4.1 (a)

$$\underset{C}{\operatorname{argmin}} \sum_{j=1}^{K} \sum_{x \in C_j} \| x - \mu_j \|^2$$

$$= \underset{C}{\operatorname{argmin}} \sum_{j=1}^{K} \sum_{x \in C_j} (x-\mu_j)^T (x-\mu_j)$$

$$= \underset{C}{\operatorname{argmin}} \left(\sum_{j=1}^{K} \sum_{x \in C_j} x^T x - 2 \sum_{j=1}^{K} \sum_{x \in C_j} x^T \mu_j + \sum_{j=1}^{K} \sum_{x \in C_j} \mu_j^T \mu_j \right)$$

$$= \underset{C}{\operatorname{argmin}} \left(-2 \sum_{j=1}^{K} \mu_j^T \sum_{x \in C_j} x + \sum_{j=1}^{K} \mu_j^T \mu_j |C_j| \right)$$

$$= \underset{}{\operatorname{argmin}} \left(- \sum_{j=1}^{K} |C_j| \mu_j^T \mu_j \right) \qquad ①$$

$$\operatorname{argmin} \sum_{j=1}^{K} \frac{1}{2|C_j|} \sum_{x \in C_j} \sum_{x' \in C_j} (x-x')^T (x-x')$$

$$= \operatorname{argmin} \sum_{j=1}^{K} \frac{1}{2|C_j|} \sum_{x \in C_j} \sum_{x' \in C_j} \left(x^T x - 2 x'^T x + x'^T x' \right)$$

$$= \operatorname{argmin} \sum_{j=1}^{K} \frac{1}{2|C_j|} \left(\sum_{x \in C_j} 2|C_j| x^T x - 2 \sum_{x \in C_j} \sum_{x' \in C_j} x'^T x \right)$$

$$= \operatorname{argmin} \sum_{j=1}^{K} \frac{1}{2|C_j|} \left(-2 \sum_{x' \in C_j} x'^T \sum_{x \in C_j} x \right)$$

$$= \operatorname{argmin} \left(- \sum_{j=1}^{K} |C_j| \mu_j^T \mu_j \right) \qquad ②$$

$$\therefore ① = ②.$$

(b) $\arg\max \sum_{i=1}^{K} \sum_{j=1}^{K} |C_i||C_j|(\mu_i-\mu_j)^T(\mu_i-\mu_j)$

$= \arg\max \sum_{i=1}^{K} \sum_{j=1}^{K} |C_i||C_j|(\mu_i^T\mu_i - 2\mu_i^T\mu_j + \mu_j^T\mu_j)$

$= \arg\max \left(2\sum_{i=1}^{K} m|C_i|\mu_i^T\mu_i - 2\sum_{i=1}^{K}\mu_i|C_i|\sum_{j=1}^{K}\mu_j|C_j|\right)$

$= \arg\max \left(2m\sum_{i=1}^{K}\mu_i^T\mu_i|C_i| - 2\left(\sum_{i=1}^{m} x_i\right)^2\right)$

$= -\arg\min \sum_{i=1}^{K} \mu_i^T\mu_i|C_i|$

4.2 (a) let u have some diction of $x^{(i)}$:

$$u = \frac{x^{(i)}}{\|x^{(i)}\|_2}$$

$$\left\| x^{(i)} - \left(x^{(i)T}\frac{x^{(i)}}{\|x^{(i)T}\|_2}\right)\frac{x^{(i)}}{\|x^{(i)}\|_2}\right\|^2$$

$$= \left\| x^{(i)} - \left(\frac{\|x^{(i)}\|_2^2}{\|x^{(i)}\|_2}\right)\frac{x^{(i)}}{\|x^{(i)}\|_2}\right\|^2$$

$= 0$

$\because \|x^{(i)} - (x^{(i)T}u)u\|_2^2 \geq 0$

$\therefore \underset{u^Tu=}{\operatorname{argmin}} \|x^{(i)} - (x^{(i)T}u)u\|^2 = \frac{x^{(i)}}{\|x^{(i)}\|_2}$

$\because \left(x^{(i)T} \frac{x^{(i)}}{\|x^{(i)}\|_2}\right)^2$

$= \left(\frac{\|x^{(i)}\|_2^2}{\|x^{(i)}\|_2}\right)^2$

$= \|x^{(i)}\|_2^2$, 而 $(x^{(i)T}u)^2 \leq \|x^{(i)}\|_2^2$

$\therefore \underset{u^Tu=1}{\operatorname{argmax}} (x^{(i)T}u)^2 = \frac{x^{(i)}}{\|x^{(i)}\|_2}$

(b) $u^* = \underset{u^Tu=1}{\operatorname{argmax}} \frac{1}{m}\sum_{i=1}^{m} (x^{(i)T}u)^2$

$= \operatorname{argmin}\left(\frac{-1}{m}\sum_{i=1}^{m} (x^{(i)T}u)^2 + \beta(u^Tu-1)\right)$

Consider $g(u) = \frac{-1}{m}\sum_{i=1}^{m}(x^{(i)T}u)^2 + \beta(u^Tu-1)$

$$\frac{\partial g}{\partial u} = \frac{-2}{m}\sum_{i=1}^{m} x^{(i)}x^{(i)T}u + 2\beta u$$

u^* is at the minimum of g

$$\frac{\partial g}{\partial u} = 0 \Leftrightarrow \quad \beta u = \Sigma u.$$

$$\therefore u^* = \underset{u^Tu=1}{\operatorname{argmin}} \frac{1}{m}\left(\sum_{i=1}^{m}\|x^{(i)}\|_2^2 - 2(x^{(i)T}u)^2 + u^T(u^Tx^{(i)})^2u\right).$$

$$= \underset{u^Tu=1}{\operatorname{argmin}} \frac{1}{m}\sum_{i=1}^{m}\left((x^{(i)T}u)^2(u^Tu-2)\right)$$

$$= \underset{u^Tu=1}{\operatorname{argmax}} \frac{1}{m}\sum_{i=1}^{m}(x^{(i)T}u)^2$$

$$= \underset{u^Tu=1}{\operatorname{argmax}} \frac{1}{m}\sum_{i=1}^{m}\left[(x_i^Tu)\right]^T(x_i^Tu)$$

$$= \underset{u^Tu=1}{\operatorname{argmax}} \frac{1}{m}\sum_{i=1}^{m}u^Tx_ix_i^Tu$$

$$= \underset{u^Tu=1}{\operatorname{argmax}}\ u^T\frac{1}{m}\sum_{i=1}^{m}x_ix_i^T\ u.$$

$$= \underset{u^Tu=1}{\operatorname{argmax}}\ u^T\Sigma u$$

$\because \beta u = \Sigma u$ ①

$\therefore u^T \beta u = \beta = u^T \Sigma u$

$\therefore u^* = \underset{u^T u = 1}{\operatorname{argmax}} \ \beta$ ②

① ② $\Rightarrow u^*$ is the lagest eigenvec

4.3 (a)

$D_{ii} = (W\mathbb{1})_i$

$\therefore D = W\mathbb{1}$

$$
\begin{aligned}
(Df)_i &= D_{ii} f_i \\
&= (W\mathbb{1})_i f_i \\
&= \sum_{j=1}^{n} w_{ij} \begin{cases} \sqrt{\frac{vol(\bar{A})}{vol(A)}}, & V_i \in A \\ -\sqrt{\frac{vol(A)}{vol(\bar{A})}}, & V_i \notin A \end{cases}
\end{aligned}
$$

$$
\begin{aligned}
(Df)^T \mathbb{1} &= \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \left(\mathbb{1}\{v_i \in A\} \sqrt{\frac{vol(\bar{A})}{vol(A)}} - \mathbb{1}\{v_i \notin A\} \sqrt{\frac{vol(A)}{vol(\bar{A})}} \right) \\
&= \sqrt{\frac{vol(\bar{A})}{vol(A)}} \sum_{V_i \in A} \sum_{j=1}^{n} w_{ij} - \sqrt{\frac{vol(A)}{vol(\bar{A})}} \sum_{V_i \notin A} \sum_{j=1}^{n} w_{ij} \\
&= \sqrt{\frac{vol(\bar{A})}{vol(A)}} \, vol(A) - \sqrt{\frac{vol(A)}{vol(\bar{A})}} \, vol(\bar{A}) \\
&= 0
\end{aligned}
$$

$$f^T D f = \sum_{i=1}^{n} f_i (Df)_i$$

$$= \sum_{i=1}^{n} \left(1\{v_i \in A\} \frac{Vol(\bar{A})}{Vol(A)} + 1\{v_i \notin A\} \frac{Vol(A)}{Vol(\bar{A})} \right) \sum_{j=1}^{n} w_{ij})$$

$$= \frac{Vol(\bar{A})}{Vol(A)} \sum_{v_i \in A} \sum_{j=1}^{n} w_{ij} + \frac{Vol(A)}{Vol(\bar{A})} \sum_{v_i \notin A} \sum_{j=1}^{n} w_{ij})$$

$$= \frac{Vol(\bar{A})}{Vol(\bar{A})} Vol(\bar{A}) + \frac{Vol(A)}{Vol(\bar{A})} Vol(\bar{A})$$

$$= Vol(V)$$

$$L = D - W$$

$$f^T L f = f^T (D - W) f = f^T D f - f^T W f$$

$$= Vol(V) - \sum_{i=1}^{n} \sum_{j=1}^{n} f_i w_{ij} f_j$$

$$= Vol(V) - \left(\frac{Vol(\bar{A})}{Vol(A)} \sum_{v_i \in A} \sum_{v_j \in A} w_{ij} + \frac{Vol(A)}{Vol(\bar{A})} \sum_{v_j \in \bar{A}} \sum_{v_i \in \bar{A}} w_{ij} \right.$$

$$\left. - 2 \sum_{v_i \in A} \sum_{v_j \in \bar{A}} w_{ij} \right)$$

$$= Vol(V) - \left(\frac{Vol(\bar{A})}{Vol(A)} (Vol(A) - Cut(A, \bar{A})) \right.$$

$$+ \frac{Vol(A)}{Vol(\bar{A})} (Vol(\bar{A}) - Cut(A, \bar{A}))$$

$$\left. - 2 Cut(A, \bar{A}) \right)$$

$$= \mathrm{Vol}(V) - \Big(\mathrm{Vol}(\bar{A}) - \frac{\mathrm{Cut}(A,\bar{A})\,\mathrm{Vol}(\bar{A})}{\mathrm{Vol}(A)} + \mathrm{Vol}(A) - \frac{\mathrm{Cut}(A,\bar{A})\,\mathrm{Vol}(A)}{\mathrm{Vol}(\bar{A})} - 2\,\mathrm{Cut}(A,\bar{A}) \Big)$$

$$= \frac{\mathrm{Cut}(A,\bar{A})\,\mathrm{Vol}(\bar{A}) + \mathrm{Cut}(A,\bar{A})\,\mathrm{Vol}(A)}{\mathrm{Vol}(A)} + \frac{\mathrm{Cut}(A,\bar{A})\,\mathrm{Vol}(A) + \mathrm{Cut}(A,\bar{A})\,\mathrm{Vol}(\bar{A})}{\mathrm{Vol}(\bar{A})}$$

$$= \mathrm{Vol}(V) \left(\frac{\mathrm{Cut}(A,\bar{A})}{\mathrm{Vol}(A)} + \frac{\mathrm{Cut}(A,\bar{A})}{\mathrm{Vol}(\bar{A})} \right)$$

$$\therefore A^* = \operatorname*{argmin}_A f^T L f$$

$$= \operatorname*{argmin}_A \mathrm{Vol}(V) \left(\frac{\mathrm{Cut}(A,\bar{A})}{\mathrm{Vol}(A)} + \frac{\mathrm{Cut}(A,\bar{A})}{\mathrm{Vol}(\bar{A})} \right)$$

$$= \operatorname*{argmin}_A \left(\frac{\mathrm{Cut}(A,\bar{A})}{\mathrm{Vol}(A)} + \frac{\mathrm{Cut}(A,\bar{A})}{\mathrm{Vol}(\bar{A})} \right)$$

QED

(b) (2 points) Now we relax the above optimization problem such that $\boldsymbol{f} \in \mathbb{R}^n$. Please prove that the optimal f^* is the second eigenvector of $L_{rw} := D^{-1}L$, and that it is the generalized eigenvector: $Lf = \lambda Df$.

4.3 (b) 1° The question is unclear, it said to relax $f \in R^n$

then the problem become:

$$\min_{A} f^T L f \quad f \in \mathbb{R}$$

Since $f_i = \begin{cases} \sqrt{\frac{\mathrm{Vol}\,\bar{A}}{\mathrm{Vol}\,A}}, & V_i \in A \\ -\sqrt{\frac{\mathrm{Vol}(A)}{\mathrm{Vol}(\bar{A})}}, & V_i \in \bar{A} \end{cases}$, f is determined by A.

The only true variable is A. The domain of f should be $\{A \in \mathcal{A} \mid f \leftarrow A\}$

$\mathcal{A}$ has 2^n element, indictcating all possible A.

So the true constraint of f is 2^n discrete point scattered in $\mathbb{R}^n$. Since the domain is nonlenear space (discrete space), it's strange to do following prove.

2° Another way to explain the question is:

$$\min_{f \in R} f^T L f$$

But this is a Quddratic opt form, Since $L \succeq 0$, $f^* = 0$. This is also a strange situation.

3° This main information I want to convey here is that this question didn't make clear about what kind of relax, how much of the relax it wants to apply.

4° If we make no relaxation:

first take

$f^T D f - Vol(V) = 0$ into account

$$L(f) = f^T L f + u(f^T D f - Vol(V))$$

$$\frac{\partial L}{\partial f} = 2Lf + 2Dfu$$

Let $\frac{\partial L}{\partial f} = 0 \iff Lf = -uDf$

$\iff L_{rw} f = -uf$

Known that for normalized Larange L_{rw}, Smallest eigenvalue is 0, eigenvector is $\mathbb{1}$

$\because$ $D\mathbb{1}\mathbb{1} \neq 0$, unfeasible.

So $u > 0$, The second Smallest eigenvector will be the f^* for $\mathcal{L}(f)$.

Acknowledge Xue Cheng Chen