EIGENMAPS ALSO TURNS OUT TO RELY ON A GENERALIZED EIGENVALUE PROBLEM INVOLVING THE GRAPH LAPLACIAN; IN THAT CASE THE STRAIN, THE DISTANCE BETWEEN POINTS WAS TO BE MINIMIZED INSTEAD OF THE RATIO-CUT COST FUNCTION; BOTH THE 2 PROBLEMS RELY ON THE SPECTRAL PROPERTIES OF THE GRAPH, NOT SURPRISINGLY.