Complex Networks
Solutions 6

---

1. (a) If $(i,j) \notin E$, then $a_{ij} = 0$ by definition. We have

$$\begin{aligned}\mathbf{x}^T L_G \mathbf{x} &= \sum_{i,j \in V} x_i L_G(i,j) x_j = \sum_{i \in V} d_i x_i^2 - \sum_{i \neq j} a_{ij} x_i x_j \\ &= \sum_{i \in V} x_i^2 \sum_{j \in V} a_{ij} - 2 \sum_{(i,j) \in E} a_{ij} x_i x_j. \qquad (1)\end{aligned}$$

We have used the fact that $a_{ij} = 0$ for $(i,j) \notin E$ and the definition of d_i to obtain the last equality. The factor of 2 in front of the last sum comes from the fact that, when summing over edges, we don't count (i,j) and (j,i) as two different edges, but in summing over pairs of vertices, we have two terms, $a_{ij} x_i x_j$ and $a_{ji} x_j x_i$, which are equal since $a_{ij} = a_{ji}$.
On the other hand, we have

$$\begin{aligned}\sum_{(i,j) \in E} a_{ij}(x_i - x_j)^2 &= \sum_{(i,j) \in E} a_{ij}(x_i^2 - 2x_i x_j + x_j^2) \\ &= \sum_{i \in V} x_i^2 \sum_{j:(i,j) \in E} a_{ij} - 2 \sum_{(i,j) \in E} a_{ij} x_i x_j. \qquad (2)\end{aligned}$$

The second equality holds because for each $i \in V$, there is an $a_{ij} x_i^2$ term in the sum above corresponding to each $(i,j) \in E$. Comparing (1) and (2), we see that they are identical since $a_{ij} = 0$ whenever $(i,j) \notin E$.
For all $x \in \mathbb{R}^n$, $x^2 L_G x$ is a weighted sum of squares with positive weights, and hence is non-negative. Thus, L_G is positive semi-definite.

(b) We have

$$L_G = \begin{pmatrix} 3 & -1 & -2 \\ -1 & 2 & -1 \\ -2 & -1 & 3 \end{pmatrix} \text{ and } \mathbf{x}^T L_G \mathbf{x} = 3x_1^2 + 2x_2^2 + 3x_3^2 - 2x_1x_2 - 4x_1x_3 - 2x_2x_3,$$

while

$$\sum_{(i,j) \in E} a_{ij}(x_i - x_j)^2 = (x_1 - x_2)^2 + (x_2 - x_3)^2 + 2(x_3 - x_1)^2.$$

Clearly, the two right-hand sides are equal.

2. (a) We have $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i \in S} \pi_i x_i y_i^*$ and $\langle \mathbf{y}, \mathbf{x} \rangle = \sum_{i \in S} \pi_i y_i x_i^*$ by definition. As π_i, being a probability, is real for all i, it follows that $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle^*$. Likewise, linearity is easy to check. Finally, $\langle \mathbf{x}, \mathbf{x} \rangle = \sum_{i \in S} \pi_i x_i x_i^* = \sum_{i \in S} \pi_i |x_i|^2$. Clearly, this quantity is non-negative, as it is a weighted sum of squared absolute values with positive weights. Moreover, if $\mathbf{x}$ is non-zero, then at least one of the $|x_i|$ is strictly positive. Also, all the π_i are strictly positive, for the reason stated in the question. Hence, $\sum_{i \in S} \pi_i |x_i|^2$ is strictly positive.

(b) Note that $(Q\mathbf{x})_i = \sum_{j \in S} q_{ij} x_j$. Hence,

$$\langle Q\mathbf{x}, \mathbf{y} \rangle = \sum_{i,j \in S} \pi_i q_{ij} x_j y_i^* = \sum_{i,j \in S} \pi_j q_{ji} x_j y_i^*,$$

where we have used reversibility to get the last equality. On the other hand,

$$\langle \mathbf{x}, Q\mathbf{y}\rangle = \sum_{i,j\in S} \pi_i x_i (q_{ij} y_j)^* = \sum_{i,j\in S} \pi_j x_j q_{ji}^* y_i^*.$$

But all the q_{ji} are real (rates are obtained from probabilities by taking differences and then a limit), so $q_{ji}^* = q_{ji}$, and we see that the above expressions are equal.

(c) Let λ be an eigenvalue of Q and $\mathbf{x}$ a corresponding eigenvector. Then,

$$\langle \mathbf{x}, Q\mathbf{x}\rangle = \langle \mathbf{x}, \lambda\mathbf{x}\rangle = \lambda^* \langle \mathbf{x}, \mathbf{x}\rangle,$$

whereas

$$\langle Q\mathbf{x}, \mathbf{x}\rangle = \langle \lambda\mathbf{x}, \mathbf{x}\rangle = \lambda \langle \mathbf{x}, \mathbf{x}\rangle.$$

As these two quantities must be equal by part (b), and as $\langle \mathbf{x}, \mathbf{x}\rangle > 0$ by part (a), it follows that $\lambda = \lambda^*$, i.e., that λ is real.

3. (a) Let S be any subset of Ω. Now, $\sum_{i\in S} p_i + \sum_{i\in S^c} p_i = 1$, and the same for q_i, because $\mathbf{p}$ and $\mathbf{q}$ are probability distributions. Hence

$$|p(S) - q(S)| = \Big| \sum_{i\in S} p_i - \sum_{i\in S} q_i \Big| = \Big| \sum_{i\in S^c} p_i - \sum_{i\in S^c} q_i \Big| . \tag{3}$$

Morevoer,

$$\Big| \sum_{i\in S} p_i - \sum_{i\in S} q_i \Big| \le \sum_{i\in S} |p_i - q_i|, \tag{4}$$

by the triangle inequality. Hence, it follows from (3) that

$$|p(S) - q(S)| \le \frac{1}{2}\Big(\sum_{i\in S} |p_i - q_i| + \sum_{i\in S^c} |p_i - q_i| \Big) = \frac{1}{2}\sum_{i=1}^n |p_i - q_i|. \tag{5}$$

Since this inequality holds for every subset S, it also holds for the maximum over all subsets, and so $d_{TV}(\mathbf{p}, \mathbf{q}) \le \frac{1}{2}\sum_{i=1}^n |p_i - q_i|$.

In order to show that equality holds, we need to find a subset S for which equality holds in (4), as equality will then hold in (5) as well. Take

$$S = \{i : p_i \ge q_i\}, \quad S^c = \{i : p_i < q_i\}.$$

It is easy to verify that equality holds in (4) and (5). Since we have found a specific set S for which

$$|p(S) - q(S)| = \sum_{i\in S} p_i - q_i = \sum_{i\in S^c} q_i - p_i = \frac{1}{2}\sum_{i=1}^n |p_i - q_i|,$$

we have $d_{TV}(\mathbf{p}, \mathbf{q}) \ge \frac{1}{2}\sum_{i=1}^n |p_i - q_i|$. Combining this with the reverse inequality shown earlier yields $d_{TV}(\mathbf{p}, \mathbf{q}) = \frac{1}{2}\|\mathbf{p} - \mathbf{q}\|_1$.

Since $0 \le p(S) \le 1$ and $0 \le q(S) \le 1$, $|p(S) - q(S)| \le 1$ for all subsets S, and hence also for the maximum over all subsets. In other words, $d_{TV}(\mathbf{p}, \mathbf{q}) \le 1$ for any two probability distribution $\mathbf{p}$ and $\mathbf{q}$. Hence $\|\mathbf{p} - \mathbf{q}\|_1 \le 2$. For any example where equality holds, take $n = 2$, $\mathbf{p} = (1\,0)$ and $\mathbf{q} = (0\,1)$.

(b) We have

$$\|\mathbf{p}-\mathbf{q}\|_1 = \sum_{i=1}^{n} |p_i - q_i| = (\mathbf{p}-\mathbf{q}) \cdot \mathbf{sgn}(\mathbf{p}-\mathbf{q}),$$

where the $\mathbf{sgn}$ function is defined for $x \in \mathbb{R}$ by $\mathbf{sgn}(x) = +1$ if x is positive, -1 if x is negative, and zero if x is zero. We define the $\mathbf{sgn}$ function on a vector to be applied componentwise. Finally $\mathbf{x} \cdot \mathbf{y}$ denotes the dot product or inner product of two vectors $\mathbf{x}$ and $\mathbf{y}$.

Hence, by the Cauchy-Schwarz inequality,

$$\|\mathbf{p}-\mathbf{q}\|_1 \leq \|\mathbf{p}-\mathbf{q}\|_2 \|\mathbf{sgn}(\mathbf{p}-\mathbf{q})\|_2. \tag{6}$$

But

$$\|\mathbf{sgn}(\mathbf{p}-\mathbf{q})\|_2 = \sqrt{\sum_{i=1}^{n} (\mathbf{sgn}(p_i - q_i))^2} \leq \sqrt{n},$$

since every element of the vector $\mathbf{sgn}(\mathbf{p}-\mathbf{q})$ is either $+1$ or -1 or 0. Substituting this in (6) gives $\|\mathbf{p}-\mathbf{q}\|_1 \leq \sqrt{n}\|\mathbf{p}-\mathbf{q}\|_2\|$, as we are asked to show.

(c) We have

$$\|\mathbf{p}-\mathbf{q}\|_2^2 = \sum_{i=1}^{n} (p_i - q_i)^2 \leq \sum_{i=1}^{n} (p_i^2 + q_i^2),$$

because the $p_i q_i$ terms are all non-negative since probabilities are non-negative. But $p_i^2 \leq p_i$ since $0 \leq p_i \leq 1$, and similarly for q_i. Hence,

$$\sum_{i=1}^{n} (p_i^2 + q_i^2) \leq \sum_{i=1}^{n} (p_i + q_i) = 2.$$

Therefore, $\|\mathbf{p}-\mathbf{q}\|_2^2 \leq 2$, which is what we are asked to show.

Equality holds for the same example as in the answer to the first part.

4. Using the fact that for any two probability distributions $\mathbf{p}$ and $\mathbf{q}$, the total variation distance between them is half the L^1-norm of $\mathbf{p}-\mathbf{q}$, as shown in Question 4(a), we get:

(a)

$$d_{TV} = \frac{1}{2}\Big(\Big| \frac{1}{4} - \frac{1}{3} \Big| + \Big| \frac{1}{2} - \frac{1}{3} \Big| + \Big| \frac{1}{4} - \frac{1}{3} \Big| \Big) = \frac{1}{6}.$$

(b)

$$d_{TV} = \frac{1}{2}\Big(\Big| \frac{1}{4} - e^{-1} \Big| + \Big| \frac{1}{2} - e^{-1} \Big| + \Big| \frac{1}{4} - \frac{1}{2}e^{-1} \Big| + \sum_{k=3}^{\infty} \frac{1}{k!} e^{-1} \Big).$$

(c) The idea is exactly the same here, but finding the L^1-norm involves integration rather than summation as we are working with continuous random variables. So we have

$$\begin{aligned} 2d_{TV} &= \int_0^1 \big| \, e^{-x} - 1 \, \big| \, dx + \int_1^{\infty} \big| \, e^{-x} - 0 \, \big| \, dx \\ &= \int_0^1 (1 - e^{-x}) dx + \int_1^{\infty} e^{-x} dx = 2e^{-1}. \end{aligned}$$

(d) As for the last part,

$$\begin{aligned} 2d_{TV} &= \int_0^{\infty} \big| \, e^{-x} - 2e^{-2x} | dx \\ &= \int_0^{\log 2} \big(2e^{-2x} - e^{-x}\big) dx + \int_{\log 2}^{\infty} \big(e^{-x} - 2e^{-2x}\big) dx = \frac{1}{2}. \end{aligned}$$

5. (a) The rate for moving from v to w is 1 if $(v,w) \in E$ and 0 otherwise. Hence, the transition rates are $q_{vw} = 1((v,w) \in E$ for $v \neq w$. Moreover, $q_{vv} = -\sum_{w \neq v} q_{vw} = -d_v$, where d_v denotes the degree of v. Thus, the transition rate matrix for the star graph is

$$Q = \begin{pmatrix} -(n-1) & 1 & 1 & \dots & 1 \\ 1 & -1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \dots & -1 \end{pmatrix}$$

This is the negative of the Laplacian matrix.

(b) All row sums of Q are zero, and Q is symmetric, so all column sums are also zero. Therefore, the all-1 row vector is an eigenvector with eigenvalue zero. Consequently,

$$\frac{1}{n}\mathbf{1}^T Q = 0.$$

This says that $\frac{1}{n}\mathbf{1}^T$ is a probability vector which solves the global balance equations. Hence, it is an invariant distribution of the Markov chain. It is unique because the graph is connected, which implies that the Markov chain is irreducible (it is possible to go from any state/node to any other).

(c) Let S be a subset of V consisting of k leaves. Then, $|E(S,S^c)| = k$, and so

$$\frac{|E(S,S^c)|}{\frac{1}{n}|S| \cdot |S^c|} = \frac{k}{\frac{1}{n}k(n-k)} = \frac{n}{n-k}.$$

Minimising this over all subsets S, equivalently, over all k between 1 and $n-1$, gives $\Phi(G) = \frac{n}{n-1}$.

(d) Cheeger's inequality says that $\Phi(G) \leq \sqrt{8 d_{\max} \lambda_2}$, where $d_{\max}$ is the maximum node degree, which is $n-1$ for the star graph. Hence,

$$\lambda_2 \geq \frac{\Phi(G)^2}{8 d_{\max}} = \frac{n^2}{8(n-1)^3}.$$

(e) If $\sqrt{n} e^{-\lambda_2 t}$ is smaller than ϵ, then so is $d_{TV}(p(t), \pi)$. So we want to choose t such that

$$e^{\lambda_2 t} \geq \frac{\sqrt{n}}{\epsilon}, \text{ i.e., } t \geq \frac{1}{\lambda_2}\Big(\frac{1}{2}\log n + \log\frac{1}{\epsilon}\Big).$$

Using the bound for λ_2 from the last part, this is ensured if we choose

$$t \geq \frac{8(n-1)^3}{n^2}\Big(\frac{1}{2}\log n + \log\frac{1}{\epsilon}\Big).$$