20170058 Keonwoo Kim · June 23, 2020
MAS109EF Introduction to Linear Algebra

Homework

1. (a) Observe that there is a solution $\mathbf{x} \in \mathrm{row}(A)$ to $A\mathbf{x} = \mathbf{b}$ iff $\mathbf{b} \in \mathrm{col}(AA^T)$. This is because ($R_i$: i-th row of A as a column vector)

$$\begin{aligned} \mathbf{x} \in \mathrm{row}(A) &\iff \mathbf{x} = a_1 R_1 + \cdots + a_m R_m = A^T [a_1 \ \cdots \ a_m]^T \text{ for some } a_i\text{'s} \\ &\iff \mathbf{x} \in \mathrm{col}(A^T) \end{aligned}$$

so that

$$\exists \mathbf{x} \in \mathrm{row}(A),\ A\mathbf{x} = \mathbf{b} \iff \exists \mathbf{y} \in \mathbb{R}^m\colon AA^T\mathbf{y} = \mathbf{b} \iff \mathbf{b} \in \mathrm{col}(AA^T).$$

Since $\mathrm{col}(AA^T) \subseteq \mathrm{col}(A)$ and the dimensions of $\mathrm{col}(AA^T)$ and $\mathrm{col}(A)$, which are $\mathrm{rk}(AA^T)$ and $\mathrm{rk}(A)$, are the same, we have $\mathrm{col}(AA^T) = \mathrm{col}(A)$. Note that $A\mathbf{x} = \mathbf{b}$ is consistent so that $\mathbf{b} \in \mathrm{col}(A) = \mathrm{col}(AA^T)$. Therefore, there is always such an $\mathbf{x} \in \mathrm{row}(A)$ satisfying $A\mathbf{x} = \mathbf{b}$ by the observation above.

(b) It is unique iff $\mathrm{null}(AA^T) = 0$, because $AA^T\mathbf{a}_1 = AA^T\mathbf{a}_2$ iff $\mathbf{a}_1 - \mathbf{a}_2 \in \mathrm{null}(AA^T)$. Note that it is equivalent to that $\mathrm{rk}(A) = \mathrm{rk}(AA^T) = m$ (i.e., A has full row-rank) by the rank–nullity theorem.

(c) The set of such solutions, $\{\mathbf{x} \in \mathrm{row}(A) : A\mathbf{x} = \mathbf{b}\}$, can be represented as

$$\{A^T\mathbf{y} : \mathbf{y} = \mathbf{y}_0 + \mathbf{z},\ (AA^T)\mathbf{z} = \mathbf{0}\}$$

for some $\mathbf{y}_0$ satisfying $AA^T\mathbf{y}_0 = \mathbf{b}$. Therefore,

1. first, find any $\mathbf{y}_0$ satisfying $AA^T\mathbf{y}_0 = \mathbf{b}$, which exists, using various method taught in the lecture;
2. the set of such solutions is

$$\left\{A^T(\mathbf{y}_0 + \mathbf{z}) : \mathbf{z} \in (AA^T)\right\}.$$

2. (a) No. When $n = 1$ and $A = (-1)$, there is clearly no real matrix $B \in \mathbb{R}^{1\times 1}$ satisfying $B^2 = A$.

(b) If and only if the multiplicity of each negative eigenvalue of A is even.

First, assume the condition above. By the diagonalization (always possible since A is real and symmetric) and rearrangement of the eigenvalues, we may assume

$$A = P^{-1}DP, \quad D = \operatorname{diag}(\lambda_1, \cdots, \lambda_n),$$

where $\lambda_1 = \lambda_2 \leq \cdots \leq \lambda_{2k-1} = \lambda_{2k} < 0 \leq \lambda_{2k+1} \leq \cdots \leq \lambda_n$. Here $2k$ is the index of the largest negative eigenvalue. Note that we can write

$$D = \begin{bmatrix} \operatorname{diag}(\lambda_2, \lambda_2) & & & \\ & \ddots & & \\ & & \operatorname{diag}(\lambda_{2k}, \lambda_{2k}) & \\ & & & \operatorname{diag}(\lambda_{2k+1}, \cdots, \lambda_n) \end{bmatrix},$$

and observe that $\operatorname{diag}(\lambda_{2j}, \lambda_{2j}) = (\Lambda_j J)^2$ for $j \leq k$, where

$$\Lambda_j = \operatorname{diag}\left(\sqrt{|\lambda_{2j}|}, \sqrt{|\lambda_{2j}|}\right), \qquad J = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

Therefore,

$$\sqrt{D} = \begin{bmatrix} \Lambda_1 J & & & \\ & \ddots & & \\ & & \Lambda_k J & \\ & & & \operatorname{diag}\left(\sqrt{\lambda_{2k+1}}, \cdots, \sqrt{\lambda_n}\right) \end{bmatrix}$$

yields $\sqrt{D}^2 = D$ so that it makes $B = P^{-1}\sqrt{D}P$ to satisfy $B^2 = P^{-1}\sqrt{D}^2 P = P^{-1}DP = A$.

Let us prove the converse. We may assume A is diagonal without loss of generality. Let $B^2 = A$ and $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ be the roots of the characteristic polynomial of B, namely:

$$\det(tI - B) = (t - \lambda_1) \cdots (t - \lambda_n).$$

Then,

$$\begin{aligned} \det(tI + B) &= (-1)^n \det(sI - B)|_{s=-t} \\ &= (-1)^n(-t - \lambda_1) \cdots (-t - \lambda_n) = (t + \lambda_1) \cdots (t + \lambda_n) \end{aligned}$$

so that

$$\det(tI - A) = \det((\sqrt{t}I - B)(\sqrt{t}I + B))$$

$$
\begin{aligned}
&= (\sqrt{t} - \lambda_1)(\sqrt{t} + \lambda_1) \cdots (\sqrt{t} - \lambda_n)(\sqrt{t} + \lambda_n) \\
&= (t - \lambda_1^2) \cdots (t - \lambda_n^2).
\end{aligned}
$$

Assume A has $\lambda < 0$ for an odd number of entries. Then, $\lambda_j^2 = \lambda < 0$ for an odd number of j, so that $\lambda_j = \pm i\sqrt{-\lambda}$ for an odd number of j. However, we know that $\det(tI - B)$ is a polynomial of real coefficients. Observe that, when a pure imaginary number ia ($a \in \mathbb{R} \setminus \{0\}$) is a root of the polynomial f of real coefficients, the multiplicity of ia and $-ia$ should be the same:

$$f(ia) = 0 = \overline{f(ia)} = f(\overline{ia}) = f(-ia)$$

with an induction process on the degree of f. Therefore, the multiplicity of $i\sqrt{-\lambda}$ and $-i\sqrt{-\lambda}$ should be the same as roots of the polynomial $\det(tI - B)$ of real coefficients. Thus, the sum of the multiplicities of $i\sqrt{-\lambda}$ and $-i\sqrt{-\lambda}$ should be even, which contradicts to the assumption. This completes the proof by contradiction.

(c) I constructed the matrix B in (b). Basically, first diagonalize $A = P^{-1}DP$ and sort the eigenvalues in nondecreasing order. Then find the $\sqrt{D}$ satisfying $\sqrt{D}^2 = D$ according to the construction in (b) and return $B = P^{-1}\sqrt{D}P$.

3. The Frobenius norm is the square root of sum of squares of the singular values of A. Note that we have

$$\|A\|_F^2 = \sum_{i=1}^{\min(m,n)} \sigma_i^2(A) \leq \left(\sum_{i=1}^{\min(m,n)} \sigma_i(A) \right)^2 = \|A\|_N^2$$

due to the existences of the cross terms (≥ 0). The equality holds iff every cross term becomes zero, which is equivalent to that there are at most one nonzero singular values.

Furthermore, the spectral norm of A is the largest singular value of A, so we have $\|A\|_2 = \sigma_1(A) = \sqrt{\sigma_1^2(A)} \leq \sqrt{\sum_{i=1}^{\min(m,n)} \sigma_i^2(A)} = \|A\|_F$ (with the ordering $\sigma_1(A) \geq \cdots \geq \sigma_{\min(m,n)}(A)$.) The equality holds iff $\sigma_2(A) = \cdots = \sigma_{\min(m,n)}(A) = 0$, i.e., there are at most one nonzero singular values.

To sum up, we have $\|A\|_2 \leq \|A\|_F \leq \|A\|_N$. Both equalities hold iff there are at most one nonzero singular values.

4. Note that the given matrix A is of rank 2. Therefore, we have

$$C = A_{1:3,1:2} = \begin{bmatrix} 3 & 4 \\ 1 & 2 \\ 5 & 8 \end{bmatrix}.$$

Corresponding Z is as follows:

$$Z = A_{1:2,1:2}^{-1} A_{1:2\times1:5} = \begin{bmatrix} 1 & 0 & -1 & 2 & -3 \\ 0 & 1 & 2 & -1 & 3 \end{bmatrix}.$$

Finally, $B = A_{1:2,1:2} = \begin{bmatrix} 3 & 4 \\ 1 & 2 \end{bmatrix}$ and $Y = CB^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 2 \end{bmatrix}$.

5. Let λ_i and $\mathbf{v}_i$ be (strictly decreasing) eigenvalues and corresponding eigenvectors of S, then we have

$$\begin{aligned} \lambda_1 &= 3, \quad \mathbf{v}_1 = \mathbf{e}_1; \\ \lambda_2 &= 2, \quad \mathbf{v}_2 = \mathbf{e}_2; \\ \lambda_3 &= 1, \quad \mathbf{v}_3 = \mathbf{e}_3. \end{aligned}$$

Set c_i's to satisfy $\mathbf{u} = \sum_{i=1}^{3} c_i \mathbf{v}_i$. Then, for an eigenvalue z of $S + \theta \mathbf{u}\mathbf{u}^T$, we have

$$\frac{1}{\theta} = \sum_{i=1}^{3} \frac{c_i^2}{z - \lambda_i} = \frac{c_1^2}{z-3} + \frac{c_2^2}{z-2} + \frac{c_3^2}{z-1}$$

under the assumption that z is not an eigenvalue of S. We have four unknowns, c_i's and θ, and four conditions: the above equation for three given eigenvalues and that $\|\mathbf{u}\|^2 = \sum_{i=1}^{3} c_i^2 = 1$. Thus, we can uniquely determine the values of θ and $\mathbf{u}$ in general situation.

Let $C_1 = c_1^2$ and $C_2 = c_2^2$ in the following discussions; note that $c_3^2 = 1 - C_1 - C_2 \geq 0$.

When $z = \lambda_i$ is an eigenvalue of S, we have

$$(\lambda_i I - S)\mathbf{v} = \theta \mathbf{u}(\mathbf{u}^T \mathbf{v}).$$

Letting $\mathbf{v} = \sum_{j=1}^n b_j \mathbf{v}_j$,

$$(\lambda_i I - S)\mathbf{v} = \sum_{j=1}^n b_j(\lambda_i - \lambda_j)\mathbf{v}_j = \theta \left(\sum_{j=1}^n c_j b_j \right) \sum_{j=1}^n c_j \mathbf{v}_j$$

implying that $b_j(\lambda_j - \lambda_i) = c_j\theta \left(\sum_{j=1}^n c_j b_j \right)$. When $j = i$, we should have $c_i = 0$ unless $\theta = 0$ or $\sum_{j=1}^n c_j b_j = \mathbf{u}^T\mathbf{v} = 0$. However, $\mathbf{u}^T\mathbf{v} = 0$ means that $S\mathbf{v} = (S + \theta\mathbf{u}\mathbf{u}^T)\mathbf{v} = \lambda_i\mathbf{v}$ so that $\mathbf{v} = k\mathbf{v}_i$ for some $0 \neq k \in \mathbb{R}$. Therefore, this implies again that $c_i = \frac{1}{k}\mathbf{u}^T\mathbf{v} = 0$. Consequently, when $z = \lambda_i$ is an eigenvalue of $S + \theta\mathbf{u}\mathbf{u}^T$, then we have $c_i = 0$ instead of the equality involving the sum of fractions above.

(a)

$$\begin{aligned} z = 4\colon \quad & \frac{1}{\theta} = \frac{C_1}{1} + \frac{C_2}{2} + \frac{1 - C_1 - C_2}{3}, \\ z = 3\colon \quad & c_1 = 0, \\ z = 2\colon \quad & c_2 = 0, \end{aligned}$$

therefore $\theta = 3$ and $\mathbf{u} = \mathbf{e}_3$.

(b)

$$\begin{aligned} z = 3.3\colon \quad & \frac{1}{\theta} = \frac{C_1}{0.3} + \frac{C_2}{1.3} + \frac{1 - C_1 - C_2}{2.3}, \\ z = 2.2\colon \quad & \frac{1}{\theta} = \frac{C_1}{-0.8} + \frac{C_2}{0.2} + \frac{1 - C_1 - C_2}{1.2}, \\ z = 1.1\colon \quad & \frac{1}{\theta} = \frac{C_1}{-1.9} + \frac{C_2}{-0.9} + \frac{1 - C_1 - C_2}{0.1}, \end{aligned}$$

this gives $C_1 = 19/50$ and $C_2 = 39/100$, with corresponding $\theta = 3/5$ and $\mathbf{u} = (\sqrt{19/50}, \sqrt{39/100}, \sqrt{23/100})^T$.

(c)

$$\begin{aligned} z = 3.5\colon \quad & \frac{1}{\theta} = \frac{C_1}{0.5} + \frac{C_2}{1.5} + \frac{1 - C_1 - C_2}{2.5} \overset{(*)}{=} \frac{C_1}{0.5} + \frac{1 - C_1}{2.5}, \\ z = 2.5\colon \quad & \frac{1}{\theta} = \frac{C_1}{-0.5} + \frac{C_2}{0.5} + \frac{1 - C_1 - C_2}{1.5} \overset{(*)}{=} \frac{C_1}{-0.5} + \frac{1 - C_1}{1.5}, \\ (*)\ z = 2\colon \quad & c_2 = 0, \end{aligned}$$

therefore $C_1 = 1/16$, $\theta = 2$ and $\mathbf{u} = (1/4, 0, \sqrt{15}/4)^T$.

6. (a) Exchanging columns about the horizontal center of the matrix, we observe that

$$\det(A) = \begin{vmatrix} x_1^{n-1} & \cdots & x_1^0 \\ \vdots & \ddots & \vdots \\ x_n^{n-1} & \cdots & x_n^0 \end{vmatrix} = \epsilon_n \begin{vmatrix} x_1^0 & \cdots & x_1^{n-1} \\ \vdots & \ddots & \vdots \\ x_n^0 & \cdots & x_n^{n-1} \end{vmatrix}$$

where $\epsilon_n = +1$ when $n \equiv 0, 1 \pmod 4$ and $\epsilon_n = -1$ otherwise; i.e., $\epsilon_n = (-1)^{n(n-1)/2}$. Since the determinant in the very right hand side is the Vandermonde determinant, we have

$$\det(A) = (-1)^{n(n-1)/2} \prod_{1 \le i < j \le n} (x_j - x_i).$$

(b) Denote

$$P = \operatorname{diag}(p_1(x_1), \cdots, p_n(x_n)), \qquad Q = \operatorname{diag}(p(y_1), \cdots, p(y_n)),$$
$$V_x = \begin{bmatrix} x_1^0 & \cdots & x_n^0 \\ x_1^1 & \cdots & x_n^1 \\ \vdots & \ddots & \vdots \\ x_1^{n-1} & \cdots & x_n^{n-1} \end{bmatrix}, \qquad V_y = \begin{bmatrix} y_1^0 & \cdots & y_n^0 \\ y_1^1 & \cdots & y_n^1 \\ \vdots & \ddots & \vdots \\ y_1^{n-1} & \cdots & y_n^{n-1} \end{bmatrix}$$

where

$$p(x) = \prod_{i=1}^{n} (x - x_i), \qquad p_j(x) = \frac{p(x)}{x - x_j}.$$

Then, the following identity holds:

$$A = -P V_x^{-1} V_y Q^{-1}.$$

In fact, let us denote $p_i(x) = \sum_{k=1}^{n} p_{ik} x^{k-1}$, then, observing $p_i(x_j) = 0$ for $i \neq j$,

$$P = \big(p_i(x_j)\big)_{ij} = \big(p_{ij}\big)_{ij} V_x \tag{$\dagger$}$$

by the definition of the matrix multiplication. Similarly, we have

$$\big(p_i(y_j)\big)_{ij} = \big(p_{ij}\big)_{ij} V_y \overset{(\dagger)}{=} P V_x^{-1} V_y.$$

Note that

$$(AQ)_{ij} = \frac{1}{x_i - y_j} \cdot p(y_j) = -p_i(y_j).$$

Therefore, we have $-AQ = PV_x^{-1}V_y$, i.e., $A = -PV_x^{-1}V_yQ^{-1}$.

Now, we can easily find the determinant of A:

$$\begin{aligned}\det(A) &= (-1)^n \det(P)\det(V_x)^{-1}\det(V_y)\det(Q)^{-1} \\ &= \frac{(-1)^n \cdot \prod_{i=1}^n p_i(x_i) \cdot \det(V_y)}{\prod_{i=1}^n p(y_i) \cdot \det(V_x)}.\end{aligned}$$

Note that $\prod_{i=1}^n p_i(x_i) = \prod_{i=1}^n \prod_{1\le j\le n,\ j\ne i}(x_j - x_i) = (-1)^{n(n-1)/2}(\det(V_x))^2$ since there are $\frac{n(n-1)}{2}$ 'swaps' of the indices and the remaining thing is the squared of the Vandermonde determinant. Therefore,

$$\begin{aligned}\det(A) &= \frac{(-1)^n \cdot (-1)^{n(n-1)/2}(\det(V_x))^2 \cdot \det(V_y)}{\prod_{i=1}^n p(y_i) \cdot \det(V_x)} \\ &= \frac{(-1)^{n(n+1)/2} \prod_{1\le i<j\le n}(x_j - x_i) \cdot \prod_{1\le i<j\le n}(y_j - y_i)}{\prod_{j=1}^n \prod_{i=1}^n (y_j - x_i)} \\ &= \frac{(-1)^{n(n+1)/2} \prod_{1\le i<j\le n}(x_j - x_i) \cdot (-1)^{n(n-1)/2} \prod_{1\le i<j\le n}(y_i - y_j)}{(-1)^{n^2} \prod_{j=1}^n \prod_{i=1}^n (x_i - y_j)} \\ &= \frac{\prod_{1\le i<j\le n}(x_j - x_i) \cdot \prod_{1\le i<j\le n}(y_i - y_j)}{\prod_{j=1}^n \prod_{i=1}^n (x_i - y_j)}\end{aligned}$$

where all the signs can be cancelled with appropriate choice of indices.

(c) By a simple calculation, $\det(A) = x_1 - y_1$ when $n = 1$, and $\det(A) = (x_1 - x_2)(y_1 - y_2)$ when $n = 2$. When $n \ge 3$, we have

$$\det(A) = \begin{vmatrix} x_1 - y_1 & \cdots & x_1 - y_n \\ x_2 - y_1 & \cdots & x_2 - y_n \\ \vdots & \ddots & \vdots \\ x_n - y_1 & \cdots & x_n - y_n \end{vmatrix} = \begin{vmatrix} x_1 - y_1 & \cdots & x_1 - y_n \\ x_2 - x_1 & \cdots & x_2 - x_1 \\ \vdots & \ddots & \vdots \\ x_n - x_1 & \cdots & x_n - x_1 \end{vmatrix} = 0$$

since there are $n - 1(\ge 2)$ rows which are a constant times of $(1, \dots, 1)$.

7. (a) $B = \mathbf{v}\mathbf{v}^T$ where $\mathbf{v} = (1,2,3,4,5)^T$. Letting $\mathbf{u} = \mathbf{v}$, we have

$$(I_5 - B)^{-1} = I_5 + \frac{\mathbf{v}\mathbf{v}^T}{1 - \mathbf{v}^T\mathbf{v}} = I_5 - \frac{1}{54}B$$
$$= \begin{bmatrix} 53/54 & -1/27 & -1/18 & -2/27 & -5/54 \\ -1/27 & 25/27 & -1/9 & -4/27 & -5/27 \\ -1/18 & -1/9 & 5/6 & -2/9 & -5/18 \\ -2/27 & -4/27 & -2/9 & 19/27 & -10/27 \\ -5/54 & -5/27 & -5/18 & -10/27 & 29/54 \end{bmatrix}.$$

(b) $C = \mathbf{v}\mathbf{v}^T - \mathbf{w}\mathbf{w}^T$ where $\mathbf{v} = (1,2,3,4,5)^T$ and $\mathbf{w} = (1,1,1,1,1)^T$. By the Sherman–Morrison–Woodbury formula,

$$\begin{aligned}(I_5 - C)^{-1} &= ((I_5 - \mathbf{v}\mathbf{v}^T) - (-\mathbf{w})\mathbf{w}^T)^{-1} \\ &= (I_5 - \mathbf{v}\mathbf{v}^T)^{-1} + (I_5 - \mathbf{v}\mathbf{v}^T)^{-1}(-\mathbf{w}) \cdot \\ &\qquad \left[1 - \mathbf{w}^T(I_5 - \mathbf{v}\mathbf{v}^T)^{-1}(-\mathbf{w})\right]^{-1} \mathbf{w}^T(I_5 - \mathbf{v}\mathbf{v}^T)^{-1}\end{aligned}$$

where we obtained $(I_5 - \mathbf{v}\mathbf{v}^T)^{-1}$ in (a). Putting all values, we have

$$(I_5 - C)^{-1} = \left(I_5 - \frac{1}{54}B\right) - \frac{\left(I_5 - \frac{1}{54}B\right)\mathbf{w}\mathbf{w}^T\left(I_5 - \frac{1}{54}B\right)}{1 + \mathbf{w}^T\left(I_5 - \frac{1}{54}B\right)\mathbf{w}}$$
$$= \begin{bmatrix} 23/33 & -7/33 & -4/33 & -1/33 & 2/33 \\ -7/33 & 9/11 & -5/33 & -4/33 & -1/11 \\ -4/33 & -5/33 & 9/11 & -7/33 & -8/33 \\ -1/33 & -4/33 & -7/33 & 23/33 & -13/33 \\ 2/33 & -1/11 & -8/33 & -13/33 & 5/11 \end{bmatrix}.$$

8. We have the following identity:

$$A^+ = \lim_{t \downarrow 0} (A^T A + tI)^{-1} A^T.$$

Calculating $A^T A$, we have

$$A^T A = \begin{bmatrix} 14 & 17 & 20 & 23 \\ 17 & 22 & 27 & 32 \\ 20 & 27 & 34 & 41 \\ 23 & 32 & 41 & 50 \end{bmatrix}$$

so that

$$A^TA + tI = \begin{bmatrix} 14+t & 17 & 20 & 23 \\ 17 & 22+t & 27 & 32 \\ 20 & 27 & 34+t & 41 \\ 23 & 32 & 41 & 50+t \end{bmatrix}$$

and hence

$$(A^TA{+}tI)^{-1} = \frac{1}{t(t^2+120t+380)} \begin{bmatrix} 114+t(106+t) & -152-17t & -38-20t & 76-23t \\ -152-17t & 266+t(98+t) & -76-27t & -38-32t \\ -38-20t & -76-27t & 266+t(86+t) & -152-41t \\ 76-23t & -38-32t & -152-41t & 114+t(70+t) \end{bmatrix}$$

and

$$(A^TA{+}tI)^{-1}A^T = \frac{1}{t^2+120t+380} \begin{bmatrix} -80+t & 2(-17+t) & 3(46+t) \\ 2(-15+t) & -8+3t & 3(22+t) \\ 20+3t & 18+4t & 3(-2+t) \\ 70+4t & 44+5t & 3(-26+t) \end{bmatrix}.$$

As $t \downarrow 0$, we have

$$A^+ = \lim_{t\downarrow 0}(A^TA{+}tI)^{-1}A^T = \frac{1}{380}\begin{bmatrix} -80 & -34 & 138 \\ -30 & -8 & 66 \\ 20 & 18 & -6 \\ 70 & 44 & -78 \end{bmatrix} = \begin{bmatrix} -4/19 & -17/190 & 69/190 \\ -3/38 & -2/95 & 33/190 \\ 1/19 & 9/190 & -3/190 \\ 7/38 & 11/95 & -39/190 \end{bmatrix}.$$

9. By the singular value decomposition of A, we have

$$A = U\Sigma V^T$$

where

$$U = [\mathbf{u}_1\ \mathbf{u}_2] = \begin{bmatrix} \frac{1}{2}\sqrt{2-\sqrt{\frac{2}{5}}} & -\frac{1}{2}\sqrt{2+\sqrt{\frac{2}{5}}} \\ \frac{1}{2}\sqrt{2+\sqrt{\frac{2}{5}}} & \frac{1}{2}\sqrt{2-\sqrt{\frac{2}{5}}} \end{bmatrix}, \quad \Sigma = \operatorname{diag}(\sigma_1,\sigma_2) = \operatorname{diag}(\sqrt{2}{+}\sqrt{5}, -\sqrt{2}{+}\sqrt{5}),$$

and

$$V = [\mathbf{v}_1\ \mathbf{v}_2] = \begin{bmatrix} \frac{1}{2}\sqrt{2-3\sqrt{\frac{2}{5}}} & -\frac{1}{2}\sqrt{2+3\sqrt{\frac{2}{5}}} \\ \frac{1}{2}\sqrt{2+3\sqrt{\frac{2}{5}}} & \frac{1}{2}\sqrt{2-3\sqrt{\frac{2}{5}}} \end{bmatrix}.$$

Since $dA/dt = \begin{bmatrix} 2 & 0 \\ 2 & 0 \end{bmatrix}$, we have

$$\begin{aligned}
\left.\frac{d\sigma_1(t)}{dt}\right|_{t=0} &= \mathbf{u}_1^T \left.\frac{dA}{dt}\right|_{t=0} \mathbf{v}_1 \\
&= \begin{bmatrix} \frac{1}{2}\sqrt{2-\sqrt{\frac{2}{5}}} & \frac{1}{2}\sqrt{2+\sqrt{\frac{2}{5}}} \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{2}\sqrt{2-3\sqrt{\frac{2}{5}}} \\ \frac{1}{2}\sqrt{2+3\sqrt{\frac{2}{5}}} \end{bmatrix} = \frac{1}{\sqrt{5}},
\end{aligned}$$

and

$$\begin{aligned}
\left.\frac{d\sigma_2(t)}{dt}\right|_{t=0} &= \mathbf{u}_2^T \left.\frac{dA}{dt}\right|_{t=0} \mathbf{v}_2 \\
&= \begin{bmatrix} -\frac{1}{2}\sqrt{2+\sqrt{\frac{2}{5}}} & \frac{1}{2}\sqrt{2-\sqrt{\frac{2}{5}}} \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} -\frac{1}{2}\sqrt{2+3\sqrt{\frac{2}{5}}} \\ \frac{1}{2}\sqrt{2-3\sqrt{\frac{2}{5}}} \end{bmatrix} = \frac{1}{\sqrt{5}}.
\end{aligned}$$

10.

$$\begin{aligned}
\det(L-\lambda I_n) &= \begin{vmatrix} L_{11}-\lambda & L_{12} & \cdots & L_{1n} \\ L_{21} & L_{22}-\lambda & \cdots & L_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ L_{n1} & L_{n2} & \cdots & L_{nn}-\lambda \end{vmatrix} \\
&= \begin{vmatrix} L_{11}-\lambda & L_{12} & \cdots & L_{1n} \\ L_{21} & L_{22}-\lambda & \cdots & L_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^n L_{i1}-\lambda & \sum_{i=1}^n L_{i2}-\lambda & \cdots & \sum_{i=1}^n L_{in}-\lambda \end{vmatrix} \\
&= \begin{vmatrix} L_{11}-\lambda & L_{12} & \cdots & L_{1n} \\ L_{21} & L_{22}-\lambda & \cdots & L_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -\lambda & -\lambda & \cdots & -\lambda \end{vmatrix} \\
&= \lambda \begin{vmatrix} L_{11}-\lambda & L_{12} & \cdots & L_{1n} \\ L_{21} & L_{22}-\lambda & \cdots & L_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -1 & -1 & \cdots & -1 \end{vmatrix} =: \lambda \det(M(\lambda)).
\end{aligned}$$

This completes the verification for $\det(L-\lambda I_n) = \lambda \det(M(\lambda))$.

Now, let L_0 be a matrix which is obtained from L by removing the i-th column and the i-th row. By interchanging the i-th column with the n-th column and the i-th row with the n-th row, we may assume that L_0 is obtained by removing the n-th row and the n-th column from L without loss of any generality: this is because the process mentioned just before is nothing but multiplying the permutation matrix

$$P_{in} = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 0 & & 1 & \leftarrow i\text{-th} \\ & & & \ddots & & \\ & & 1 & & 0 & \leftarrow n\text{-th} \end{bmatrix} = P_{in}^{-1}$$

on the left and the right simultaneously, which is under the matrix similarity, hence preserving the determinant $(\det(L) = \det(P_{in}^{-1} L P_{in}).)$

Now, we have

$$\begin{aligned}
\det(L - \lambda I) &= \lambda \begin{vmatrix} L_{11} - \lambda & \cdots & L_{1,n-1} & L_{1n} \\ \vdots & \ddots & \vdots & \vdots \\ L_{n-1,1} & \cdots & L_{n-1,n-1} - \lambda & L_{n-1,n} \\ -1 & \cdots & -1 & -1 \end{vmatrix} \\
&= \lambda \begin{vmatrix} L_{11} - \lambda & \cdots & L_{1,n-1} & L_{1n} \\ \vdots & \ddots & \vdots & \vdots \\ L_{n-1,1} & \cdots & L_{n-1,n-1} - \lambda & L_{n-1,n} \\ -1 & \cdots & -1 & -1 \end{vmatrix} \\
&= -\lambda \left(\det(L_0 - \lambda I_{n-1}) + \sum_{i=1}^{n-1} (-1)^i \det([L_0 - \lambda I_{n-1}]_{n-i}^*) \right)
\end{aligned}$$

where $[A]_j^*$ is the matrix obtained from A by deleting j-th column of A and gluing $(L_{1n}, \cdots, L_{n-1,n})^T$ at the very right side of it. By exchaning the order of columns, we can observe that

$$\det([A]_j^*) = (-1)^{n-1-j} \det([A]_j)$$

where $[A]_j^*$ is the matrix obtained from A by replacing j-th column by $(L_{1n}, \cdots, L_{n-1,n})^T$. By the Cramer's rule, we know that

$$\det([L_0 - \lambda I_{n-1}]_{n-i}) = \det(L_0 - \lambda I_{n-1})\, x_{n-i}(\lambda)$$

where $\mathbf{x} = \mathbf{x}(\lambda) = (x_1(\lambda), \cdots, x_{n-1}(\lambda))^T$ satisfies

$$(L_0 - \lambda I_{n-1})\mathbf{x} = (L_{1n}, \cdots, L_{n-1,n})^T.$$

Therefore, we have

$$\det(L - \lambda I_n) = -\lambda \det(L_0 - \lambda I_{n-1}) \left(1 + \sum_{i=1}^{n-1} (-1)^i \cdot (-1)^{n-1-(n-i)} x_{n-i}(\lambda)\right).$$

Since L has n eigenvalues, $\lambda_1, \ldots, \lambda_{n-1}$ and 0, the characteristic polynomial of L is

$$\det(L - \lambda I_n) = -\lambda(\lambda_1 - \lambda) \cdots (\lambda_{n-1} - \lambda)$$

whence

$$(\lambda_1 - \lambda) \cdots (\lambda_{n-1} - \lambda) = \det(L_0 - \lambda I_{n-1}) \left(1 - \sum_{i=1}^{n-1} x_{n-i}(\lambda)\right). \qquad (\dagger)$$

However, with $\mathbf{v} = (1, 1, \cdots, 1)^T$, we know that

$$L_0\mathbf{v} + (L_{1n}, \cdots, L_{n-1,n})^T = \mathbf{0}$$

so that $\mathbf{x}(0) = -\mathbf{v}$, i.e., $x_i(0) = -1$ for any $i = 1, \cdots, n-1$. By putting $\lambda = 0$ in ($\dagger$), we obtain

$$\lambda_1 \cdots \lambda_{n-1} = \det(L_0) \left(1 + \sum_{i=1}^{n-1} 1\right) = n \det(L_0).$$

This completes the proof.

11. (a) Since the determinant is a multilinear form,

$$\begin{aligned} T(a_1A_1 + a_2A_2) &= \det([(a_1A_1 + a_2A_2)B_1 \ B_2 \ \cdots \ B_n]) + \cdots \\ &\quad + \det([B_1 \ B_2 \ \cdots \ (a_1A_1 + a_2A_2)B_n]) \\ &= \Big[a_1 \det([A_1B_1 \ B_2 \ \cdots \ B_n]) + a_2 \det([A_2B_1 \ B_2 \ \cdots \ B_n])\Big] + \cdots \\ &\quad + \Big[a_1 \det([B_1 \ B_2 \ \cdots \ A_1B_n]) + a_2 \det([B_1 \ B_2 \ \cdots \ A_2B_n])\Big] \\ &= a_1T(A_1) + a_2T(A_2). \end{aligned}$$

(b) Denote a matrix which is obtained from B by removing the i-th row and the k-th column as $\hat{B}_{ik}$, and let $B^{(i)}$ be the i-th row of B: $B =$

$[(B^{(1)})^T \ \cdots \ (B^{(n)})^T]^T$. Then, we have

$$\begin{aligned}
T(E_{ij}) &= \sum_{k=1}^{n} \det([B_1 \ \cdots \ E_{ij}B_k \ \cdots \ B_n]) \\
&= \sum_{k=1}^{n} \det([B_1 \ \cdots \ b_{jk}\mathbf{e}_i \ \cdots \ B_n]) \\
&= \sum_{k=1}^{n} (-1)^{i+k} b_{jk} \det(\hat{B}_{ik}) \\
&= \det\left(\begin{array}{c} \begin{bmatrix} — & B^{(1)} & — \\ & \vdots & \\ — & B^{(j)} & — \\ & \vdots & \\ — & B^{(j)} & — \\ & \vdots & \\ — & B^{(n)} & — \end{bmatrix} \begin{array}{l} \\ \\ \leftarrow j\text{-th} \\ \\ \leftarrow i\text{-th} \\ \\ \\ \end{array} \end{array}\right) \\
&= \det(B)\delta_{ij} = \operatorname{tr}(E_{ij}) \det(B).
\end{aligned}$$

(c) By the linearity of T and $\operatorname{tr}^{(*)}$, for $A = (a_{ij})$,

$$\begin{aligned}
T(A) &= T\left(\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}E_{ij}\right) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}T(E_{ij}) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} \operatorname{tr}(E_{ij}) \cdot \det(B) \\
&= \operatorname{tr}\left(\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}E_{ij}\right) \cdot \det(B) \\
&= \operatorname{tr}(A) \det(B).
\end{aligned}$$

This completes the proof.

$(*)$: Trace is linear because $\alpha \operatorname{tr}(A) + \beta \operatorname{tr}(B) = \sum_i (\alpha a_{ii} + \beta b_{ii}) = \operatorname{tr}(\alpha A + \beta B)$.