7.8* Show or confirm that the eigenvalues and eigenvectors of the P matrix of equation (7.42), as defined in equation (7.36), are given as

$$\Lambda = \begin{bmatrix} 1 & & & & & \\ & 0 & & & & \\ & & 0 & & & \\ & & & \ddots & & \\ & & & & 0 & \\ & & & & & 1-\frac{1}{\pi_1} \end{bmatrix},$$

$$E = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & \left(1-\frac{1}{\pi_1}\right)a_K \\ 1 & a_2 & a_3 & \cdots & a_{K-1} & a_K \\ 1 & -\frac{\pi_2}{\pi_3}a_2 & 0 & \cdots & 0 & a_K \\ 1 & 0 & -\frac{\pi_2}{\pi_4}a_3 & \cdots & 0 & a_K \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 0 & 0 & \cdots & -\frac{\pi_2}{\pi_K}a_{K-1} & a_K \end{bmatrix}, \tag{7.105}$$

where the factors $a_k = \left[\frac{\pi_{k+1}}{\pi_2(\pi_2+\pi_{k+1})}\right]^{\frac{1}{2}}, k = 2, \ldots, K-1$, and $a_K = \left[\frac{\pi_1}{1-\pi_1}\right]^{\frac{1}{2}}$ are for normalizing the eigenvectors so that $E^T BE = I$ or $\sum_{i=1}^{K} \pi_i e_{ik}^2 = 1$ for each k. Note that the kth column in E is the eigenvector corresponding to λ_k.

[Hint. (**a**) To confirm the results, simply check that Λ and E above satisfy $Px = \lambda x$, with x to be a column in E. (**b**) To derive the eigenvalues, one way is to solve the characteristic equation $|P-\lambda I| = 0$. By Laplace's formula, the determinant $|P-\lambda I| = p_{11} \cdot |P_{11}| - p_{12} \cdot |P_{12}| + p_{13} \cdot |P_{13}| - \ldots$, where P_{1k} is the $(K-1) \times (K-1)$ matrix that results from removing the 1st row and kth column of $P-\lambda I$. Thus show that $|P-\lambda I| = \frac{1}{\pi_1}(-\lambda)^{K-2}(\lambda-1)(\pi_1\lambda + 1 - \pi_1)$. Note that the determinant of a triangular matrix is the product of the diagonal elements and that interchanging two rows of a matrix multiplies its determinant by –1.]

Solution.

We actually have reservations about part of this problem. The reason is because E does not meet the criterion $E^T BE = I$ as given in the problem. As detailed below, denote

$$E^T = [e_1 \quad \cdots \quad e_K],$$

where $e_1 = [e_{11} \quad \cdots \quad e_{1K}]^T, e_1 = [e_{21} \quad \cdots \quad e_{1K}]^T, \ldots, e_K = [e_{11} \quad \cdots \quad e_{1K}]^T$.

According to the problem statement, we have

$$E^T BE = [e_1 \quad \cdots \quad e_K] \times \begin{bmatrix} \pi_1 & 0 & \cdots & 0 \\ 0 & \pi_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \pi_K \end{bmatrix} \times \begin{bmatrix} e_1^T \\ \vdots \\ e_K^T \end{bmatrix}$$

$$= \pi_1 e_1 e_1^T + \cdots + \pi_K e_K e_K^T$$

$$= \begin{bmatrix} \sum_{i=1}^{K} \pi_i e_{i1}^2 & \sum_{i=1}^{K} \pi_i e_{i2} e_{i1} & \cdots & \sum_{i=1}^{K} \pi_i e_{iK} e_{i1} \\ \sum_{i=1}^{K} \pi_i e_{i1} e_{i2} & \sum_{i=1}^{K} \pi_i e_{i2}^2 & \cdots & \sum_{i=1}^{K} \pi_i e_{iK} e_{i2} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{K} \pi_i e_{i1} e_{iK} & \sum_{i=1}^{K} \pi_i e_{i2} e_{iK} & \cdots & \sum_{i=1}^{K} \pi_K e_{iK}^2 \end{bmatrix}$$

$$= I.$$

Hence, we have

$$\sum_{i=1}^{K} \pi_i e_{ik}^2 = 1, \text{for } k = 1, \dots, K,$$

$$\sum_{i=1}^{K} \pi_i e_{im} e_{in} = 0, \text{for } m \neq n, 1 \leq m, n \leq K.$$

Unfortunately, E as given in the problem does not seem to meet the second criterion above (readers can easily verify themselves).

Nevertheless, as follows, we still follow the “answer” given in the problem to present a solution but readers may want to be aware of the above. Interestingly, this does not affect the calculation of Problem 7.9.

a)

Refer to Eq. 7.42 in (Yang 2014a) as follows

$$P = \begin{bmatrix} 1 - \frac{1-\pi_1}{\pi_1} & \frac{\pi_2}{\pi_1} & \frac{\pi_3}{\pi_1} & \cdots & \frac{\pi_K}{\pi_1} \\ 1 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \cdots & 0 \end{bmatrix}.$$

To find the eigenvalue λ, set $|P - \lambda I| = 0$. Assume $\lambda \neq 0$, it follows that

$|P - \lambda I|$

$$= \begin{vmatrix} 1 - \frac{1-\pi_1}{\pi_1} - \lambda & \frac{\pi_2}{\pi_1} & \frac{\pi_3}{\pi_1} & \cdots & \frac{\pi_K}{\pi_1} \\ 1 & -\lambda & 0 & \cdots & 0 \\ 1 & 0 & -\lambda & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \cdots & -\lambda \end{vmatrix}$$

$$\underline{\underline{\pi_1 r_1}} \frac{1}{\pi_1} \begin{vmatrix} (2-\lambda)\pi_1 - 1 & \pi_2 & \pi_3 & \cdots & \pi_K \\ 1 & -\lambda & 0 & \cdots & 0 \\ 1 & 0 & -\lambda & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \cdots & -\lambda \end{vmatrix}$$

$$\xlongequal{\substack{\frac{1}{\lambda}c_2 \\ \vdots \\ \frac{1}{\lambda}c_{K-1}}} \frac{\lambda^{K-1}}{\pi_1}\begin{vmatrix} (2-\lambda)\pi_1 - 1 & \frac{\pi_2}{\lambda} & \frac{\pi_3}{\lambda} & \cdots & \frac{\pi_K}{\lambda} \\ 1 & -1 & 0 & \cdots & 0 \\ 1 & 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \cdots & -1 \end{vmatrix}$$

$$\xlongequal{\substack{c_1 + c_2 \\ \vdots \\ c_1 + c_K}} \frac{\lambda^{K-1}}{\pi_1}\begin{vmatrix} (2-\lambda)\pi_1 - 1 + \sum_{i=2}^{K} \frac{\pi_i}{\lambda} & \frac{\pi_2}{\lambda} & \frac{\pi_3}{\lambda} & \cdots & \frac{\pi_K}{\lambda} \\ 0 & -1 & 0 & \cdots & 0 \\ 0 & 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -1 \end{vmatrix}$$

$$= (-1)^{K-1}\frac{\lambda^{K-1}}{\pi_1}\left((2-\lambda)\pi_1 - 1 + \sum_{i=2}^{K}\frac{\pi_i}{\lambda}\right)$$

$$= (-1)^{K-1} \times \frac{\lambda^{K-1}}{\pi_1} \times \left(\lambda(2-\lambda)\pi_1 - \lambda + (1-\pi_1)\right)$$

$$= \frac{1}{\pi}(-1)^{\mathrm{K}-1}(\lambda - 0)^{K-2}(\lambda - \lambda_1)(\lambda - \lambda_K),$$

where λ_1 and λ_K are the two roots of the unary quadratic equation $\lambda(2-\lambda)\pi_1 - \lambda + (1-\pi_1) = 0$. By solving it, we have

$$\begin{aligned} \lambda &= \frac{-(1-2\pi_1) \pm \sqrt{(1-2\pi_1)^2 - 4\pi_1(\pi_1 - 1)}}{2\pi_1} \\ &= \frac{-(1-2\pi_1) \pm 1}{2\pi_1} \\ &= 1 \text{ or } \frac{\pi_1 - 1}{\pi_1}. \end{aligned}$$

Note that in Section 7.3.2.1 of (Yang 2014a) it is indicated that $1 = \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_K \geq -1$, therefore we have $\lambda_1 = 1$, $\lambda_K = \frac{\pi_1 - 1}{\pi_1}$, $\lambda_2 = \lambda_3 = \cdots = \lambda_{K-1} = 0$. Accordingly,

$$\Lambda = \begin{bmatrix} 1 & & & & & \\ & 0 & & & & \\ & & 0 & & & \\ & & & \ddots & & \\ & & & & 0 & \\ & & & & & \frac{\pi_1 - 1}{\pi_1} \end{bmatrix}.$$

b)

As given in the problem, the eigenvectors should be normalized so that the following holds for each k

$$\sum_{i=1}^{K} \pi_i e_{ik}^2 = 1.$$

Denote $e_k = (e_{1k}, e_{2k}, \dots, e_{Kk})^T$ as the k^{th} eigenvector of the matrix E. In other words, $E = (e_1, e_2, \dots, e_K)$. Solve the normalizing eigenvector(s) for each eigenvalue as follows.

i) For $k = 1$, we have $\lambda_k = \lambda_1 = 1$. The eigenvector can be found by solving $(P - \lambda_1)e_1 = \mathbf{0}$. Hence,

$$\begin{bmatrix} 1 - \frac{1-\pi_1}{\pi_1} - 1 & \frac{\pi_2}{\pi_1} & \frac{\pi_3}{\pi_1} & \cdots & \frac{\pi_K}{\pi_1} \\ 1 & -1 & 0 & \cdots & 0 \\ 1 & 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \cdots & -1 \end{bmatrix} e_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

By solving the above, we get $e_1 = c(1,1,\dots,1)^T$ where c is a constant. Due to the constraint $\sum_{i=1}^{K} \pi_i e_{i1}^2 = 1$, it is easy to see $c^2(\pi_1 + \pi_2 + \cdots + \pi_K) = 1$, thus $c = 1$ and $e_1 = (1,1,\dots,1)^T$.

For $k = K$, we have $\lambda_k = \lambda_K = \frac{1-\pi_1}{\pi_1}$. It follows that

$$(P - \lambda_K)e_K = \begin{bmatrix} 1 - \frac{1-\pi_1}{\pi_1} - \frac{\pi_1 - 1}{\pi_1} & \frac{\pi_2}{\pi_1} & \frac{\pi_3}{\pi_1} & \cdots & \frac{\pi_K}{\pi_1} \\ 1 & \frac{1-\pi_1}{\pi_1} & 0 & \cdots & 0 \\ 1 & 0 & \frac{1-\pi_1}{\pi_1} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \cdots & \frac{1-\pi_1}{\pi_1} \end{bmatrix} e_K = \mathbf{0}.$$

It can be further shown that

$$\begin{cases} \pi_1 e_{1K} + \sum_{i \geq 2}^{K} \pi_i e_{iK} = 0, \\ e_{1K} = \left(\frac{1-\pi_1}{\pi_1}\right) e_{2K}, \\ e_{1K} = \left(\frac{1-\pi_1}{\pi_1}\right) e_{3K}, \\ \vdots \\ e_{1K} = \left(\frac{1-\pi_1}{\pi_1}\right) e_{KK}. \end{cases}$$

This is equivalent to saying

$$\frac{e_{1K}}{1 - \frac{1}{\pi_1}} = e_{2K} = \cdots = e_{KK}.$$

Applying the constraint $\sum_{i=1}^{K} \pi_i e_{iK}^2 = 1$, it can be shown that

$$\pi_1 \left(\left(1 - \frac{1}{\pi_1}\right) e_{KK} \right)^2 + \pi_2 (e_{KK})^2 + \pi_3 (e_{KK})^2 + \cdots + \pi_K (e_{KK})^2 = \frac{e_{KK}^2 (1 - \pi_1)}{\pi_1} = 1.$$

Hence

$$e_{iK} = \begin{cases} -\dfrac{\sqrt{1-\pi_1}}{\pi_1}, i = 1 \\ \sqrt{\dfrac{\pi_1}{1-\pi_1}}, i \neq 1 \end{cases}.$$

Further, define $a_K = \sqrt{\frac{\pi_1}{1-\pi_1}}$, such that e_K can be rewritten as

$$e_K = \left(\frac{\pi_1 - 1}{\pi_1} a_K, a_K, \dots, a_K\right)^T$$

where $a_K = \sqrt{\frac{\pi_1}{1-\pi_1}}$.

For $2 \leq k \leq K-1$, $\lambda_k = 0$. It follows that

$$(P - \lambda_k)e_k = \begin{bmatrix} 1 - \frac{1-\pi_1}{\pi_1} & \frac{\pi_2}{\pi_1} & \frac{\pi_3}{\pi_1} & \cdots & \frac{\pi_K}{\pi_1} \\ 1 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \cdots & 0 \end{bmatrix} e_k = \mathbf{0}.$$

It can be shown that

$$\begin{cases} e_{1k} = 0, \\ \pi_2 e_{2k} + \pi_3 e_{3k} + \cdots \pi_K e_{Kk} = 0. \end{cases}$$

By solving the above, we get

$$e_{2k} = \begin{bmatrix} 0 \\ 1 \\ -\pi_2/\pi_3 \\ 0 \\ \vdots \\ 0 \end{bmatrix} a_2, e_{3k} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ -\pi_2/\pi_4 \\ 0 \\ \vdots \\ 0 \end{bmatrix} a_3, \dots, e_{K-1,k} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ -\pi_2/\pi_K \end{bmatrix} a_{K-1},$$

where a_k is a normalizing constant. Applying the constraint $\sum_{i=1}^{K} \pi_i e_{ik}^2 = 1$, it can be calculated that $\left(\pi_2 + \pi_{k+1} \times \frac{\pi_2^2}{\pi_{k+1}^2}\right) a_k^2 = 1$, thus $a_k = \sqrt{\frac{\pi_{k+1}}{\pi_2 \pi_{k+1} + \pi_2^2}}$ for $2 \leq k \leq K-1$.

By integrating the above results, finally we obtain

$$E = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & \left(1 - \dfrac{1}{\pi_1}\right) a_K \\ 1 & a_2 & a_3 & \cdots & a_{K-1} & a_K \\ 1 & -\dfrac{\pi_2}{\pi_3} a_2 & 0 & \cdots & 0 & a_K \\ 1 & 0 & -\dfrac{\pi_2}{\pi_4} a_3 & \cdots & 0 & a_K \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 0 & 0 & \cdots & -\dfrac{\pi_2}{\pi_K} a_{K-1} & a_K \end{bmatrix},$$

where $a_k = \sqrt{\frac{\pi_{k+1}}{\pi_2 \pi_{k+1} + \pi_2^2}}$ for $2 \leq k \leq K - 1$ and $a_K = \sqrt{\frac{\pi_1}{1-\pi_1}}$.