Math 227
Suggested solutions to Homework Set 5

Problem 1. Since the set $\{\bar{u}_1, \bar{u}_2, \ldots, \bar{u}_m\}$ is a basis of V, given any $\bar{x}$ in V, we can find $\lambda_1, \lambda_2, \ldots, \lambda_m \in \mathbb{F}$ (where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$) such that

$$\bar{x} = \lambda_1 \bar{u}_1 + \lambda_2 \bar{u}_2 + \cdots + \lambda_m \bar{u}_m. \tag{1}$$

Since in addition $\{\bar{u}_1, \bar{u}_2, \ldots, \bar{u}_m\}$ is an orthonormal set, we can write

$$\begin{aligned}
\langle \bar{x}, \bar{u}_1 \rangle &= \langle \lambda_1 \bar{u}_1 + \lambda_2 \bar{u}_2 + \cdots + \lambda_m \bar{u}_m, \bar{u}_1 \rangle \\
&= \langle \lambda_1 \bar{u}_1, \bar{u}_1 \rangle + \langle \lambda_2 \bar{u}_2, \bar{u}_1 \rangle + \cdots + \langle \lambda_m \bar{u}_m, \bar{u}_1 \rangle \quad \text{(by linearity in the 1st argument)} \\
&= \lambda_1 \langle \bar{u}_1, \bar{u}_1 \rangle + \lambda_2 \langle \bar{u}_2, \bar{u}_1 \rangle + \cdots + \lambda_m \langle \bar{u}_m, \bar{u}_1 \rangle \quad \text{(again, by linearity in the 1st argument)} \\
&= \lambda_1 \cdot 1 \,+\, \lambda_2 \cdot 0 \,+ \cdots +\, \lambda_m \cdot 0 = \lambda_1.
\end{aligned}$$

Completely analogously we see that

$$\langle \bar{x}, \bar{u}_2 \rangle = \lambda_2, \quad \langle \bar{x}, \bar{u}_3 \rangle = \lambda_3,$$
$$\text{and in general, for every } 1 \leqslant i \leqslant m,\ \langle \bar{x}, \bar{u}_i \rangle = \lambda_i.$$

This also shows that the choice of $\lambda_1, \lambda_2, \ldots, \lambda_m \in \mathbb{F}$ so that (1) holds true is unique (since each λ_i has to equal the inner product $\langle \bar{x}, \bar{u}_i \rangle$, which is completely determined by the vector $\bar{x}$ and the basis vector $\bar{u}_i$); of course, observe that we already knew that the choice of coefficients has to be unique, given that the set $\{\bar{u}_1, \bar{u}_2, \ldots, \bar{u}_m\}$ is a basis of V, however in the case of an orthonormal basis we have an additional way of justifying it.

We can now write

$$\begin{aligned}
\|x\|^2 = \langle \bar{x}, \bar{x} \rangle &= \left\langle \sum_{i=1}^m \langle \bar{x}, \bar{u}_i \rangle \bar{u}_i,\ \sum_{i=1}^m \langle \bar{x}, \bar{u}_i \rangle \bar{u}_i \right\rangle \\
&= \sum_{i=1}^m \sum_{j=1}^m \big\langle \langle \bar{x}, \bar{u}_i \rangle \bar{u}_i,\ \langle \bar{x}, \bar{u}_j \rangle \bar{u}_j \big\rangle \quad \text{(by linearity in the 1st argument, and linearity (or conjugate linearity) in the 2nd argument)} \\
&= \sum_{i=1}^m \sum_{j=1}^m \Big(\langle \bar{x}, \bar{u}_i \rangle \cdot \big\langle \bar{u}_i,\ \langle \bar{x}, \bar{u}_j \rangle \bar{u}_j \big\rangle \Big) \quad \text{(by linearity in the 1st argument)} \\
&= \sum_{i=1}^m \sum_{j=1}^m \Big(\langle \bar{x}, \bar{u}_i \rangle \cdot \overline{\langle \bar{x}, \bar{u}_j \rangle} \cdot \big\langle \bar{u}_i,\ \bar{u}_j \big\rangle \Big) \quad \text{(by linearity (or conjugate linearity) in the 2nd argument)} \\
&= \sum_{i=1}^m \Big(\langle \bar{x}, \bar{u}_i \rangle \cdot \overline{\langle \bar{x}, \bar{u}_i \rangle} \cdot \big\langle \bar{u}_i,\ \bar{u}_i \big\rangle \Big) \quad \text{(since } \langle \bar{u}_i,\ \bar{u}_j \rangle = 0 \text{ when } i \neq j) \\
&= \sum_{i=1}^m \big| \langle \bar{x}, \bar{u}_i \rangle \big|^2. \quad \text{(since } \langle \bar{u}_i,\ \bar{u}_i \rangle = 1)
\end{aligned}$$

Problem 2. (i) Consider $\bar{x} \in V$. Then, since $V = S + T$, we can find $\bar{y}_1 \in S$ and $\bar{z}_1 \in T$ such that $\bar{x} = \bar{y}_1 + \bar{z}_1$.

Assume now that we can also write $\bar{x} = \bar{y}_2 + \bar{z}_2$ for some other $\bar{y}_2 \in S$ and $\bar{z}_2 \in T$. Then

$$\bar{y}_2 + \bar{z}_2 = \bar{x} = \bar{y}_1 + \bar{z}_1 \qquad \Rightarrow \qquad \bar{y}_2 - \bar{y}_1 = \bar{z}_1 - \bar{z}_2.$$

But $\bar{y}_2 - \bar{y}_1 \in S$ given that both $\bar{y}_1$ and $\bar{y}_2$ are elements of S, and S is a subspace, while similarly $\bar{z}_1 - \bar{z}_2 \in T$. Therefore,

$$\begin{gathered} \bar{y}_2 - \bar{y}_1 = \bar{z}_1 - \bar{z}_2 \in S \cap T = \{\bar{0}_V\} \\ \Rightarrow \quad \bar{y}_2 - \bar{y}_1 = \bar{0}_V = \bar{z}_1 - \bar{z}_2 \quad \Rightarrow \quad \bar{y}_2 = \bar{y}_1 \text{ and } \bar{z}_1 = \bar{z}_2. \end{gathered}$$

Thus there are unique $\bar{y}_1 \in S$ and $\bar{z}_1 \in T$ such that $\bar{x} = \bar{y}_1 + \bar{z}_1$.

(ii) Consider $\bar{x}_1, \bar{x}_2 \in V$ and $r \in \mathbb{F}$. By part (i) we know that we can find unique $\bar{y}_1, \bar{y}_2 \in S$ and $\bar{z}_1, \bar{z}_2 \in T$ such that $\bar{x}_1 = \bar{y}_1 + \bar{z}_1$ and $\bar{x}_2 = \bar{y}_2 + \bar{z}_2$.

Then

$$\bar{x}_1 + \bar{x}_2 = (\bar{y}_1 + \bar{z}_1) \;+\; (\bar{y}_2 + \bar{z}_2) = (\bar{y}_1 + \bar{y}_2) \;+\; (\bar{z}_1 + \bar{z}_2)$$

with $\bar{y}_1 + \bar{y}_2 \in S$ and $\bar{z}_1 + \bar{z}_2 \in T$, and thus by part (i) we know that the above is the unique way we can write $\bar{x}_1 + \bar{x}_2$ as a sum of an element in S and an element in T. But then, by definition of P,

$$P(\bar{x}_1 + \bar{x}_2) = \bar{y}_1 + \bar{y}_2 = P(\bar{x}_1) + P(\bar{x}_2).$$

Similarly, we can observe that

$$r \cdot \bar{x}_1 = r \cdot (\bar{y}_1 + \bar{z}_1) = r \cdot \bar{y}_1 + r \cdot \bar{z}_1$$

with $r \cdot \bar{y}_1 \in S$ and $r \cdot \bar{z}_1 \in T$, thus by part (i) we know that the above is the unique way we can write $r \cdot \bar{x}_1$ as a sum of an element in S and an element in T. It follows that

$$P(r \cdot \bar{x}_1) = r \cdot \bar{y}_1 = r \cdot P(\bar{x}_1).$$

Combining the above, we conclude that P is linear.

Next, we note that, for every $\bar{y} \in S$, we have that $\bar{y} = \bar{y} + \bar{0}_V$ is one way of writing $\bar{y}$ as a sum of an element in S and an element in T, thus it must also be the only way. This shows that

$$P(\bar{y}) = \bar{y},$$

and therefore every $\bar{y} \in S$ is an image of some element in V under P. In other words, $S \leqslant \text{Range}(P)$, and since by definition we also clearly have $\text{Range}(P) \leqslant S$, we can conclude that $\text{Range}(P) = S$.

Moreover, for every $\bar{x} \in V$, once we write $\bar{x} = \bar{y} + \bar{z}$ for some $\bar{y} \in S$ and $\bar{z} \in T$, we obtain that

$$P^2(\bar{x}) = P\big(P(\bar{x})\big) = P\big(P(\bar{y} + \bar{z})\big) = P(\bar{y}) = \bar{y} = P(\bar{x}).$$

In other words, for every $\bar{x} \in V$, we have that $P^2(\bar{x}) = P(\bar{x})$, which shows that $P^2 = P$.

We can finally observe that, if $\bar{x} \in V$ is such that

$$P(\bar{x}) = \bar{0}_V,$$

then $\bar{0}_V$ is the unique $\bar{y} \in S$ such that $\bar{x} - \bar{y} \in T$. In other words, we have that $\bar{x} = \bar{x} - \bar{0}_V \in T$. This shows that $\text{Ker}(P) \leqslant T$. Moreover, if $\bar{z} \in T$, then $\bar{z} = \bar{0}_V + \bar{z}$ is one way of writing $\bar{z}$ as a sum of an element in S and an element in T, so it should also be the only way. This then implies that

$$P(\bar{z}) = P\big(\bar{0}_V + \bar{z}\big) = \bar{0}_V,$$

which gives that $T \leqslant \text{Ker}(P)$, and finally allows us to conclude that $\text{Ker}(P) = T$.

(iii) First of all, the zero vector $\bar{0}_V$ is orthogonal to any vector in V, therefore $\bar{0}_V \in S^{\perp}$ and $S^{\perp}$ is nonempty.

We now check that $S^{\perp}$ is closed under vector addition. Let $\bar{w}_1, \bar{w}_2 \in S^{\perp}$. Then, by definition of $S^{\perp}$, we have that $\langle \bar{w}_i, \bar{y} \rangle = 0$ for every $\bar{y} \in S$ and $i = 1, 2$. But then, because of linearity in the 1st argument, we get that

$$\langle \bar{w}_1 + \bar{w}_2, \bar{y} \rangle = \langle \bar{w}_1, \bar{y} \rangle + \langle \bar{w}_2, \bar{y} \rangle = 0 + 0 = 0 \qquad \text{for every } \bar{y} \in S.$$

In other words, $\bar{w}_1 + \bar{w}_2 \in S^{\perp}$ too.

Similarly, we verify that $S^{\perp}$ is closed under scalar multiplication. Let $\bar{w} \in S^{\perp}$ and let $r \in \mathbb{F}$ (recall that here $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$). For every $\bar{y} \in S$, we have that $\langle \bar{w}, \bar{y} \rangle = 0$. But then, by linearity in the 1st argument, we get that, for every $\bar{y} \in S$,

$$\langle r \cdot \bar{w}, \bar{y} \rangle = r \cdot \langle \bar{w}, \bar{y} \rangle = r \cdot 0 = 0.$$

In other words, $r \cdot \bar{w} \in S^{\perp}$ too.

Combining the above, we see that $S^{\perp}$ is a subspace of V.

We now note that $S \cap S^{\perp} = \{\bar{0}_V\}$. Indeed, let $\bar{z} \in S \cap S^{\perp}$. Then $\bar{z} \in S$, and also $\bar{z} \in S^{\perp}$, so by definition of $S^{\perp}$ we must have $\langle \bar{z}, \bar{y} \rangle = 0$ for every $\bar{y} \in S$. In particular, we must have $\langle \bar{z}, \bar{z} \rangle = 0$, which implies that $\bar{z} = \bar{0}_V$ by the positive-definiteness of the inner product. Thus $S \cap S^{\perp} \subseteq \{\bar{0}_V\}$, and since both S and $S^{\perp}$ are subspaces, we conclude that $S \cap S^{\perp} = \{\bar{0}_V\}$.

We now note that, to show $V = S \oplus S^{\perp}$, it remains to show that $V = S + S^{\perp}$. Given that $\big(S, \langle \cdot, \cdot \rangle_S\big)$ is an inner product space (where $\langle \cdot, \cdot \rangle_S$ is the restriction of the inner product on V to pairs of vectors from S), we can find an orthonormal basis $\mathcal{B}_S = \{\bar{u}_i : i \in I\}$ of S *(note that one way to see this is that, if we have a basis $\mathcal{C}_S$ of S, then we can use the Gram-Schmidt orthogonalisation process to get from $\mathcal{C}_S$ a new basis of S that is also orthonormal)*.

But then, we can extend this basis of S to an orthonormal basis $\mathcal{B}_V = \{\bar{u}_i : i \in I\} \cup \{\bar{v}_j : j \in J\}$ of the entire space V *(again, one way to justify why this is possible is to first extend $\mathcal{B}_S$ to some basis of V, not necessarily orthonormal, and then to use the Gram-Schmidt orthogonalisation process again)*.

We now claim that $\{\bar{v}_j : j \in J\} \subset S^{\perp}$. Indeed, for every $\bar{y} \in S$, we can find $\bar{u}_{i_1}, \bar{u}_{i_2}, \ldots, \bar{u}_{i_k} \in \mathcal{B}_S = \{\bar{u}_i : i \in I\}$ for some $k \geqslant 1$, and $\mu_1, \mu_2, \ldots, \mu_k \in \mathbb{F}$, so that

$$\bar{y} = \mu_1 \bar{u}_{i_1} + \mu_2 \bar{u}_{i_2} + \cdots + \mu_k \bar{u}_{i_k}.$$

But then, for every $\bar{v}_{j_0} \in \{\bar{v}_j : j \in J\}$,

$$\begin{aligned}
\langle \bar{y}, \bar{v}_{j_0} \rangle &= \langle \mu_1 \bar{u}_{i_1} + \mu_2 \bar{u}_{i_2} + \cdots + \mu_k \bar{u}_{i_k}, \bar{v}_{j_0} \rangle \\
&= \mu_1 \langle \bar{u}_{i_1}, \bar{v}_{j_0} \rangle + \mu_2 \langle \bar{u}_{i_2}, \bar{v}_{j_0} \rangle + \cdots + \mu_k \langle \bar{u}_{i_k}, \bar{v}_{j_0} \rangle \quad \textit{(by linearity in the 1st argument)} \\
&= \mu_1 \cdot 0 + \mu_2 \cdot 0 + \cdots + \mu_k \cdot 0 = 0,
\end{aligned}$$

given that $\langle \bar{u}_i, \bar{v}_{j_0} \rangle = 0$ for every $\bar{u}_i \in \mathcal{B}_S$. Thus $\bar{v}_{j_0} \in S^{\perp}$, as we wanted.

Consider now an arbitrary $\bar{x} \in V$. Given that $\mathcal{B}_V = \{\bar{u}_i : i \in I\} \cup \{\bar{v}_j : j \in J\}$ is a basis of V, we can find $\bar{u}_{i_1}, \bar{u}_{i_2}, \ldots, \bar{u}_{i_k} \in \mathcal{B}_S = \{\bar{u}_i : i \in I\}$ and $\bar{v}_{j_1}, \bar{v}_{j_2}, \ldots, \bar{v}_{j_m} \in \{\bar{v}_j : j \in J\}$ for some $k, m \geqslant 1$, and also scalars $\lambda_1, \lambda_2, \ldots, \lambda_k$, $\mu_1, \mu_2, \ldots, \mu_m$, so that

$$\bar{x} \;=\; \lambda_1 \bar{u}_{i_1} + \lambda_2 \bar{u}_{i_2} + \cdots + \lambda_k \bar{u}_{i_k} \;+\; \mu_1 \bar{v}_{j_1} + \mu_2 \bar{v}_{j_2} + \cdots + \mu_m \bar{v}_{j_m}.$$

But then note that $\lambda_1 \bar{u}_{i_1} + \lambda_2 \bar{u}_{i_2} + \cdots + \lambda_k \bar{u}_{i_k} \in \mathrm{span}(\mathcal{B}_S) = S$, while $\mu_1 \bar{v}_{j_1} + \mu_2 \bar{v}_{j_2} + \cdots + \mu_m \bar{v}_{j_m} \in \mathrm{span}\big(\{\bar{v}_j : j \in J\}\big) \subseteq \mathrm{span}(S^{\perp}) = S^{\perp}$. In other words, this shows that the arbitrary $\bar{x} \in V$ belongs to $S + S^{\perp}$, proving that $V = S + S^{\perp}$.

Combining what we showed above, that $V = S + S^\perp$ and that $S \cap S^\perp = \{\bar{0}_V\}$, we conclude that $V = S \oplus S^\perp$.

(iv) In the same way as was justified in part (iii), we can extend the orthonormal basis $\{\bar{a}_1, \bar{a}_2, \ldots, \bar{a}_k\}$ of S to an orthonormal basis of V of the form $\{\bar{a}_1, \bar{a}_2, \ldots, \bar{a}_k\} \cup \{\bar{b}_j : j \in J\}$.

Consider now $\bar{x} \in V$. We can find $\bar{b}_{j_1}, \bar{b}_{j_2}, \ldots, \bar{b}_{j_m} \in \{\bar{b}_j : j \in J\}$ for some $m \geqslant 1$, and scalars $\lambda_1, \lambda_2, \ldots, \lambda_k$, $\mu_1, \mu_2, \ldots, \mu_m$, so that

$$\bar{x} \;=\; \lambda_1 \bar{a}_1 + \lambda_2 \bar{a}_2 + \cdots + \lambda_k \bar{a}_k \;+\; \mu_1 \bar{b}_{j_1} + \mu_2 \bar{b}_{j_2} + \cdots + \mu_m \bar{b}_{j_k}.$$

But then, by definition of the projection $P_{S;S^\perp}$,

$$P_{S;S^\perp}(\bar{x}) = \lambda_1 \bar{a}_1 + \lambda_2 \bar{a}_2 + \cdots + \lambda_k \bar{a}_k\,.$$

Moreover, exactly as we did in Problem 1 of this homework, we can check that, since $\{\bar{a}_1, \bar{a}_2, \ldots, \bar{a}_k\} \cup \{\bar{b}_j : j \in J\}$ is an orthonormal set,

$$\lambda_i = \langle \bar{x}, \bar{a}_i \rangle \qquad \text{for every } 1 \leqslant i \leqslant k.$$

We thus conclude that

$$P_{S;S^\perp}(\bar{x}) = \langle \bar{x}, \bar{a}_1 \rangle \bar{a}_1 + \langle \bar{x}, \bar{a}_2 \rangle \bar{a}_2 + \cdots + \langle \bar{x}, \bar{a}_k \rangle \bar{a}_k\,.$$

Problem 3. (I) In Homework 4, Problem 2, part (i), we found that the matrix

$$A_1 = \begin{pmatrix} 3 & 0 & 1 \\ -2.5 & -2 & -1.5 \\ 2 & 4 & -2 \end{pmatrix} \in \mathbb{C}^{3\times 3}$$

has eigenvalues $3, -2+2\boldsymbol{i}$ and $-2-2\boldsymbol{i}$. Given that these are 3 different eigenvalues (as many as the number of rows or columns of A_1), we can conclude that A_1 is diagonalisable (and hence also triangularisable). Indeed, since the geometric multiplicity of any eigenvalue is at least 1, we must have that the geometric multiplicities of all these eigenvalues add to 3, and hence by Theorem 1 (as stated in the file for Homework 6 before Problem 3) we will have that A_1 is diagonalisable.

In fact, we have also found that

- an eigenvector for the eigenvalue 3 is the vector $\left(\begin{smallmatrix} -2 \\ 1 \\ 0 \end{smallmatrix}\right)$;
- an eigenvector for the eigenvalue $-2+2\boldsymbol{i}$ is the vector $\begin{pmatrix} \frac{0.8-3\boldsymbol{i}}{2-0.8\boldsymbol{i}} - 3.75 \\ \frac{3.75+\boldsymbol{i}}{2-0.8\boldsymbol{i}} \\ 1 \end{pmatrix}$;
- an eigenvector for the eigenvalue $-2-2\boldsymbol{i}$ is the vector $\begin{pmatrix} \frac{0.8+3\boldsymbol{i}}{2+0.8\boldsymbol{i}} - 3.75 \\ \frac{3.75-\boldsymbol{i}}{2+0.8\boldsymbol{i}} \\ 1 \end{pmatrix}$.

Since these correspond to different eigenvalues, as is recalled before Problem 3 in the file for Homework 6, they will form a linearly independent set. Given also that this set will be of size 3, it will be a basis of $\mathbb{C}^3$. Thus, if we set

$$E_1 = \begin{pmatrix} -2 & \frac{0.8-3\boldsymbol{i}}{2-0.8\boldsymbol{i}} - 3.75 & \frac{0.8+3\boldsymbol{i}}{2+0.8\boldsymbol{i}} - 3.75 \\ 1 & \frac{3.75+\boldsymbol{i}}{2-0.8\boldsymbol{i}} & \frac{3.75-\boldsymbol{i}}{2+0.8\boldsymbol{i}} \\ 0 & 1 & 1 \end{pmatrix},$$

then, by Theorem 1 in Homework 6 again, we will have

$$E_1^{-1} A_1 E_1 = \begin{pmatrix} 3 & 0 & 0 \\ 0 & -2+2\boldsymbol{i} & 0 \\ 0 & 0 & -2-2\boldsymbol{i} \end{pmatrix}.$$

(II) In Homework 4, Problem 2, part (ii), we found that the matrix

$$A_2 = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 4 & 3 \\ 3 & 2 & 4 \end{pmatrix} \in \mathbb{Z}_5^{3\times 3}$$

has only one eigenvalue in $\mathbb{Z}_5$ with algebraic multiplicity $1 < 3$. Therefore, by Theorems 1 and $1'$ in Homework 6, we can conclude that this matrix is neither diagonalisable nor triangularisable over $\mathbb{Z}_5$.

(III) In Homework 4, Problem 2, part (iii), we found that the matrix

$$A_3 = \begin{pmatrix} 2 & 0 & 6 \\ 0 & 1 & 2 \\ 3 & 0 & 5 \end{pmatrix} \in \mathbb{Z}_7^{3\times 3}$$

has eigenvalues 1 and 6 in $\mathbb{Z}_7$, and that the eigenvalue 1 has algebraic multiplicity 2). We thus see that the algebraic multiplicities of the eigenvalues of A_3 add to 3 (the number of rows or columns of A_3), and thus by Theorem $1'$ in Homework 6 we can conclude that A_3 is (upper or lower) triangularisable.

On the other hand, we also found in the solution to Homework 4, Problem 2, part (iii) that the eigenspace corresponding to the eigenvalue 1 has dimension 1, or equivalently that the geometric multiplicity of 1 is strictly smaller than its algebraic multiplicity. But then, the geometric multiplicities of the eigenvalues of A_3 will not add to 3, and hence by Theorem 1 in Homework 6 we obtain that A_3 is **not** diagonalisable over $\mathbb{Z}_7$.

(IV) In Problem 1 of the 2nd Midterm Exam we found that the matrix

$$A = \begin{pmatrix} 6 & 0 & 3 \\ 2 & 2 & 5 \\ 1 & 0 & 1 \end{pmatrix} \in \mathbb{Z}_7^{3\times 3}$$

has eigenvalues 2 and 5, and that the eigenvalue 2 has both algebraic and geometric multiplicity equal to 2. But then, we see that the geometric multiplicities of the eigenvalues of A add to 3, and hence by Theorem 1 in Homework 6 we get that A is diagonalisable (and thus triangularisable too) over $\mathbb{Z}_7$.

In fact, we have also found that

- a basis for the eigenspace corresponding to the eigenvalue 2 of A is the set
$$\left\{ \left(\begin{smallmatrix} 0 \\ 1 \\ 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 \\ 0 \\ 1 \end{smallmatrix}\right) \right\};$$
- an eigenvector for the eigenvalue 5 is the vector $\left(\begin{smallmatrix} 4 \\ 2 \\ 1 \end{smallmatrix}\right)$.

Thus a basis of $\mathbb{Z}_7^3$ consisting of eigenvectors of A is the set

$$\left\{ \left(\begin{smallmatrix} 0 \\ 1 \\ 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 \\ 0 \\ 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 4 \\ 2 \\ 1 \end{smallmatrix}\right) \right\}.$$

It follows that, if we set

$$E = \begin{pmatrix} 0 & 1 & 4 \\ 1 & 0 & 2 \\ 0 & 1 & 1 \end{pmatrix},$$

then, by Theorem 1 in Homework 6, we will have

$$E^{-1}AE = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 5 \end{pmatrix}.$$

Problem 4. (i) Let $q_1, q_2 \in \mathcal{P}_4$. Then we can write $q_1(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4$ and $q_2(x) = b_0 + b_1x + b_2x^2 + b_3x^3 + b_4x^4$ for some coefficients $a_0, a_1, \ldots, a_4$, $b_0, b_1, \ldots, b_4$. But then

$$(q_1 + q_2)(x) = (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2 + (a_3 + b_3)x^3 + (a_4 + b_4)x^4,$$

$$\text{and thus } \; f(q_1 + q_2) = \begin{pmatrix} (a_0 + b_0) + (a_2 + b_2) & a_3 + b_3 \\ -a_1 - b_1 & (a_2 + b_2) + (a_4 + b_4) \end{pmatrix}$$

$$= \begin{pmatrix} a_0 + a_2 & a_3 \\ -a_1 & a_2 + a_4 \end{pmatrix} + \begin{pmatrix} b_0 + b_2 & b_3 \\ -b_1 & b_2 + b_4 \end{pmatrix} = f(q_1) + f(q_2).$$

Similarly, we see that, if $r \in \mathbb{R}$, then

$$(r \cdot q_1)(x) = r \cdot (a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4) = ra_0 + ra_1x + ra_2x^2 + ra_3x^3 + ra_4x^4,$$

$$\begin{aligned} \text{and thus } \; f(r \cdot q_1) &= \begin{pmatrix} ra_0 + ra_2 & ra_3 \\ -ra_1 & ra_2 + ra_4 \end{pmatrix} \\ &= r \cdot \begin{pmatrix} a_0 + a_2 & a_3 \\ -a_1 & a_2 + a_4 \end{pmatrix} = r \cdot f(q_1). \end{aligned}$$

We conclude that f is a linear function.

(ii) Given that $\dim_{\mathbb{R}} \mathcal{P}_4 = 5$ and $\dim_{\mathbb{R}} \mathbb{R}^{2\times 2} = 4$, we need to find a matrix $A \in \mathbb{R}^{4\times 5}$ so that, for every $q \in \mathcal{P}_4$,

$$A[q]_{\mathcal{B}} = \big[f(q)\big]_{\mathcal{C}}.$$

Given that, for every $1 \leqslant i \leqslant 5$, we will have $[p_i]_{\mathcal{B}} = \bar{e}_i \in \mathbb{R}^5 \equiv \mathbb{R}^{5\times 1}$, we can observe that

$$\mathrm{Col}_i(A) = A\bar{e}_i = A[p_i]_{\mathcal{B}} = \big[f(p_i)\big]_{\mathcal{C}}.$$

Thus we can determine the matrix A completely if we find $\big[f(p_i)\big]_{\mathcal{C}}$ for every $1 \leqslant i \leqslant 5$.

We have:

1. $f(p_1) = \begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix} = \frac{1}{2}\big(E_1 - E_2\big) + \frac{1}{2}\big(E_4 - E_3\big)$, therefore

$$\big[f(p_1)\big]_{\mathcal{C}} = \begin{pmatrix} 0.5 \\ -0.5 \\ -0.5 \\ 0.5 \end{pmatrix}.$$

2. $f(p_2) = \begin{pmatrix} -1 & 0 \\ -1 & 0 \end{pmatrix} = \frac{1}{2}\big(E_2 - E_1\big) + \frac{1}{2}\big(E_4 - E_3\big)$, therefore

$$\big[f(p_2)\big]_{\mathcal{C}} = \begin{pmatrix} -0.5 \\ 0.5 \\ -0.5 \\ 0.5 \end{pmatrix}.$$

3. $f(p_3) = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} = E_1 + \frac{1}{2}\big(E_4 - E_3\big)$, therefore

$$\big[f(p_3)\big]_{\mathcal{C}} = \begin{pmatrix} 1 \\ 0 \\ -0.5 \\ 0.5 \end{pmatrix}.$$

4. $f(p_4) = \begin{pmatrix} 0 & 2 \\ 0 & 1 \end{pmatrix} = \frac{1}{2}\big(E_1 + E_2\big) + E_3 + E_4$, therefore

$$\big[f(p_4)\big]_{\mathcal{C}} = \begin{pmatrix} 0.5 \\ 0.5 \\ 1 \\ 1 \end{pmatrix}.$$

5. Finally, $f(p_5) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \frac{1}{2}\big(E_3 + E_4\big)$, therefore

$$\big[f(p_5)\big]_{\mathcal{C}} = \begin{pmatrix} 0 \\ 0 \\ 0.5 \\ 0.5 \end{pmatrix}.$$

We conclude that the matrix representation of f that we want is

$$A = \begin{pmatrix} 0.5 & -0.5 & 1 & 0.5 & 0 \\ -0.5 & 0.5 & 0 & 0.5 & 0 \\ -0.5 & -0.5 & -0.5 & 1 & 0.5 \\ 0.5 & 0.5 & 0.5 & 1 & 0.5 \end{pmatrix}.$$

(iii) To determine $\mathrm{Ker}(f)$, we are looking for those polynomials $q(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 \in \mathcal{P}_4$ such that

$$\begin{pmatrix} a_0 + a_2 & a_3 \\ -a_1 & a_2 + a_4 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

This implies that $a_1 = a_3 = 0$, while $a_0 + a_2 = a_2 + a_4 = 0 \Leftrightarrow a_0 = -a_2 = a_4$.

It's also immediate to see that, if the coefficients of a polynomial in $\mathcal{P}_4$ satisfy the above, then the polynomial will be in $\mathrm{Ker}(f)$.

In other words,

$$\mathrm{Ker}(f) = \{q \in \mathcal{P}_4 : \ q(x) = r - rx^2 + rx^4 \text{ for some } r \in \mathbb{R}\} = \mathrm{span}\big(\{1 - x^2 + x^4\}\big).$$

We can similarly verify that $\mathrm{Range}(f) = \mathbb{R}^{2\times 2}$ (or in other words, that f is surjective).

Alternatively, we could use Main Theorem E here: we must have that

$$5 = \dim_{\mathbb{R}} \mathcal{P}_4 = \dim_{\mathbb{R}} \mathrm{Ker}(f) + \dim_{\mathbb{R}} \mathrm{Range}(f) = 1 + \dim_{\mathbb{R}} \mathrm{Range}(f),$$

which shows that $\mathrm{Range}(f)$ is a subspace of $\mathbb{R}^{2\times 2}$ of dimension 4, and thus it must coincide with $\mathbb{R}^{2\times 2}$.

Problem 5. As we did in Homework 5 too, given $\bar{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ and $\bar{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n$, let us write $\bar{x} \bullet \bar{y} := \sum_{s=1}^{n} x_s y_s$. Note that in $\mathbb{R}^n$ this coincides with the standard inner product of $\bar{x}$ and $\bar{y}$, however we will use this alternative notation too because it allows us to describe matrix multiplication more easily (and this remains true even in settings that do not relate to inner product spaces, so it's convenient to distinguish the two uses).

We recall that the (i,j)-th entry of $C^T \cdot C$ is given by

$$\mathrm{Row}_i(C^T) \bullet \mathrm{Col}_j(C) = \mathrm{Col}_i(C) \bullet \mathrm{Col}_j(C) = \bar{u}_i \bullet \bar{u}_j = \langle \bar{u}_i, \bar{u}_j \rangle.$$

Thus, recalling also that $\{\bar{u}_1, \bar{u}_2, \ldots, \bar{u}_n\}$ is an orthonormal set, we get that the (i,j)-th entry of $C^T \cdot C$ is equal to 0 if $i \neq j$, while it is equal to $1 = \langle \bar{u}_i, \bar{u}_i \rangle = \|\bar{u}_i\|^2$ if $i = j$. In other words, $C^T \cdot C = I_n$.

But this shows that C has a left inverse. We now recall that this implies that C also has a right inverse, and moreover that we can then conclude that the left and right inverses of C coincide. In other words, C is invertible and $C^{-1} = C^T$. Thus we also get that $C \cdot C^T = I_n$.

We now use analogous reasoning to above:

- if $i \neq j$, the (i,j)-th entry of $C \cdot C^T = I_n$ is equal to 0, which gives

  $$0 = \mathrm{Row}_i(C) \bullet \mathrm{Col}_j(C^T) = \mathrm{Row}_i(C) \bullet \mathrm{Row}_j(C) = \langle \mathrm{Row}_i(C), \mathrm{Row}_j(C) \rangle\,;$$

- similarly, for every $1 \leqslant i \leqslant n$, the (i,i)-th entry of $C \cdot C^T = I_n$ is equal to 1, which gives

$$1 = \mathrm{Row}_i(C) \bullet \mathrm{Col}_i(C^T) = \mathrm{Row}_i(C) \bullet \mathrm{Row}_i(C) = \langle \mathrm{Row}_i(C), \mathrm{Row}_j(C) \rangle = \|\mathrm{Row}_i(C)\|^2.$$

Combining these observations, we can conclude that the set

$$\{\mathrm{Row}_1(C), \mathrm{Row}_2(C), \ldots, \mathrm{Row}_n(C)\}$$

is an orthonormal subset of $\mathbb{R}^n$.

Finally, we observe that

$$1 = \det(I_n) = \det(C^T \cdot C) = \det(C^T)\det(C) = \big(\det(C)\big)^2$$

(given that $\det(C^T) = \det(C)$), and thus the real number $\det(C)$ is equal either to 1 or to -1.