MATH 245 Linear Algebra 2, Solutions to the Exercises for Chapter 8

**1:** (a) Let $A=\begin{pmatrix}0&2&4\\2&-3&2\\4&2&0\end{pmatrix}$. Find an orthogonal matrix $P$ and a diagonal matrix $D$ such that $P^TAP=D$.

Solution: The characteristic polynomial of $A$ is

$$\begin{aligned}f_A(x)&=\begin{vmatrix}-x&2&4\\2&-3-x&2\\4&2&-x\end{vmatrix}\\&=-x^2(x+3)+16+16+4x+4x+16(x+3)\\&=-(x^3+3x^2-24x-80)=-(x-5)(x^2+8x+16)=-(x-5)(x+4)^2\end{aligned}$$

so the eigenvalues are $\lambda_1=5$ and $\lambda_2=\lambda_3=-4$. When $\lambda=\lambda_1=5$ we have

$$A-\lambda I=\begin{pmatrix}-5&2&4\\2&-8&2\\4&2&-5\end{pmatrix}\sim\begin{pmatrix}-5&2&4\\1&-4&1\\4&2&-5\end{pmatrix}\sim\begin{pmatrix}1&-4&1\\0&-18&9\\0&18&-9\end{pmatrix}\sim\begin{pmatrix}1&-4&1\\0&1&-\frac12\\0&17&-9\end{pmatrix}\sim\begin{pmatrix}1&0&-1\\0&1&-\frac12\\0&0&0\end{pmatrix}$$

so we can take $v_1=\begin{pmatrix}2\\1\\2\end{pmatrix}$ so that $\{v_1\}$ is a basis for the eigenspace $E_5$, then let $u_1=\dfrac{v_1}{|v_1|}=\frac13\begin{pmatrix}2\\1\\2\end{pmatrix}$. Since $A$ is symmetric, we know that its eigenspaces are orthogonal so we have $E_{-4}={E_5}^{\perp}$. To find a basis for $E_{-4}$ we can, by inspection, choose a unit vector $u_2$ which is orthogonal to $u_1$ and then choose $u_3=u_1\times u_2$. We choose $u_2=\frac{1}{\sqrt2}\begin{pmatrix}-1\\0\\1\end{pmatrix}$ and $u_3=\frac{1}{3\sqrt2}\begin{pmatrix}2\\1\\2\end{pmatrix}\times\begin{pmatrix}-1\\0\\1\end{pmatrix}=\frac{1}{3\sqrt2}\begin{pmatrix}1\\-4\\1\end{pmatrix}$. Thus we can orthogonally diagonalize $A$ using $P=(u_1,u_2,u_3)=\begin{pmatrix}\frac23&-\frac{1}{\sqrt2}&\frac{1}{3\sqrt2}\\\frac13&0&-\frac{4}{3\sqrt2}\\\frac23&\frac{1}{\sqrt2}&\frac{1}{3\sqrt2}\end{pmatrix}$ and $D=\text{diag}(\lambda_1,\lambda_2,\lambda_3)=\begin{pmatrix}5&0&0\\0&-4&0\\0&0&-4\end{pmatrix}$.

(b) Let $A=\begin{pmatrix}5+i&-6\\2&-2+i\end{pmatrix}$. Find a unitary matrix $P$ and an upper-triangular matrix $T$ so that $P^*AP=T$.

Solution: The characteristic polynomial of $A$ is

$$\begin{aligned}f_A(x)&=\begin{vmatrix}(5+i)-x&-6\\2&(-2+i)-x\end{vmatrix}\\&=x^2-(3+2i)x+(-11+3i)+12=x^2-(3+2i)x+(1+3i).\end{aligned}$$

The eigenvalues are

$$\lambda=\frac{(3+2i)\pm\sqrt{(5+12i)-(4+12i)}}{2}=\frac{(3+2i)\pm1}{2}=2+i\,,\,1+i.$$

say $\lambda_1=2+i$ and $\lambda_2=1+i$. When $\lambda=\lambda_1=2+i$ we have

$$A-\lambda I=\begin{pmatrix}3&-6\\2&-4\end{pmatrix}\sim\begin{pmatrix}1&-2\\0&0\end{pmatrix}$$

so we can take $v_1=\begin{pmatrix}2\\1\end{pmatrix}$ so that $\{v_1\}$ is a basis for $E_{\lambda_1}$. By inspection, the vector $v_2=\begin{pmatrix}-1\\2\end{pmatrix}$ is orthogonal to $v_1$. Normalize these vectors to get $u_1=\dfrac{v_1}{|v_1|}=\frac{1}{\sqrt5}\begin{pmatrix}2\\1\end{pmatrix}$ and $u_2=\dfrac{v_2}{|v_2|}=\frac{1}{\sqrt5}\begin{pmatrix}-1\\2\end{pmatrix}$ so that $\{u_1,u_2\}$ is an orthonormal basis for $\mathbb{C}^2$. Thus we can take

$$P=(u_1,u_2)=\tfrac{1}{\sqrt5}\begin{pmatrix}2&-1\\1&2\end{pmatrix}\ \text{, and}$$

$$\begin{aligned}T=P^*AP&=\tfrac15\begin{pmatrix}2&1\\-1&2\end{pmatrix}\begin{pmatrix}5+i&-6\\2&-2+i\end{pmatrix}\begin{pmatrix}2&-1\\1&2\end{pmatrix}\\&=\tfrac15\begin{pmatrix}2&1\\-1&2\end{pmatrix}\begin{pmatrix}4+2i&-17-i\\2+i&-6+2i\end{pmatrix}=\tfrac15\begin{pmatrix}10+5i&-40\\0&5+5i\end{pmatrix}=\begin{pmatrix}2+i&-8\\0&1+i\end{pmatrix}.\end{aligned}$$

**2:** For $0 \neq u \in \mathbb{R}^3$ and $\theta \in \mathbb{R}$, let $R_{u,\theta} : \mathbb{R}^3 \to \mathbb{R}^3$ denote the rotation about the vector $u$ by the angle $\theta$ (where the direction of rotation is determined by the right-hand rule: the right thumb points in the direction of $u$ and the fingers curl in the direction of rotation).

(a) Let $u = \big(1,1,-1\big)^T$ and let $\theta = \frac{\pi}{3}$. Find $A = \big[R_{u,\theta}\big]_{\mathcal{S}}$ where $\mathcal{S}$ is the standard basis for $\mathbb{R}^3$.

Solution: Let $v = \big(0,1,1\big)^T$ and $w = \big(1,0,1\big)^T$. Let $\mathcal{U} = \{u,v,w\}$ and let $B = \big[R_{u,\theta}\big]_{\mathcal{U}}$. Note that $v$ and $w$ are othogonal to $u$ with $|v| = |w| = \sqrt{2}$ and $v \times w = u$, and we have $\theta(v,w) = \cos^{-1}\frac{v \bullet w}{|v||w|} = \cos^{-1}\frac{1}{2} = \frac{\pi}{3}$. Thus $R_{u,\theta}(u) = u$, $R_{u,\theta}(v) = w$ and $R_{u,\theta}(w) = w - v$, and so

$$B = \big[R_{u,\theta}\big]_{\mathcal{U}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix}.$$

We have $A = PBP^{-1}$ where $P$ is the change of basis matrix $P = \big[I\big]_{\mathcal{S}}^{\mathcal{U}} = \big(u_1, u_2, u_3\big)$. We calculate $P^{-1}$.

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ -1 & 1 & 1 & 0 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & -1 & -1 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & -1 & -1 & 1 & 0 \\ 0 & 0 & 3 & 2 & -1 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & -\frac{1}{3} \\ 0 & 1 & 0 & -\frac{1}{3} & \frac{2}{3} & \frac{1}{3} \\ 0 & 0 & 1 & \frac{2}{3} & -\frac{1}{3} & \frac{1}{3} \end{array}\right)$$

Thus

$$\begin{aligned} A = PBP^{-1} &= \frac{1}{3}\begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ -1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 & -1 \\ -1 & 2 & 1 \\ 2 & -1 & 1 \end{pmatrix} \\ &= \frac{1}{3}\begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ -1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 & -1 \\ -2 & 1 & -1 \\ 1 & 1 & 2 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 2 & 2 & 1 \\ -1 & 2 & -2 \\ -2 & 1 & 2 \end{pmatrix}. \end{aligned}$$

(b) Let $B = \begin{pmatrix} 2 & 3 & -6 \\ -3 & 6 & 2 \\ 6 & 2 & 3 \end{pmatrix}$. Find $c > 0$, $0 \neq u \in \mathbb{R}^3$ and $0 \leq \theta \leq \pi$ such that $B = \big[c\, R_{u,\theta}\big]$.

Solution: First, let us find the eigenvalues of the rotation $R_{u,\theta}$, where $0 \neq u \in \mathbb{R}^3$ and $\theta \in \mathbb{R}$. Let $u_1 = \frac{u}{|u|}$ and extend $\{u_1\}$ to an orthonormal basis $\mathcal{U} = \big\{u_1, u_2, u_3\big\}$ for $\mathbb{R}^3$. Writing $R = R_{u,\theta}$, we have

$$\big[R\big]_{\mathcal{U}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$$

and so

$$\begin{aligned} f_R(t) &= \begin{vmatrix} 1-t & 0 & 0 \\ 0 & \cos\theta - t & -\sin\theta \\ 0 & \sin\theta & \cos\theta - t \end{vmatrix} = (1-t)\big((\cos\theta - t)^2 + \sin^2 t\big) \\ &= (1-t)(\cos^2\theta - 2\cos\theta\, t + t^2 + \sin^2\theta) = -(t-1)(t^2 - 2\cos\theta\, t + 1) \\ &= -(t-1)\big(t - e^{i\,\theta}\big)\big(t - e^{-i\,\theta}\big)\,. \end{aligned}$$

Thus the eigenvalues of $R = R_{u,\theta}$ are $1, e^{\pm i\,\theta}$. It follows that for $c > 0$, the eigenvalues of $cR$ are $c, c\,e^{\pm i\,\theta}$. Now let us find the eigenvalues of $B$. We have

$$\begin{aligned} f_B(t) = \big|B - tI\big| &= \begin{vmatrix} 2-t & 3 & -6 \\ -3 & 6-t & 2 \\ 6 & 2 & 3-t \end{vmatrix} \\ &= (2-t)(18 - 9t + t^2) + 36 + 36 - 4(2-t) + 9(3-t) + 36(6-t) \\ &= 36 - 36\,t + 11\,t^2 - t^3 + 72 - 8 + 4t + 27 - 9t + 216 - 36\,t \\ &= -(t^3 - 11\,t^2 + 77\,t - 343) = -(t-7)(t^2 - 4t + 49) \end{aligned}$$

so the eigenvalues of $B$ are $\lambda = 7$ or $\lambda = \frac{4 \pm \sqrt{16 - 4\cdot 49}}{2} = 2 \pm \sqrt{-45} = 2 \pm 3\sqrt{5}\,i$. Thus in order to have $B = \big[c\,R_{u,\theta}\big]$ with $c > 0$ and $0 \leq \theta \leq \pi$, we must have $c = 7$ and $\theta = \cos^{-1}\frac{2}{7}$. To find the required vector $u$, we find a basis for the eigenspace $E_7$. We have

$$B - 7I = \begin{pmatrix} -5 & 3 & -6 \\ -3 & -1 & 2 \\ 6 & 2 & -4 \end{pmatrix} \sim \begin{pmatrix} 1 & 5 & -10 \\ -3 & -1 & 2 \\ 0 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 5 & -10 \\ 0 & 14 & -28 \\ 0 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 5 & -10 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix}$$

so a basis for $E_7$ is given by $\{u\}$ where $u = \pm\big(0,2,1\big)^T$. We still need to take some care in our choice of the vector $u$. If we chose $u = \big(0,2,1\big)^T$, $v = \big(0,-1,2\big)^T$, $w = \big(\sqrt{5},0,0\big)^T$ so that $\mathcal{U} = \{u, v, w\}$ is a positively oriented orthogonal basis with $|u| = |v| = |w|$, then we would have

$$B(u,v,w) = \begin{pmatrix} 2 & 3 & -6 \\ -3 & 6 & 2 \\ 6 & 2 & 3 \end{pmatrix} \begin{pmatrix} 0 & 0 & \sqrt{5} \\ 2 & -1 & 0 \\ 1 & 2 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -15 & 2\sqrt{5} \\ 14 & -2 & -3\sqrt{5} \\ 7 & 4 & 6\sqrt{5} \end{pmatrix} = (u,v,w) \begin{pmatrix} 7 & 0 & 0 \\ 0 & 2 & 3\sqrt{5} \\ 0 & -3\sqrt{5} & 2 \end{pmatrix}$$

so that $B = \big[c\,R_{u,-\theta}\big]$, which is not quite what we need. Instead we must choose $u = \big(0,-2,-1\big)^T$ (or some positive multiple of that) in order to get $B = \big[c\,R_{u,\theta}\big]$.

**3:** Find a singular value decomposition $Q^*AP = S$ for the matrix $A = \begin{pmatrix} 1 & 1 \\ 2 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix} \in M_{4\times 2}(\mathbb{R})$.

Solution: We have

$$A^*A = \begin{pmatrix} 6 & 2 \\ 2 & 3 \end{pmatrix} .$$

The characteristic polynomial of $A^*A$ is

$$f_{A^*A}(x) = \left|A^*A - xI\right| = \begin{vmatrix} 6-x & 2 \\ 2 & 3-x \end{vmatrix} = x^2 - 9x + 14 = (t-7)(t-2)$$

so the eigenvalues of $A^*A$ are $\lambda_1 = 7$, $\lambda_2 = 2$, and hence the singular values of $A$ are $\sigma_1 = \sqrt{7}$, $\sigma_2 = \sqrt{2}$, so we can take

$$S = \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \sqrt{7} & 0 \\ 0 & \sqrt{2} \\ 0 & 0 \\ 0 & 0 \end{pmatrix} .$$

When $\lambda = \lambda_1 = 7$ we have

$$A^*A - \lambda I = \begin{pmatrix} -1 & 2 \\ 2 & -4 \end{pmatrix} \sim \begin{pmatrix} 1 & -2 \\ 0 & 0 \end{pmatrix}$$

so $u_1 = \frac{1}{\sqrt{5}}\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ is a unit eigenvector for $\lambda_1$. The eigenspace for $\lambda_2$ is orthogonal so $u_2 = \frac{1}{\sqrt{5}}\begin{pmatrix} -1 \\ 2 \end{pmatrix}$ is a unit eigenvector for $\lambda_2$, and so we can take

$$P = \left(u_1, u_2\right) = \begin{pmatrix} \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix} .$$

Next let $v_1 = \frac{Au_1}{\sigma_1} = \frac{1}{\sqrt{7}\sqrt{5}}\begin{pmatrix} 1 & 1 \\ 2 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{35}}\begin{pmatrix} 3 \\ 4 \\ 1 \\ 3 \end{pmatrix}$ and $v_2 = \frac{Au_2}{\sigma_2} = \frac{1}{\sqrt{2}\sqrt{5}}\begin{pmatrix} 1 & 1 \\ 2 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} -1 \\ 2 \end{pmatrix} = \frac{1}{\sqrt{10}}\begin{pmatrix} 1 \\ -2 \\ 2 \\ 1 \end{pmatrix}$ then extend $\{v_1, v_2\}$ to an orthonormal basis $\mathcal{V} = \{v_1, v_2, v_3, v_4\}$ for $\mathbb{R}^4$. We have

$$\begin{pmatrix} 3 & 4 & 1 & 3 \\ 1 & -2 & 2 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & -2 & 2 & 1 \\ 0 & 10 & -5 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & -2 & 2 & 1 \\ 0 & 1 & -\frac{1}{2} & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & -\frac{1}{2} & 0 \end{pmatrix}$$

So the orthogonal complement of $\text{Span}\{v_1, v_2\}$ has basis $\left\{(-1,0,0,1)^T, (02,1,2,0)^T\right\}$. We apply the Gram-Schmidt Procedure, replacing the second vector in the basis by

$$(-2,1,2,0)^T - \tfrac{2}{2}(-1,0,0,1)^T = (-1,1,2,-1)^T ,$$

and then we normalize to obtain $v_3 = \frac{1}{\sqrt{2}}(-1,0,0,1)^T$ and $v_4 = \frac{1}{\sqrt{7}}(-1,1,2,-1)^T$. Thus we can take

$$Q = \left(v_1, v_2, v_3, v_4\right) = \begin{pmatrix} \frac{3}{\sqrt{35}} & \frac{1}{\sqrt{10}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{7}} \\ \frac{4}{\sqrt{35}} & -\frac{2}{\sqrt{10}} & 0 & \frac{1}{\sqrt{7}} \\ \frac{1}{\sqrt{35}} & \frac{2}{\sqrt{10}} & 0 & \frac{2}{\sqrt{7}} \\ \frac{3}{\sqrt{35}} & \frac{1}{\sqrt{10}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{7}} \end{pmatrix} .$$

**4:** Let $A \in M_n(\mathbb{C})$. Let $\lambda_1, \lambda_2, \cdots, \lambda_n$ be the eigenvalues of $A$ (listed with repetition according to algebraic multiplicity). Show that the following are equivalent.

1. $A\,A^* = A^*A$.
2. $A^* = f(A)$ for some polynomial $f$.
3. $A^* = AP$ for some unitary matrix $P$.
4. $\sum_{i,j} \left|A_{i,j}\right|^2 = \sum_i |\lambda_i|^2$.

Solution: First we show that $1 \iff 2$. Suppose first that $A^*A = A\,A^*$. Choose a unitary matrix $P$ so that $P^*AP = D = \text{diag}\big(\lambda_1, \cdots, \lambda_1, \lambda_2, \cdots, \lambda_2, \cdots, \lambda_k, \cdots, \lambda_k\big)$ where $\lambda_1, \cdots, \lambda_k$ are the distinct eigenvalues of $A$. Note that $A = P\,DP^*$ and $A^* = P\,D^*P^*$ and that $D^* = \overline{D}$. Let $f$ be the polynomial of degree at most $k$ such that $f(\lambda_i) = \overline{\lambda_i}$ for $i = 1, 2, \cdots, k$. Note that $f(D) = \overline{D}$, so we have

$$f(A) = f(PDP^*) = Pf(D)\,P^* = P\,\overline{D}\,P^* = PD^*P^* = A^*\,.$$

Conversely, for any square matrix $A$ and any polynomial $f$, it is clear that $A$ commutes with $f(A)$, indeed if $f(x) = \sum c_k x^k$ then we have $Af(A) = \sum c_k A^{k+1} = f(A)A$. Thus if $A^* = f(A)$ then $A$ commutes with $A^*$.

Next we show that $1 \iff 3$. Suppose first that $A\,A^* = A^*A$. Choose a unitary matrix $Q$ so that $Q^*AQ = D = \text{diag}(\lambda_1, \cdots, \lambda_n)$ where $\lambda_1, \cdots, \lambda_n$ are the eigenvalues of $A$ (listed with repetition according to multiplicity). Note that $A = QDQ^*$ and $A^* = QD^*Q^*$. For $k = 1, 2, \cdots, n$, let $r_k = |\lambda_k|$ and choose $\theta_k \in [0, 2\pi)$ so that $\lambda_k = r_k e^{i\,\theta_k}$. Let $E = \text{diag}\big(e^{-i\,2\theta_1}, \cdots, e^{-i\,2\theta_n}\big)$. Note that $E$ is unitary and

$$DE = \text{diag}\big(r_1 e^{i\,\theta_1}, \cdots, r_n e^{i\,\theta_n}\big)\,\text{diag}\big(e^{-i\,2\theta_1}, \cdots, e^{-i\,2\theta_n}\big) = \text{diag}\big(r_1 e^{-i\theta_1}, \cdots, r_n e^{-i\theta_n}\big) = D^*\,.$$

Let $P = QEQ^*$. Then $P$ is unitary and

$$AP = QDQ^*QEQ^* = QDEQ^* = QD^*Q^* = A^*\,.$$

Conversely, suppose that $A^* = AP$ where $P$ is unitary. Then we have

$$A^* = AP \Longrightarrow A^*P^* = APP^* = A \Longrightarrow (A^*P^*)^* = A^* \Longrightarrow PA = A^*\,.$$

Thus $A^*A = (AP)A = A(PA) = AA^*$.

Finally we show that $1 \iff 4$. Suppose first that $AA^* = A^*A$. Choose a unitary matrix $P$ so that $P^*AP = D = \text{diag}(\lambda_1, \cdots, \lambda_n)$. Recall that similar matrices have the same trace. Note that $A^*A$ is similar to $D^*D$ since $D^*D = PA^*P^*PAP^* = PA^*AP^*$. Using the standard inner product $\langle A, B\rangle = \text{trace}(B^*A)$, we have

$$\sum_{i,j} |A_{i,j}|^2 = |A|^2 = \text{trace}(A^*A) = \text{trace}(D^*D) = |D|^2 = \sum_i |\lambda_i|^2\,.$$

Conversely, suppose that $A^*A \neq AA^*$. Choose a unitary matrix $P$ so that $P^*AP = T$ is upper triangular with diagonal entries $T_{i,i} = \lambda_i$. Since $A^*A \neq AA^*$, we know that $T$ is not diagonal, so we have $T_{k,l} \neq 0$ for some $k < l$. Then

$$\sum_{i,j} |A_{i,j}|^2 = |A|^2 = |T|^2 = \sum_{i \leq j} |T_{i,j}|^2 \leq |T_{k,l}|^2 + \sum_i |T_{i,i}|^2 = |T_{k,l}|^2 + \sum_i |\lambda_i|^2 < \sum_i |\lambda_i|^2\,.$$

**5:** A matrix $A \in M_{n\times n}(\mathbb{C})$ is called Hermitian positive-definite when $A^* = A$ and the eigenvalues of $A$ are all positive. Let $H_n(\mathbb{C})$ denote the set of Hermitian positive-definite matrices in $M_{n\times n}(\mathbb{C})$.

(a) Let $A \in H_n(\mathbb{C})$. Show that if $Q^*AP = S$ is a singular value decomposition of $A$, then $Q = P$.

Solution: Let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ be the eigenvalues of $A^*A$, let $\sigma_i = \sqrt{\lambda_i}$ be the singular values of $A$, let $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_n > 0$ be the eigenvalues of $A$, and let $\mathcal{U} = \{u_1, \cdots, u_n\}$ be a corresponding basis of eigenvectors of $A$. Since $A$ is Hermitian (that is $A^* = A$) and since $Au_i = \mu_i u_i$, we have

$$A^*Au_i = AAu_i = A\mu_i u_i = \mu_i Au_i = {\mu_i}^2 u_i$$

and so each ${\mu_i}^2$ is an eigenvalue of $A^*A$ with eigenvector $u_i$. Since $A$ is positive semi-definite we must have $\mu_i = \sqrt{\lambda_i} = \sigma_i$, and so the eigenvalues of $A$ are equal to the singular values of $A$, and $A$ and $A^*A$ have the same eigenvectors.

Let $Q^*AP = S$, with $S = \mathrm{diag}(\sigma_1, \cdots, \sigma_n) = \mathrm{diag}(\mu_1, \cdots, \mu_n)$, be a singular value decomposition of $A$. Let $w_1, \cdots, w_n$ be the columns of $P$ and let $\mathcal{W} = \{w_1, \cdots, w_n\}$. Recall that $\mathcal{W}$ is a basis of eigenvectors of $A^*A$. Since $A$ and $A^*A$ have the same eigenvectors, $\mathcal{W}$ is also a basis of eigenvectors of $A$, and so it diagonalizes $A$, that is $A = PSP^*$. Thus we have $A = QSP^* = PSP^*$. Multiply on the right by $P$ and then by $S^{-1} = \mathrm{diag}\big(\frac{1}{\sigma_1}, \cdots, \frac{1}{\sigma_n}\big)$ to get $Q = P$.

(b) Show that every element of $H_n(\mathbb{C})$ has a unique square root in $H_n(\mathbb{C})$.

Solution: Let $A \in H_n$. Since $A = A^*$ we can unitarily diagonalize $A$. Choose a unitary matrix $P$ so that $P^*AP = D = \mathrm{diag}\{\lambda_1, \cdots, \lambda_n\}$ with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n > 0$ (we know the eigenvalues of $A$ are positive since $A \in H_n$). Let $E = \mathrm{diag}\big(\sqrt{\lambda_1}, \cdots, \sqrt{\lambda_n}\big)$ and note that $E^2 = D$. Let $B = PEP^*$. Note that $B \in H_n(\mathbb{C})$ since $B^* = B$ and the eigenvalues of $B$ are $\sqrt{\lambda_1}, \cdots, \sqrt{\lambda_n}$ which are positive. Also note that $B^2 = (PEP^*)^2 = PEP^*PEP^* = PEP^* = A$. Thus $B$ is a square root of $A$ in $H_n(\mathbb{C})$.

It remains to show that this square root is unique. Suppose that $C \in H_n(\mathbb{C})$ and $C^2 = A$. Since $C^* = C$ we can unitarily diagonalize $C$. Choose a unitary matrix $Q$ so that $Q^*CQ = F = \mathrm{diag}\big(\sigma_1, \cdots, \sigma_n\big)$ with $\sigma_1 \geq \cdots \geq \sigma_n > 0$ (each $\sigma_i > 0$ since $C \in H_n(\mathbb{C})$) and let $v_1, \cdots, v_n$ be the columns of $Q$. For each $i$ we have

$$Au_i = C^2u_i = C(Cu_i) = C(\sigma_i u_i) = \sigma_i Cu_i = {\sigma_i}^2 u_i$$

so the eigenvalues of $C$ are the square roots of the eigenvalues of $A$ and each eigenvector of $C$ is also an eigenvector for $A$. It follows that $\sigma_i = \sqrt{\lambda_i}$ for each $i$, and that $\{v_1, \cdots, v_n\}$ is both a basis of eigenvectors for $C$ and a basis of eigenvectors for $A$, so the matrices $C$ and $A$ have the same eigenvectors. Similarly the above matrix $B$ also has the same eigenvectors. It follows that the unitary matrix $Q$ which we used to diagonalize $C$ can also be used to diagonalize $B$. Thus we have

$$Q^*CQ = \mathrm{diag}(\sqrt{\lambda_1}, \cdots, \sqrt{\lambda_n}) = Q^*BQ\,.$$

Multiplying on the left by $Q$ and on the right by $Q^*$ gives $C = B$.

(c) Show that for every $A \in GL_n(\mathbb{C})$ there exist unique $R \in H_n(\mathbb{C})$ and $\Theta \in U_n(\mathbb{C})$ such that $A = R\Theta$.

Solution: By the Singular Value Decomposition Theorem, we can choose unitary matrices $P$ and $Q$ so that $Q^*AP = D = \mathrm{diag}(\sigma_1, \cdots, \sigma_n)$ with $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n > 0$ (so the $\sigma_i$ are the singular values of $A$, and they are strictly positive since $A$ is invertible). Then we have

$$A = PDQ^* = PDP^*PQ^* = R\Theta \text{ with } R = PDP^* \text{ and } \Theta = PQ^*\,.$$

Note that $\Theta \in U_n(\mathbb{C})$ since $\Theta^*\Theta = (PQ^*)^*(PQ^*) = QP^*PQ^* = QIQ^* = QQ^* = I$, and $R \in H_n(\mathbb{C})$ since $R^* = (PDP^*)^* = PD^*P^* = PDP^* = R$ and since $R$ is similar to $D$ so that the eigenvalues of $R$ are the singular values $\sigma_i$ which are positive.

It remains to show that $R$ and $\Theta$ are unique. Suppose that $A = R\Theta = S\Phi$ where $R, S \in H_n(\mathbb{C})$ and $\Theta, \Phi \in U_n(\mathbb{C})$. Note that $R$ and $S$ are invertible with $R^{-1} = \Theta A^{-1}$ and $S^{-1} = \Phi A^{-1}$. Since $R\Theta = S\Phi$, we have $SR^{-1} = \Theta\Phi^{-1}$. Since $\Theta\Phi^{-1}$ is unitary, we see that $R^{-1}S$ is unitary, so we have

$$I = (R^{-1}S)(R^{-1}S)^* = R^{-1}SS^*{R^{-1}}^* = R^{-1}S^2R^{-1}$$

Multiply on the left and right by $R$ to get $R^2 = S^2$. By the uniqueness of square roots proven in Part (b), we have $R = S$. Since $R\Theta = S\Phi$ and $R = S$ with $R$ and $S$ invertible, we also have $\Theta = \Phi$.

**6:** Let $U$ be a finite-dimensional inner product space over $\mathbb{R}$ and let $L : U \to U$ be linear. Suppose $L^*L = L\,L^*$. Show that there is an orthonormal basis $\mathcal{U}$ for $U$ such that $[L]_\mathcal{U}$ is in the block-diagonal form

$$[L]_\mathcal{U} = \begin{pmatrix} \lambda_1 &&&&&&& \\ & \ddots &&&&&& \\ && \lambda_k &&&&& \\ &&& a_1 & b_1 &&& \\ &&& -b_1 & a_1 &&& \\ &&&&& \ddots && \\ &&&&&& a_l & b_l \\ &&&&&& -b_l & a_l \end{pmatrix}$$

where each $1 \times 1$ block corresponds to a real eigenvalue $\lambda_j$ of $L$, and each $2 \times 2$ block corresponds to a pair of conjugate complex eigenvalues $a_j \pm i\,b_j$.

Solution 1: Choose a basis $\mathcal{U}_0$ for $U$ and let $A = \big[L\big]_{\mathcal{U}_0} \in M_{n\times n}(\mathbb{R})$. Since $L^*L = L\,L^*$, we have $A^*A = A\,A^*$ (since $A$ is real, this is equivalent to $A^TA = A\,A^T$). We need to show that there exists a real orthogonal matrix $P$ such that $P^*A\,P$ is in the required block-diagonal form. Since $A^*A = A\,A^*$, we know that $A$ can be unitarily diagonalized over $\mathbb{C}$. Choose an orthonormal basis of complex eigenvectors for $\mathbb{C}^n$, say

$$\big\{v_1, \cdots, v_k, w_1, \cdots, w_l, z_1, \cdots, z_m\big\}$$

where each $v_i$ is a complex eigenvector for a real eigenvalue $\lambda_i$, each $w_\alpha$ is a complex eigenvector for a complex eigenvalue $a_\alpha + i\,b_\alpha$ with $b_\alpha > 0$, and each $z_\beta$ is a complex eigenvector for a complex eigenvalue $a_\beta - i\,b_\beta$ with $b_\beta > 0$. We make several preliminary remarks.

First we remark that the eigenvectors $v_i$ can be chosen to be real vectors. This is because for a real eigenvalue $\lambda$, when we use our standard procedure to find a basis for the eigenspace $E_\lambda = \text{Null}(A - \lambda I)$ by reducing the matrix $A - \lambda I$, we obtain a basis of real vectors, and when we apply the Gram Schmidt procedure, the vectors remain real.

Next we remark that we must have $m = l$ because the characteristic polynomial $f_A$ is a real polynomial, so each complex root $\mu = a + i\,b$ occurs with the same algebraic multiplicity as its conjugate $\overline{\mu} = a - i\,b$. and since $A$ is diagonalizable, the geometric and algebraic multiplicities are equal, so $\dim E_\lambda = \dim E_{\overline{\lambda}}$.

Finally, we remark that we can choose to have $z_\alpha = \overline{w}_\alpha$ for each $\alpha = 1, \cdots, l$. To see this, suppose that $\{w_1, \cdots, w_r\}$ is an orthonormal basis for the eigenspace of $\mu = a + i\,b$. We claim that $\{\overline{w}_1, \cdots, \overline{w}_r\}$ is an orthonormal basis for the eigenspace of $\overline{\mu}$. Since $A\overline{w}_\alpha = \overline{Aw_\alpha} = \overline{\mu w_\alpha} = \overline{\mu}\,\overline{w}_\alpha$, we see that each $\overline{w}_\alpha$ does lie in the eigenspace $E_{\overline{\mu}}$. Since $\langle \overline{w}_\alpha, \overline{w}_\beta\rangle = \overline{\langle w_\alpha, w_\beta\rangle} = \overline{\delta_{\alpha,\beta}} = \delta_{\alpha,\beta}$, we see that $\{\overline{w}_1, \cdots, \overline{w}_r\}$ is orthonormal. Since $\dim E_\mu = \dim E_{\overline{\mu}}$, it follows that $\{\overline{w}_1, \cdots, \overline{w}_r\}$ is an orthonormal basis for $E_{\overline{\mu}}$, as claimed.

From these remarks, it follows that we can choose an orthonormal basis of complex eigenvectors

$$\big\{v_1, \cdots, v_k, w_1, \overline{w}_1, \cdots, w_l, \overline{w}_l\big\}$$

for $\mathbb{C}^n$ such that each $v_i$ is a real eigenvector for a real eigenvalue $\lambda_i$ and each $w_\alpha$ is a complex eigenvector for a complex eigenvalue $\mu_\alpha = a + \alpha + i\,b_\alpha$ with $b_\alpha > 0$.

For each $\alpha = 1, \cdots, k$, write $w_\alpha = x_\alpha + i\,y_\alpha$ with $x_\alpha, y_\alpha \in \mathbb{R}^n$. Let

$$\mathcal{V} = \big\{v_1, \cdots, v_k, \sqrt{2}\,x_1, \sqrt{2}\,y_1, \cdots, \sqrt{2}\,x_n, \sqrt{2}\,y_n\big\}$$

and let $P$ be the matrix whose columns are the vectors in $\mathcal{V}$. We claim that $\mathcal{V}$ is orthonormal, or equivalently that $P^*P = I$, and that $P^*A\,P$ is of the required block-diagonal form.

First let us show that $P^*A\,P$ is of the required form. Since $Av_i = \lambda_i v_i$ for $1 \le i \le k$, it follows that the first $k$ columns of $A$ are of the required form. For $1 \le \alpha \le l$, since $Aw_\alpha = \mu_\alpha w_\alpha$ we have

$$A(x_\alpha + i\,y_\alpha) = (a_\alpha + i\,b_\alpha)(x_\alpha + i\,y_\alpha) = (a_\alpha x_\alpha - b_\alpha y_\alpha) + i(a_\alpha y_\alpha + b_\alpha x_\alpha)\,.$$

Equating real and imaginary parts gives

$$Ax_\alpha = a_\alpha x_\alpha - b_\alpha y_\alpha\ ,\quad Ay_\alpha = b_\alpha x_\alpha + b_\alpha y_\alpha\,.$$

After scaling by $\sqrt{2}$ we see that the remaining columns of $P^*A\,P$ also have the required form.

Finally, it remains to show that $\mathcal{V}$ is orthonormal. Since $\langle v_i, v_j\rangle = \delta_{i,j}$, we have $v_i \cdot v_j = \langle v_i, v_j\rangle = \delta_{i,j}$. Since $\langle v_i, w_\alpha\rangle = 0$ we have $0 = \langle v_i, w_\alpha\rangle = v_i \cdot \overline{w}_\alpha = v_i \cdot (x_\alpha - i\,y_\alpha) = v_i \cdot x_\alpha - i\,v_i \cdot y_\alpha$. Equating real and imaginary parts gives $v_i \cdot x_\alpha = v_i \cdot y_\alpha = 0$. Since $\langle w_\alpha, \overline{w}_\alpha\rangle = 0$ we have $0 = \langle w_\alpha, \overline{w}_\alpha\rangle = w_\alpha \cdot w_\alpha = (x_\alpha + i\,y_\alpha)\cdot(x_\alpha + i\,y_\alpha) = (x_\alpha \cdot x_\alpha - y_\alpha \cdot y_\alpha) + 2i\,(x_\alpha \cdot y_\alpha)$. Equating real and imaginary parts gives $x_\alpha \cdot y_\alpha = 0$ and $|x_\alpha|^2 = |y_\alpha|^2$. Since $\langle w_\alpha, w_\alpha\rangle = 1$ we also have $|x_\alpha|^2 + |y_\alpha|^2 = 1$, and so we have $|x_\alpha|^2 = |y_\alpha|^2 = \frac{1}{2}$ and hence $|\sqrt{2}\,x_\alpha|^2 = |\sqrt{2}\,y_\alpha|^2 = 1$. Finally, since $\langle w_\alpha, w_\beta\rangle = 0 = \langle w_\alpha, \overline{w}_\beta\rangle$, a similar calculation shows that $x_\alpha \cdot x_\beta = x_\alpha \cdot y_\beta = y_\alpha \cdot x_\beta = y_\alpha \cdot y_\beta = 0$.

Solution 2: Choose an orthonormal basis $\mathcal{U}_0$ for $U$ and let $A = \big[L\big]_{\mathcal{U}_0} \in M_{n\times n}(\mathbb{R})$. Since $L^*L = L\,L^*$ we have $A^*A = A\,A^*$. We must show that $A$ can be put into the required block-diagonal form using an orthogonal change of basis. We shall use induction on $n$. When $n = 1$, $A$ is in diagonal form. Let $n = 2$, say $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Note that $A^*A = \begin{pmatrix} a^2+c^2 & ab+cd \\ ab+cd & b^2+d^2 \end{pmatrix}$ and $A\,A^* = \begin{pmatrix} a^2+b^2 & ac+bd \\ ac+bd & c^2+d^2 \end{pmatrix}$, so the condition that $A^*A = A\,A^*$ gives $a^2+c^2 = a^2+b^2$ so that $c = \pm b$, and $ab+cd = ac+bd$ so that $(a-d)(b-c) = 0$. Thus either $c = b$, in which case $A$ is symmetric and hence orthogonally diagonalizable, or $c = -b$ and $a = d$, in which case $A$ is in the form $A = \begin{pmatrix} a & b \\ -b & d \end{pmatrix}$.

Let $n \geq 3$, suppose, inductively, that for $1 \leq k < n$ every $k\times k$ matrix $B$ can be put into the required block-diagonal form using an orthogonal change of basis, and let $A \in M_{n\times n}(\mathbb{R})$ with $A^*A = A\,A^*$. Suppose first that $A$ has a real eigenvalue $\lambda_1$, and let $u_1$ be a corresponding unit eigenvector. Note that since $A^*A = A\,A^*$, we have $Au_1 = \lambda_1 u_1 \iff A^*u_1 = \overline{\lambda}u_1$ (see Theorem 6.15 in section 6.4 of the text). Extend $\{u_1\}$ to an orthonormal basis $\{u_1,\cdots,u_n\}$ for $\mathbb{R}^n$, and let $P = (u_1,\cdots,u_n) \in M_{n\times n}(\mathbb{R})$. Consider the matrix $P^*A\,P$. For $j \geq 1$ we have

$$\begin{aligned}
\big(P^*A\,P\big)_{1,1} &= {u_1}^*Au_1 = {u_1}^*\lambda_1 u_1 = \lambda_1(u_1 \bullet u_1) = \lambda_1 \\
\big(P^*A\,P\big)_{j,1} &= {u_j}^*A\,u_1 = {u_j}^*\lambda_1 u_1 = \lambda_1(u_1 \bullet u_j) = 0 \\
\big(P^*A\,P\big)_{1,j} &= {u_1}^*Au_j = (A^*u_1)^*u_j = (\overline{\lambda}_1 u_1)^*u_j = \lambda_1(u_j \bullet u_1) = 0
\end{aligned}$$

and so the matrix $P^*A\,P$ is in the form

$$P^*A\,P = \begin{pmatrix} \lambda_1 & 0 \\ 0 & B_1 \end{pmatrix}$$

where $B_1 = (u_2,\cdots,u_n)^*A\,(u_1,\cdots,u_n)$. Since $A$ commutes with its adjoint (or transpose) it follows that $P^*A\,P$ commutes with its adjoint, and hence $B_1$ commutes with its adjoint. By the induction hypothesis, we can put $B_1$ into the required block diagonal form.

Now suppose that $A$ does not have any real eigenvalues. Let $\lambda = a + i\,b$ be a complex eigenvalue of $A$ and let $w = x + i\,y$ be a corresponding complex eigenvector, where $x, y \in \mathbb{R}^n$. Since $Aw = \lambda w$ we have

$$A(x+i\,y) = (a+i\,b)(x+i\,y) = (ax-by) + i(ay+bx)\,.$$

Equating the real and imaginary parts gives

$$Ax = ax - by\;, \quad Ay = ay + bx\,.$$

Also, since $A^*A = A\,A^*$ we have $A^*w = \overline{\lambda}w$, so $A^*(x+i\,y) = (a-i\,b)(x+i\,y) = (ax+by) + i(ay-bx)$, hence

$$A^*x = ax + by\;, \quad A^*y = ay - bx\,.$$

This shows that $A$ and $A^*$ both map $\text{Span}\{x,y\}$ to itself. We also note that $\{x,y\}$ must be linearly independent, since otherwise $A$ would have a real eigenvector (indeed if say $y = cx$ then we would have $w = x+i\,y = x+i\,cx = (1+ic)x$, so $x = \frac{1}{1+ic}\,w$ would be a real eigenvector) and hence $A$ would have a real eigenvalue (indeed if we had $0 \neq w \in \mathbb{R}^n$ then we would have $Aw \in \mathbb{R}^n$ and so $Aw = \lambda w$ would imply that $\lambda \in \mathbb{R}$). Let $\{u_1,u_2\}$ be an orthonormal basis for $\text{Span}\{x,y\}$. Extend $\{u_1,u_2\}$ to an orthonormal basis $\{u_1,\cdots,u_n\}$ for $\mathbb{R}^n$ and let $P = (u_1,\cdots,u_n) \in M_{n\times n}(\mathbb{R})$. Consider the matrix $P^*A\,P$. For $i = 1,2$ and $j = 3,4,\cdots,n$ we have

$$\big(P^*A\,P\big)_{j,i} = {u_j}^*A\,u_i = (Au_i) \bullet u_j = 0$$

since $Au_i \in \text{Span}\{u_1,u_2\}$ and $u_j \in \text{Span}\{u_1,u_2\}^\perp$, and

$$\big(P^*A\,P\big)_{i,j} = {u_i}^*A\,u_j = (A^*u_i)^*u_j = u_j \bullet (A^*u_i) = 0$$

since $A^*u_i \in \text{Span}\{u_1,u_2\}$ and $u_j \in \text{Span}\{u_1,u_2\}^\perp$. Thus we see that $P^*A\,P$ is in the form

$$P^*A\,P = \begin{pmatrix} A_1 & 0 \\ 0 & B_1 \end{pmatrix}$$

where $A_1 = (u_1,u_2)^*A\,(u_1,u_2) \in M_{2\times 2}(\mathbb{R})$ and $B_1 = (u_3,\cdots,u_n)^*A\,(u_3,\cdots,u_n) \in M_{(n-2)\times(n-2)}(\mathbb{R})$. Note that since $A$ commutes with its adjoint, so do $A_1$ and $B_1$. Since $A_1$ is a $2\times 2$ matrix with $A^*A = A\,A^*$, the first paragraph shows that $A$ is of the form $A_1 = \begin{pmatrix} a_1 & b_1 \\ -b_1 & a_1 \end{pmatrix}$. Since $B_1$ is an $(n-2)\times(n-2)$ matrix with ${B_1}^*B_1 = B_1\,{B_1}^*$, the induction hypothesis ensures that we can put $B_1$ into the required block-diagonal form.