Problem 1: Normalized Cuts (15 points)

Answer:

Proof 1:

To prove that $L = \sum_{(i,j)\in E}(1_i - 1_j)(1_i - 1_j)^T$, can expand the expression. Let's consider a specific edge (i, j) in E. The corresponding entry in the adjacency matrix A is $A_{ij} = 1$ indicating that there is an edge between nodes i and j.

Now, let's look at the corresponding entry in the Laplacian matrix L. Using the definition of L, we have $L_{ij} = D_{ii} - A_{ij} = d_i - 1$, where d_i is the degree of node i. Similarly, $L_{ji} = d_j - 1$. Considering the expression $(1_i - 1)(1_i - 1_j)^T$, the entry at position (i, i) is 1, the entry at position (i, j) is -1, and all other entries are 0.

Since $L_{ij} = K_{ji} = -1$, we can see that the entry (i, j) in the Laplacian matrix L matches the entry at position (i, j) in the sum $\sum_{i,j)\in E}(1_i - 1_j)(1_i - 1_j)^T$ This holds for all edges in E, so the equality $L = \sum_{(i,j)\in E}(1_i - 1_j)(1_i - 1_j)^T$ is proven.

Proof 2:

To prove that $x^T Lx = \sum_{(i,j)\in E}(x_i - x_j)^2$, expand the expression for $x^T Lx$:

$x^T Lx = x^T(D - A)x = x^T Dx - x^T Ax$

Expanding further, we can write $x^T Dx$ as:

$x^T Dx = \sum_{i=1}^{n} x_i D_{ii} x_i = \sum_{i=1}^{n} x_i^2 d_i$

Similarly, we can write $x^T Ax$ as:

$x^T Ax = \sum_{(i,j)\in E} x_i A_{ij} x_j = \sum_{(i,j)\in E} x_i x_j$

Substituting these expressions back into $x^T Lx$, get:

$x^T Lx = \sum_{i=1}^{n} x_i^2 d_i - \sum_{(i,j)\in E} x_i x_j$

Now, can rewrite the summation over edges $(i, j) \in E$ as:

$\sum_{(i,j)\in E} x_i x_j = \sum_{(i,j)\in E}(x_i - x_j)^2 + 2\sum_{(i,j)\in E} x_i x_j$

Note that the second term on the right-hand side is the sum of all terms $x_i x_j$ for which the edge (i, j) appears twice in the summation. However, in a simple undirected graph, each edge appears only once, so this term cancels out.

Therefore, left with:

$x^T Lx = \sum_{i=1}^{n} x_i^2 d_i - \sum_{(i,j)\in E} x_i x_j = \sum_{i=1}^{n} x_i^2 d_i - \sum_{(i,j)\in E}(x_i - x_j)^2$

This equation shows that $x^T Lx$ is equal to $\sum_{(i,j)\in E}(x_i - x_j)^2$, which proves theproperty.

Proof 3:

To prove that $x^T Lx = c \cdot ncut(S)$ for some constant c, we need to rewrite the sum in terms of S and $\bar{S}$.

Recall that the assignment vector x is defined as:

$$x_i = \begin{cases} \sqrt{\dfrac{cut(\bar{S})}{vol(S)}}, & i \in S \\ \sqrt{\dfrac{cut(S)}{vol(\bar{S})}}, & i \in \bar{S} \end{cases}$$

Now, let's consider the expression $x^T Lx$:

$x^T Lx = \sum_{i=1}^n \sum_{j=1}^n x_i L_{ij} x_j$

Substituting the expression for L_{ij} from Property 1, we have:

$x^T Lx = \sum_{i=1}^n \sum_{j=1}^n x_i \left(\sum_{(k,l) \in E} (1_k - 1_l)(1_k - 1_l)^T \right) x_j$

Expanding the summation, get:

$x^T Lx = \sum_{(k,l) \in E} \sum_{i=1}^n \sum_{j=1}^n x_i (1_k - 1_l)(1_k - 1_l)^T x_j$

Proof 4:

To prove that $x^T D1 = 0$, where 1 is the vector of all ones, we need to show that the dot product of x and $D1$ is equal to zero.

Let's consider the dot product $x^T D1$:

$x^T D1 = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} \begin{bmatrix} D_{11} \; D_{22} \vdots D_{nn} \end{bmatrix}$

Expanding the dot product, get:

$x^T D1 = x_1 D_{11} + x_2 D_{22} + \ldots + x_n D_{nn}$

Now, recall that the diagonal elements of the degree matrix D are given by $D_{ii} = \sum_j A_{ij}$, where A is the adjacency matrix. Therefore, we can rewrite the equation as:

$x^T D1 = x_1 \left(\sum_j A_{1j} \right) + x_2 \left(\sum_j A_{2j} \right) + \ldots + x_n \left(\sum_j A_{nj} \right)$

Expanding further, we have:

$x^T D1 = \sum_j x_1 A_{1j} + \sum_j x_2 A_{2j} + \ldots + \sum_j x_n A_{nj}$

Now, note that each term $\sum_j x_i A_{ij}$ represents the sum of weights of edges incident to node i multiplied by the corresponding entry x_i. In other words, it represents the contribution of node i to the dot product. Since we are summing over all nodes, the sum of these contributions should be zero if the graph is undirected.

This is because for every edge (i, j) in the graph, there is a corresponding edge (j, i) with the same weight. Therefore, the contribution of node i to the dot product is canceled out by the contribution of node j, resulting in a total sum of zero.

Hence, I can conclude that $x^T D1 = 0$.

Proof 5:

To prove that $x^T Dx = 2m$, need to show that the dot product of x and Dx is equal to twice the total weight of the graph.

Let's consider the dot product $x^T Dx$:

$$x^T D x = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} \begin{bmatrix} D_{11} & D_{12} & \dots & D_{1n} \\ D_{21} & D_{22} & \dots & D_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ D_{n1} & D_{n2} & \dots & D_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

Expanding the dot product, get:

$x^T D x = x_1(D_{11}x_1 + D_{12}x_2 + \ldots + D_{1n}x_n) + x_2(D_{21}x_1 + D_{22}x_2 + \ldots + D_{2n}x_n) + \ldots + x_n(D_{n1}x_1 + D_{n2}x_2 + \ldots + D_{nn}x_n)$

Now, recall that the diagonal elements of the degree matrix D are given by $D_{ii} = \sum_j A_{ij}$, where A is the adjacency matrix. Therefore, can rewrite the equation as:

$x^T D x = x_1 \left(\sum_j A_{1j}x_j\right) + x_2 \left(\sum_j A_{2j}x_j\right) + \ldots + x_n \left(\sum_j A_{nj}x_j\right)$

Expanding further, have:

$x^T D x = \sum_j x_1 A_{1j} x_j + \sum_j x_2 A_{2j} x_j + \ldots + \sum_j x_n A_{nj} x_j$

Since the graph is undirected, each edge (i, j) contributes to the dot product twice: once through the term $x_i A_{ij} x_j$ and once through the term $x_j A_{ji} x_i$. Therefore, the total contribution of all edges in the graph is twice the dot product.

Furthermore, the sum of weights of all edges in the graph is equal to twice the total weight of the graph (since each edge is counted twice in an undirected graph). Therefore, we can conclude that $x^T D x$ is equal to twice the total weight of the graph, i.e., $x^T D x = 2m$.

Problem 2: Solution to Normalized Cut Minimization (15 points)

Answer:

To prove that the minimizer of Problem (6) is $D^{-1/2}v$, where v is the eigenvector corresponding to the second smallest eigenvalue of the normalized graph Laplacian $\tilde{L} = D^{-1/2} L D^{-1/2}$, have follow steps:

1. Make the substitution $z = D^{1/2}x$. This allows us to rewrite the problem in terms of z.

2. Substitute $x = D^{-1/2}z$ into the objective function $\frac{x^T L x}{x^T D x}$. get:

$$\frac{(D^{-1/2}z)^T L (D^{-1/2}z)}{D^{-1/2}z)^T D (D^{-1/2}z)}$$

3. Simplify the expression using the fact that $L = D - W$ (where W is the weighted adjacency matrix). Have:

$$\frac{z^T D^{-1/2} L D^{-1/2} z}{z^T D^{-1/2} D^{-1/2} z}$$

4. Substitute $\tilde{L} = D^{-1/2} L D^{-1/2}$ (normalized graph Laplacian) and $\tilde{z} = D^{-1/2}z$ to obtain:

$$\frac{\tilde{z}^T \tilde{L} \tilde{z}}{\tilde{z}^T \tilde{z}}$$

5. Note that the denominator $\tilde{z}^T\tilde{z}$ is a constant since $\tilde{z}$ is just a rescaled version of z. Therefore, minimizing this expression is equivalent to maximizing the numerator $\tilde{z}^T\tilde{L}\tilde{z}$.

6. Recall that the eigenvectors of $\tilde{L}$ are orthonormal and form a basis for $\mathbb{R}^n$. This means we can write any vector $\tilde{z}$ as a linear combination of the eigenvectors of $\tilde{L}$: $\tilde{z} = \sum_{i=1}^{n} a_i v_i$, where v_i are the eigenvectors of $\tilde{L}$.

7. Substitute $\tilde{z} = \sum_{i=1}^{n} a_i v_i$ into the numerator $\tilde{z}^T\tilde{L}\tilde{z}$:

$$\tilde{z}^T\tilde{L}\tilde{z} = (\sum_{i=1}^{n} a_i v_i)\tilde{L}(\sum_{j=1}^{n} a_j v_j)$$

8. Apply the properties of matrix transpose and distributivity to expand the expression:

$$\tilde{z}^T\tilde{L}\tilde{z} = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j v_i^T\tilde{T}v_j$$

9. Recall that the eigenvectors v_i are orthonormal, which means $v_i^T v_j = \delta_{ij}$ (Kronecker delta). Using this property, the expression simplifies to:

$$\tilde{z}^T\tilde{L}\tilde{z} = \sum_{i=1}^{n} a_i^2 v_i^T\tilde{L}v_i$$

10. Notice that $v_i^T\tilde{L}v_i$ corresponds to the i-th eigenvalue of $\tilde{L}$. Let's denote the eigenvalues of $\tilde{L}$ as $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$. Then have:

$$\tilde{z}^T\tilde{L}\tilde{z} = \sum_{i=1}^{n} a_i^2\lambda_i$$

11. Since we want to maximize $\tilde{z}^T\tilde{L}\tilde{z}$, we want to assign larger weights a_i^2 to the eigenvectors corresponding to smaller eigenvalues λ_i. Therefore, to maximize the expression, we choose $a_i = 0$ for $i \neq 2$ and $a_2 = 1$.

12. Finally, substitute $\tilde{z} = \sum_{i=1}^{n} a_i v_i$ back into the expression for $\tilde{z}^T\tilde{L}\tilde{z}$:

$$\tilde{z}^T\tilde{L}\tilde{z} = (\sum_{i=1}^{n} a_i v_i)^T\tilde{L}(\sum_{j=1}^{n} a_j v_j) = (\sum_{i=1}^{n} a_i v_i^T)\tilde{L}(\sum_{j=1}^{n} a_j v_j)$$

Since the eigenvectors v_i are orthonormal, we have $v_i^T v_j = \delta_{ij}$ (Kronecker delta). Therefore, the expression simplifies to:

$$\tilde{z}^T\tilde{L}\tilde{z} = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j v_i^T\tilde{T}v_j = \sum_{i=1}^{n} a_i^2 v_i^T\tilde{L}v_i$$

The term $v_i^T\tilde{L}v_i$ corresponds to the i-th eigenvalue of $\tilde{L}$, denoted as λ_i.

13. Let's denote the eigenvalues of $\tilde{L}$ as $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$. have:

$$\tilde{z}^T\tilde{L}\tilde{z} = \sum_{i=1}^{n} a_i^2\lambda_i$$

14. Since we want to maximize $\tilde{z}^T \tilde{L} \tilde{z}$, we want to assign larger weights a_i^2 to the eigenvectors corresponding to smaller eigenvalues λ_i. Therefore, to maximize the expression, we choose $a_i = 0$ for $i \neq 2$ and $a_2 = 1$.

15. Finally, substitute $\tilde{z} = \sum_{i=1}^{n} a_i v_i$ back into the expression: $\tilde{z} = a_2 v_2 = v_2$ (since $a_2 = 1$).

16. Recall that $\tilde{z} = D^{-1/2} z$ and $z = D^{1/2} x$. Therefore, we have:

$$\begin{aligned}
D^{-1/2} z &= v_2 \\
D^{-1/2} D^{1/2} x &= v_2 \\
D^{-1/2} D^{1/2} D^{1/2} x &= v_2 \\
D^{-1/2} D D^{1/2} D^{1/2} x &= v_2 \\
D^{-1/2} D x D &= v_2 \\
D^{-1/2} D^2 x D &= v_2 \\
D^{-1/2} (D^2 x) D &= v_2 \\
(D^{-1/2} D^2 x) D &= v_2
\end{aligned}$$

17. Recall that $x^T D x = 2m$, where m is the number of edges in the graph. Therefore, $D^2 x = 2mDx$. Substituting this into the previous expression, have:

$$\begin{aligned}
(D^{-1/2}(2mDx))D &= v_2 \\
(2mD^{1/2}x)D &= v_2 \\
2mDxD &= v_2
\end{aligned}$$

18. Finally, since $x^T D x = 2m$, we can divide both sides by $2m$ to obtain:

$$D^{-1/2} x = v_2$$

19. Therefore, the minimizer of Problem (6) is $D^{-1/2} v_2$, where v_2 is the eigenvector corresponding to the second smallest eigenvalue of the normalized graph Laplacian $\tilde{L} = D^{-1/2} L D^{-1/2}$.

This completes the proof.