Solution to 2131 HW8

Ex. 7.47(4)

We introduce

$$\vec{v}_n = \begin{pmatrix} x_n \\ y_n \\ z_n \end{pmatrix}, \quad \vec{v}_0 = \begin{pmatrix} a \\ b \\ c \end{pmatrix}, \quad \vec{v}_n = \begin{pmatrix} x_{n-1} + 3y_{n-1} - 3z_{n-1} \\ -3x_{n-1} + 7y_{n-1} - 3z_{n-1} \\ -6x_{n-1} + 6y_{n-1} - 2z_{n-1} \end{pmatrix} = A\vec{v}_{n-1}, \quad A = \begin{pmatrix} 1 & 3 & -3 \\ -3 & 7 & -3 \\ -6 & 6 & -2 \end{pmatrix}.$$

Then we have $\vec{v}_n = A^n \vec{x}_0$.

The characteristic polynomial

$$\det(tI - A) = \det \begin{pmatrix} t-1 & -3 & 3 \\ 3 & t-7 & 3 \\ 6 & -6 & t+2 \end{pmatrix} = (t-4)^2(t+2)$$

has two roots

$$\lambda_1 = 4, \quad \lambda_2 = -2.$$

By finding the null space of the corresponding matrices, we have

$$\mathrm{Nul}(A - 4I) = \mathbb{R}(1,1,0) \oplus \mathbb{R}(-1,0,1), \quad \mathrm{Nul}(A + 2I) = \mathbb{R}(1,1,2).$$

To find $\vec{v}_n$, we decompose $\vec{v}_0$ according to the basis of eigenvectors

$$\vec{v}_0 = \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \frac{3b - a - c}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + (b - a) \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} + \frac{a + c - b}{2} \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}.$$

Then

$$\begin{aligned}
\vec{v}_n &= \frac{3b - a - c}{2} A^n \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + (b - a)A^n \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} + \frac{a + c - b}{2} A^n \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \\
&= \frac{3b - a - c}{2} 4^n \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + (b - a)4^n \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} + \frac{a + c - b}{2} (-2)^n \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \\
&= \begin{pmatrix} \big(2^{2n-1} + (-1)^n 2^{n-1}\big)a + \big(2^{2n-1} + (-2)^{n-1}\big)b - \big(2^{2n-1} + (-2)^{n-1}\big)c \\ \big(2^{2n-1} + (-2)^{n-1}\big)a + \big(3 \times 2^{2n-1} + (-2)^{n-1}\big)b - \big(2^{2n-1} + (-2)^{n-1}\big)c \\ \big((-2)^n - 4^n\big)a + \big(4^n + (-2)^{n-1}\big)b + (-2)^n c \end{pmatrix}.
\end{aligned}$$

Therefore, we conclude that

$$\begin{aligned}
x_n &= \big(2^{2n-1} + (-1)^n 2^{n-1}\big)a + \big(2^{2n-1} + (-2)^{n-1}\big)b - \big(2^{2n-1} + (-2)^{n-1}\big)c, \\
y_n &= \big(2^{2n-1} + (-2)^{n-1}\big)a + \big(3 \times 2^{2n-1} + (-2)^{n-1}\big)b - \big(2^{2n-1} + (-2)^{n-1}\big)c, \\
z_n &= \big((-2)^n - 4^n\big)a + \big(4^n + (-2)^{n-1}\big)b + (-2)^n c.
\end{aligned}$$

Ex. 7.48

The recursive relation should be noted as

$$x_k = a_{n-1}x_{k-1} + a_{n-2}x_{k-2} + \cdots + a_1 x_{k-n+1} + a_0 x_{k-n}$$

for any integer $k \geq n$. Inspired by Example 7.1.5, we introduce

$$\vec{x}_k = \begin{pmatrix} x_k \\ x_{k+1} \\ \vdots \\ x_{k+n-1} \end{pmatrix}, \quad \vec{x}_0 = \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_{n-1} \end{pmatrix}, \quad \vec{x}_{k+1} = A\vec{x}_k,$$

and we can easily find out that A is

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & \cdots & 0 & a_0 \\ 1 & 0 & 0 & 0 & \cdots & 0 & a_1 \\ 0 & 1 & 0 & 0 & \cdots & 0 & a_2 \\ 0 & 0 & 1 & 0 & \cdots & 0 & a_3 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & a_{n-2} \\ 0 & 0 & 0 & 0 & \cdots & 1 & a_{n-1} \end{pmatrix}^T.$$

Exercise 7.34 gives us that the characteristic polynomial of A is exactly $t^n - a_{n-1}t^{n-1} - \cdots - a_1 t - a_0$, thus the eigenvalues of A are $\lambda_1, \lambda_2, \ldots, \lambda_n$. To find x_k, we have

$$\vec{x}_k = A^k \vec{x}_0 = \lambda_1^k \vec{v}_1 + \lambda_2^k \vec{v}_2 + \cdots + \lambda_n^k \vec{v}_n$$

where we decompose $\vec{x}_0$ into eigenvectors $\vec{v}_i$ corresponding to their eigenvalues. Picking the first coordinate of $\vec{x}_k$, we conclude that

$$x_k = c_1\lambda_1^k + c_2\lambda_2^k + \cdots + c_n\lambda_n^k.$$

Here c_i is the first coordinate of $\vec{v}_i$, which can be calculated from $x_0, x_1, \ldots, x_{n-1}$.

The result can indeed be applied to Fibonacci numbers, by choosing $n = 2$ and $a_1 = a_0 = 1$. From Example 7.1.17, the result is in the form of $c_1\lambda_1^k + c_2\lambda_2^k$

Ex. 7.51

Note $L_1 : H_1 \to H_1$ and $L_2 : H_2 \to H_2$, then for any $\vec{h}_1, \vec{h}_1' \in H_1$ and $\vec{h}_2, \vec{h}_2' \in H_2$, we have

$$\langle (L_1 \perp L_2)(\vec{h}_1 + \vec{h}_2), \vec{h}_1' + \vec{h}_2' \rangle = \langle \vec{h}_1 + \vec{h}_2, (L_1 \perp L_2)^*(\vec{h}_1' + \vec{h}_2') \rangle.$$

On the other hand, we have

$$\begin{aligned} \langle (L_1 \perp L_2)(\vec{h}_1 + \vec{h}_2), \vec{h}_1' + \vec{h}_2' \rangle &= \langle L_1(\vec{h}_1) + L_2(\vec{h}_2), \vec{h}_1' + \vec{h}_2' \rangle \\ &= \langle L_1(\vec{h}_1), \vec{h}_1' \rangle + \langle L_1(\vec{h}_1), \vec{h}_2' \rangle + \langle L_2(\vec{h}_1), \vec{h}_1' \rangle + \langle L_2(\vec{h}_1), \vec{h}_2' \rangle \\ &= \langle L_1(\vec{h}_1), \vec{h}_1' \rangle + 0 + 0 + \langle L_2(\vec{h}_2), \vec{h}_2' \rangle \\ &= \langle \vec{h}_1, L_1^*(\vec{h}_1') \rangle + 0 + 0 + \langle \vec{h}_2, L_2^*(\vec{h}_2') \rangle \\ &= \langle \vec{h}_1, L_1^*(\vec{h}_1') \rangle + \langle \vec{h}_2, L_1^*(\vec{h}_1') \rangle + \langle \vec{h}_1, L_2^*(\vec{h}_2') \rangle + \langle \vec{h}_2, L_2^*(\vec{h}_1') \rangle \\ &= \langle \vec{h}_1 + \vec{h}_2, L_1^*(\vec{h}_1') + L_2^*(\vec{h}_2') \rangle. \end{aligned}$$

By comparison, since all the vectors above are arbitrary, we conclude that

$$(L_1 \perp L_2)^* = L_1^* \perp L_2^*.$$

Ex. 7.53

Since $(L^*)^* = L$, we clearly have

$$L^*L = LL^* \iff L^*(L^*)^* = (L^*)^*L^*.$$

This means 1. and 2. are equivalent.

We have

$$\begin{aligned} ||L(\vec{v})|| = ||L^*(\vec{v})||,\ \forall \vec{v} &\iff \langle L(\vec{v}), L(\vec{v})\rangle = \langle L^*(\vec{v}), L^*(\vec{v})\rangle,\ \forall \vec{v} \\ &\iff \langle \vec{v}, L^*L(\vec{v})\rangle = \langle \vec{v}), LL^*(\vec{v})\rangle,\ \forall \vec{v} \\ &\iff \langle \vec{v}, (L^*L - LL^*)(\vec{v})\rangle = 0,\ \forall \vec{v} \\ &\iff (L^*L - LL^*)(\vec{v}) = \vec{0},\ \forall \vec{v} \\ &\iff L^*L - LL^* = O \iff L^*L = LL^*. \end{aligned}$$

This means 1. and 3. are equivalent.

We have

$$\begin{aligned} L^*L = (L_1^* + L_2^*)(L_1 + L_2) = L_1^*L_1 + L_1^*L_2 + L_2^*L_1 + L_2^*L_2 = L_1^2 + L_1L_2 - L_2L_1 - L_2^2, \\ LL^*(L_1 + L_2)(L_1^* + L_2^*) = L_1L_1^* + L_1L_2^* + L_2L_1^* + L_2L_2^* = L_1^2 - L_1L_2 + L_2L_1 - L_2^2. \end{aligned}$$

Then it is easy to verify that

$$L_1L_2 = L_2L_1 \iff L^*L = LL^*.$$

This means 1. and 4. are equivalent.

In conclusion, all four statements are equivalent.

Ex. 7.55

1. Exercise 6.24 gives that

$$(\mathrm{Ker}\, L)^\perp = \mathrm{Ran}\, L^*, \quad (\mathrm{Ker}\, L^*)^\perp = \mathrm{Ran}\, L$$

and Exercise 7.54 gives that

$$\mathrm{Ker}\, L = \mathrm{Ker}\, L^*.$$

We then clearly have

$$\mathrm{Ran}\, L = (\mathrm{Ker}\, L^*)^\perp = (\mathrm{Ker}\, L)^\perp = \mathrm{Ran}\, L^*.$$

2. Proposition 4.3.6 remains valid in complex inner product spaces with Hermitian inner product. By the proposition, we easy derive that

$$V = \mathrm{Ker}\, L \perp (\mathrm{Ker}\, L)^\perp = \mathrm{Ker}\, L \perp \mathrm{Ran}\, L.$$

3. For any integer $k > 1$ and any vector $\vec{v} \in V$, since L is an linear operator, L^{k-1} is also a linear operator and we have $L^{k-1}(\vec{v}) \in V$. Therefore

$$L^k(\vec{v}) = L\big(L^{k-1}(\vec{v})\big) \in \mathrm{Ran}\, L.$$

Since $\vec{v}$ is arbitrary, by the definition of range, we conclude that

$$\mathrm{Ran}\, L^k = \mathrm{Ran}\, L.$$

Since L^k is still a linear operator, we use 2. to find out that

$$\operatorname{Ker} L^k = \operatorname{Ker} L.$$

4. Since L^* is a linear operator on V, we have $(L^*)^k$ is also a linear operator, thus for any vector $\vec{v} \in V$ we have

$$L^j (L^*)^k (\vec{v}) = L^j \big((L^*)^k (\vec{v}) \big) \in \operatorname{Ran} L^j$$

and therefore

$$\operatorname{Ran} L^j (L^*) = \operatorname{Ran} L^j = \operatorname{Ran} L.$$

Since $L^j (L^*)^k$ is still a linear operator, we use 2. and 3. to find out that

$$\operatorname{Ker} L^j (L^*)^k = \operatorname{Ker} L.$$

Ex. 7.56

L is orthogonally diagonalisable and

$$L = \lambda_1 I \perp \lambda_2 I \perp \cdots \lambda_k I$$

with respect to

$$V = H_1 \perp H_2 \perp \cdots \perp H_k.$$

Moreover, we have

$$L^* = \bar{\lambda}_1 I \perp \bar{\lambda}_2 I \perp \cdots \bar{\lambda}_k I.$$

Exercise 7.56 gives us that H is in the form of

$$H = W_1 \perp W_2 \perp \cdots \perp W_k$$

where $W_i \subset H_i$. For any vector $\vec{h} \in H$, we decompose it into

$$\vec{h} = \vec{w}_1 + \vec{w}_2 + \cdots + \vec{w}_k$$

and then

$$L^*(\vec{h}) = \bar{\lambda}_1 \vec{w}_1 + \bar{\lambda}_2 \vec{w}_2 + \cdots + \bar{\lambda}_k \vec{w}_k \in W_1 \perp W_2 \perp \cdots \perp W_k = H.$$

Therefore, we conclude that H is an invariant subspace of L^*. By Proposition 4.3.6, we can easily deduce that $H^\perp$ is in the form of

$$H^\perp = U_1 \perp U_2 \perp \cdots \perp U_k$$

where $U_i \subset H_i$ and $W_i \perp U_i = H_i$. Then similarly we conclude that $H^\perp$ is an invariant subspace of L and L^*.

Ex. 7.57

1. Note $L : V \to V$. Notice that

$$(I - P)LP = O \iff LP = PLP.$$

For any vector $\vec{v} \in V$, we note $P(\vec{v}) = \vec{h} \in H$, then

$$H \text{ is an invariant subspace of } L \iff L(\vec{h}) \in H \iff LP(\vec{v}) \in H \iff PLP(\vec{v}) = P\big(LP(\vec{v})\big) = LP(\vec{v}).$$

Therefore, we conclude that H is an variant subspace of L if and only if $(I - P)LP = O$.

2. We have

$$X = PL(I - P) = PL - PLP \implies X^* = L^*P^* - P^*L^*P^* = L^*P - PL^*P = (I - P)L^*P,$$

and

$$XX^* = PL(I-P)(I-P)L^*P = PL(I-P)L^*P = PLL^*P - PLPL^*P.$$

From 1. we know that $LP = PLP$, and we also notice that $P^2 = P$, $\operatorname{tr} AB = \operatorname{tr} BA$ and $\operatorname{tr}(A-B) = \operatorname{tr} A - \operatorname{tr} B$. Using all these equalities, we have

$$\begin{aligned}\operatorname{tr} XX^* &= \operatorname{tr}(PLL^*P - PLPL^*P) = \operatorname{tr}(PLL^*)P - \operatorname{tr}(PLPL^*)P = \operatorname{tr} P^2LL^* - \operatorname{tr} P^2LPL^* \\ &= \operatorname{tr} PLL^* - \operatorname{tr}(PLP)L^* = \operatorname{tr} PLL^* - \operatorname{tr}(LP)L^* = \operatorname{tr} PLL^* - \operatorname{tr} L(PL^*) \\ &= \operatorname{tr} PLL^* - \operatorname{tr} PL^*L = \operatorname{tr} P(LL^* - L^*L) = \operatorname{tr} PO = \operatorname{tr} O = 0.\end{aligned}$$

3. It is known that (by the positivity of Hermitian inner product)

$$\operatorname{tr} XX^* = 0 \iff X = O.$$

By Proposition 4.3.7, we know that $P' = I - P$ is the orthogonal projection to $H^\perp$. Then from 1. we have

$$X = PL(I-P) = (I-P')LP' = O \iff H^\perp \text{ is an invariant subspace of } L.$$

Ex. 7.60

Exercise 7.59 gives us that L is Hermitian if and only if

$$\langle L(\vec{v}), \vec{v}\rangle = \langle \vec{v}, L(\vec{v})\rangle = \overline{\langle L(\vec{v}), \vec{v}\rangle},\ \forall \vec{v} \iff \langle L(\vec{v}), \vec{v}\rangle \in \mathbb{R}$$

and we derive the result.

Ex. 7.61

A Hermitian operator L is of course orthogonally diagonalisable. Suppose $L = \lambda_1 I \perp \lambda_2 I \perp \cdots \lambda_k I$ with respect to $V = H_1 \perp H_2 \perp \cdots \perp H_k$. Then Exercise 7.17 gives that

$$\det L = \lambda_1^{\dim H_1} \lambda_2^{\dim H_2} \cdots \lambda_k^{\dim H_k}.$$

Since we have

$$L^* = \bar{\lambda}_1 I \perp \bar{\lambda}_2 I \perp \cdots \bar{\lambda}_k I,$$

we clearly derive that

$$L = L^* \implies \lambda_i = \bar{\lambda}_i\ \forall i. \implies \lambda_i \in \mathbb{R}\ \forall i.$$

Therefore, we conclude that

$$\det L = \lambda_1^{\dim H_1} \lambda_2^{\dim H_2} \cdots \lambda_k^{\dim H_k} \in \mathbb{R}.$$

Ex. 7.63

By proposition 7.2.7, we have the matrix A has orthogonal diagonalisation

$$A = UDU^{-1} = UDU^T$$

where

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}, \quad U = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & x \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & y \\ 0 & 0 & z \end{pmatrix}.$$

By $UU^T = I$ we have

$$\begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & x \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & y \\ 0 & 0 & z \end{pmatrix} \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0 \\ x & y & z \end{pmatrix} = \begin{pmatrix} 1+x^2 & xy & xz \\ xy & 1+y^2 & yz \\ xz & yz & z^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

We derive that

$$x = y = 0, \quad z = \pm 1.$$

Choose $z = 1$ and we conclude that

$$A = UDU^T = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \frac{3}{2} & -\frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{3}{2} & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

Ex. 7.65

By proposition 7.2.7, we have the matrix A has orthogonal diagonalisation

$$A = UDU^{-1} \implies A^3 = UD^3U^{-1},\ A^2 = UD^2U^{-1}.$$

If $A^3 = O$, then $D^3 = O$ and we easily derive that $D = O$, which means $A = O$.

If $A^2 = I$, we have

$$A^2 = UD^2U^{-1} = I \implies UD^2U^{-1}U = U \implies UD^2 = U \implies D^2 = I \implies D = \operatorname{diag}[\pm 1].$$

Ex. 7.67

We have

$$L = \frac{1}{2}(L + L^*) + \frac{1}{2}(L - L^*) = L_1 + L_2.$$

Since

$$L_1^* = \frac{1}{2}(L + L^*)^* = \frac{1}{2}\big(L^* + (L^*)^*\big) = \frac{1}{2}(L^* + L) = L_1,$$
$$L_2^* = \frac{1}{2}(L - L^*)^* = \frac{1}{2}\big(L^* - (L^*)^*\big) = \frac{1}{2}(L^* - L) = -L_2,$$

we indeed have L_1 is Hermitian and L_2 is skew Hermitian. This prove that any linear operator L can be expressed as that. Now we prove that the expression is unique. Suppose $L = L_1 + L_2 = L_1' + L_2'$ where L_1, L_1' are Hermitian and L_2, L_2' are skew Hermitian. We have

$$\left.\begin{aligned} L &= L_1 + L_2 = L_1' + L_2' \\ L^* &= L_1 - L_2 = L_1' - L_2' \end{aligned}\right\} \implies \begin{cases} 2L_1 = 2L_1' \implies L_1 = L_1' \\ 2L_2 = 2L_2' \implies L_2 = L_2' \end{cases}.$$

Hence the expression is unique. The latter result is the same as the four statement in Exercise 7.53.

Ex. 7.68

Note that L is skew-Hermitian if and only if iL is Hermitian. Suppose $iL = \lambda_1 I \perp \lambda_2 I \perp \cdots \lambda_k I$, then we have $L = -i\lambda_1 I \perp -i\lambda_2 I \perp \cdots - i\lambda_k I$ with respect to $V = H_1 \perp H_2 \perp \cdots \perp H_k$, where λ_i are real numbers. Then Exercise 7.17 gives that

$$\det L = (-i\lambda_1)^{\dim H_1}(-i\lambda_2)^{\dim H_2} \cdots (-i\lambda_k)^{\dim H_k} = (-i)^n \lambda_1^{\dim H_1}\lambda_2^{\dim H_2} \cdots \lambda_k^{\dim H_k}.$$

Here $n = \dim H_1 + \dim H_2 + \cdots + \dim H_k = \dim V$. Therefore, we conclude that if n is even, then $\det L$ is real; if n is odd, then $\det L$ is purely imaginary (here we consider 0 as a purely imaginary number).

For skew-symmetric real operator, it of course is skew-Hermitian. Also, its determinant is always real. Then if n is odd, its determinant can only be 0.

Ex. 7.71

By Proposition 7.2.9, an orthogonal operator on $\mathbb{R}^2$ is either identity, rotation or reflection. In terms of matrix, we have

$$[L]_{\alpha\alpha} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \text{ or } [L]_{\alpha\alpha} = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}.$$

Note that identity is contained in the first case.