Solution 1.1.1.

1. To begin, inspect the case when $n = 2$.
When $n = 2$, it is equivalent to finding a linear transformation that, when it is a composite transformation of its own, that composite transformation maps any vector on the plane to another vector in its opposite direction. That can be done by rotation matrix of $\pm\frac{\pi}{2}$, or any matrix similar to that rotation matrix. Then we have

$$\text{Rot}\left(\frac{\pi}{2}\right) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \text{Rot}\left(\frac{\pi}{2}\right)^2 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

Let $A_1 = \text{Rot}\left(\frac{\pi}{2}\right)$. The following proves that when $n = 2k$, the matrix would be in the form that the main diagonal of the matrix is filled with blocks of A_1, and elsewhere zero, or any matrix similar to that.
Let $S(k) : (A_k)^2 = -I_{2k\times 2k}$.
The proof of $S(1)$ is shown above.
Assume $S(k)$ is true, then for $S(k+1)$,

$$A_{k+1} = \begin{bmatrix} A_k & \\ & A_1 \end{bmatrix}, A_{k+1}^2 = \begin{bmatrix} A_k & \\ & A_1 \end{bmatrix}\begin{bmatrix} A_k & \\ & A_1 \end{bmatrix} = \begin{bmatrix} -I_{2k} & \\ & -I_2 \end{bmatrix} = -I_{2(k+1)\times 2(k+1)}$$

So we have $S(k) \Rightarrow S(k+1)$. By the first principle of mathematical induction, $S(k)$ is true for any positive integer k, and hence

$$A = \begin{bmatrix} A_1 & & \\ & \ddots & \\ & & A_1 \end{bmatrix}$$

or any other matrices similar to that are possible solutions.

2. First, consider a Jordan block. Every Jordan block would be in the form of

$$J = \begin{bmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{bmatrix}, J^2 = \begin{bmatrix} \lambda^2 & 2\lambda & 1 & & & \\ & \lambda^2 & 2\lambda & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & \lambda^2 & 2\lambda & 1 \\ & & & & \lambda^2 & 2\lambda \\ & & & & & \lambda^2 \end{bmatrix}$$

Assume that $A \in \mathbf{R}^{n\times n}$, then its characteristic polynomial would be a polynomial with real coefficients. By Vieta's theorem, if a complex number z is an eigenvalue of A, then $\bar{z}$ is another eigenvalue of A, and hence when n is odd, there exists real eigenvalue to A.
Factorize A in the form that $A = XJX^{-1}$, where X are generalized eigenvectors w.r.t. Jordan blocks in J, then we have

$$A^2 = XJ^2X^{-1} = -I \Leftrightarrow J^2 = -I,$$

so A should be diagonalizable, and nothing other than $\pm i$ can be eigenvalues of A, contradicting the fact that there exists real eigenvalue to A. ■

Solution 1.1.2.

1. A is known as a linear complex structure.
Let $k = x + iy$, where $x, y \in \mathbf{R}, k \in \mathbf{C}$.
Let $P = B(kv) - kB(v)$, then

$$P = B(kv) - kB(v) = B(x+iy)v - (x+iy)Bv = B(xI + Ay)v - (xI + Ay)Bv$$

Since $A, B \in M_n(\mathbf{R})$, A and B are both $\mathbf{R}$–linear, then

$$P = xBv + yBAv - xBv - yABv = yBAv - yABv$$

As v and y are both arbitary,

$$B(kv) = k(Bv) \Leftrightarrow P = 0 \Leftrightarrow AB = BA$$

2. No. Consider the case when
$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, X = \begin{bmatrix} 1 & -2 \\ 1 & -1 \end{bmatrix}$$
Then
$$AX - XA = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & -2 \\ 1 & -1 \end{bmatrix} - \begin{bmatrix} 1 & -2 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 3 & -2 \\ -2 & 1 \end{bmatrix} \neq 0$$

3. Set A=$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, then $C = \mathrm{Ref}(\theta) = \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix}$, $\theta \in [0, \pi)$.
Clearly, $\mathrm{Ref}(\theta)$ represents a reflection matrix, which implies $C^2 = I$.
$$\begin{aligned} CA + AC &= \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix} \\ &= \begin{bmatrix} \sin 2\theta & -\cos 2\theta \\ -\cos 2\theta & -\sin 2\theta \end{bmatrix} + \begin{bmatrix} -\sin 2\theta & \cos 2\theta \\ \cos 2\theta & \sin 2\theta \end{bmatrix} = 0 \end{aligned}$$
For example,
$$C_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, C_2 = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}.$$

Solution 1.1.3.

1. C is real linear, but not complex linear. $\forall a \in \mathbf{R}$, $\forall v \in \mathbf{C}^n$,
$$C(a\mathbf{v}) = C\begin{bmatrix} ax_1 + iay_1 \\ ax_2 + iay_2 \\ \vdots \\ ax_n + iay_n \end{bmatrix} = \begin{bmatrix} ax_1 - iay_1 \\ ax_2 - iay_2 \\ \vdots \\ ax_n - iay_n \end{bmatrix}$$
$$a(C\mathbf{v}) = a\left(C\begin{bmatrix} x_1 + iy_1 \\ x_2 + iy_2 \\ \vdots \\ x_n + iy_n \end{bmatrix}\right) = a\begin{bmatrix} x_1 - iy_1 \\ x_2 - iy_2 \\ \vdots \\ x_n - iy_n \end{bmatrix} = \begin{bmatrix} ax_1 - iay_1 \\ ax_2 - iay_2 \\ \vdots \\ ax_n - iay_n \end{bmatrix}$$
However, C is not complex linear. For instance,
$$C\left(i \cdot \begin{bmatrix} 1+i \\ 2 \\ 3+i \end{bmatrix}\right) = C\begin{bmatrix} -1+i \\ 2i \\ -1+3i \end{bmatrix} = \begin{bmatrix} -1-i \\ -2i \\ -1-3i \end{bmatrix},$$
but then
$$i \cdot C\begin{bmatrix} 1+i \\ 2 \\ 3+i \end{bmatrix} = i \cdot \begin{bmatrix} 1-i \\ 2 \\ 3-i \end{bmatrix} = \begin{bmatrix} 1+i \\ 2i \\ 1+3i \end{bmatrix}$$

2. **C**-linear $\Rightarrow$ **R**-linear. Sufficiency is proved by the fact that reals are complex number with zero imaginary component. Necessity is disproved in (1).

3. **R**-basis: $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} i \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ i \end{bmatrix} \right\}$, **C**-basis: $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$.
Real-dimension of $\mathbf{C}^2$=size of **R**-basis=4, complex-dimension of $\mathbf{C}^2$=size of **C**-basis=2.

4. **C**-linearly independent $\Rightarrow$ **R**-linearly independent. Consider matrix B with basis column vectors. Let $\mathbf{v} = \mathbf{x} + i\mathbf{y}$, where $\mathbf{x}, \mathbf{y} \in \mathbf{R}^n$. B is filled by basis, so $B\mathbf{v} = B\mathbf{x} + Bi\mathbf{y} = 0$ has $\mathbf{x} = \mathbf{y} = 0$, proving sufficiency. Necessity is disproved by considering (3) that
$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} + i \cdot \begin{bmatrix} i \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} + i \cdot \begin{bmatrix} 0 \\ i \end{bmatrix} = 0$$

5. **R**-spanning $\Rightarrow$ **C**-spanning. Any vector in a space w.r.t. the set of spanning vectors can be expressed as a linear combination of the spanning vectors. **R**-spanning means the coefficients corresponding to the linear combination are real, which is a special case of expressing the vectors in a linear combination with coefficients of complex numbers, in the sense that the imaginary part is zero. Necessity is disproved by (3) that $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$ failed to span $\mathbf{C}^2$ as non real vectors in $\mathbf{C}^2$ are missed out.

Solution 1.1.4.

By given conditions,

$$P = \begin{bmatrix} & 1 & & \\ & & 1 & \\ & & & 1 \\ 1 & & & \end{bmatrix}$$

1.

$$P\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, P\begin{bmatrix} 1 \\ i \\ -1 \\ -i \end{bmatrix} = \begin{bmatrix} i \\ -1 \\ -i \\ 1 \end{bmatrix}$$

2.

$$F_4 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{bmatrix}, F_4^{-1}PF_4 = \frac{1}{4}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \\ 1 & i & -1 & -i \end{bmatrix}\begin{bmatrix} 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -i \end{bmatrix}$$

D is similar to P, so $\lambda_k = i^k, (k = 0, 1, 2, 3)$ are eigenvalues of P, corresponding eigenvectors are

$$\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ i \\ -1 \\ -i \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -i \\ -1 \\ i \end{bmatrix}.$$

3.

$$C = \begin{bmatrix} c_0 & c_1 & c_2 & c_3 \\ c_3 & c_0 & c_1 & c_2 \\ c_2 & c_3 & c_0 & c_1 \\ c_1 & c_2 & c_3 & c_0 \end{bmatrix}$$

$$C\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} c_0+c_1+c_2+c_3 \\ c_0+c_1+c_2+c_3 \\ c_0+c_1+c_2+c_3 \\ c_0+c_1+c_2+c_3 \end{bmatrix}, C\begin{bmatrix} 1 \\ i \\ -1 \\ -i \end{bmatrix} = \begin{bmatrix} (c_0-c_2)+(c_1-c_3)i \\ (c_3-c_1)+(c_0-c_2)i \\ (c_2-c_0)+(c_3-c_1)i \\ (c_1-c_3)+(c_2-c_0)i \end{bmatrix}$$

4.

$$C = c_0 I + c_1 P + c_2 P^2 + c_3 P^3$$

As C is similar to $F_4^{-1}CF_4$, C and $F_4^{-1}CF_4$ shares the same eigenvalues.

$$F_4^{-1}CF_4 = c_0 I + c_1 F_4^{-1}PF_4 + c_2 F_4^{-1}P^2F_4 + c_3 F_4^{-1}P^3F_4$$

$$= c_0 I + c_1 D + c_2 D^2 + c_3 D^3$$

$$= \begin{bmatrix} c_0+c_1+c_2+c_3 & 0 & 0 & 0 \\ 0 & (c_0-c_2)+(c_1-c_3)i & 0 & 0 \\ 0 & 0 & c_0-c_1+c_2-c_3 & 0 \\ 0 & 0 & 0 & (c_0-c_2)+(c_3-c_1)i \end{bmatrix}$$

Thus $\lambda_0 = c_0 + c_1 + c_2 + c_3$,
$\lambda_1 = (c_0 - c_2) + (c_1 - c_3)i$,
$\lambda_2 = c_0 - c_1 + c_2 - c_3$ and
$\lambda_3 = (c_0 - c_2) + (c_3 - c_1)i$ are the eigenvalues associated with C, the four associated eigenvectors are

$$\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ i \\ -1 \\ -i \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -i \\ -1 \\ i \end{bmatrix}.$$

In the following, "linearly independent" is used interchangeably with "independent", and "linearly dependent" is used interchangeably with "dependent".

Solution 1.2.1.

1. True.
   Suppose that any three of the four given subspaces are linearly independent, but the four subspaces are not. Then assume
   $$v_i \in V_i, i = 1, 2, 3, 4$$
   Since the four subspaces are dependent, W.L.O.G. assume $\dim(V_4 \cap (V_1 + V_2 + V_3)) > 0$, then
   $$v_4 = -(v_1 + v_2 + v_3)$$
   So $||v_1|| + ||v_2|| + ||v_3|| > 0$, that means V_4 has intersection with at least one subspace, contradicting the assumption that V_4 and any other two of the three spaces are linearly independent. ■

2. False.
   Consider
   $$V_1 = \operatorname{span}\left\{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right\}, V_2 = \operatorname{span}\left\{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right\}, V_3 = \operatorname{span}\left\{\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right\}, V_4 = \operatorname{span}\left\{\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}\right\}$$
   Here $V_4 = V_1 + V_2$, so V_1, V_2, V_4 are linearly dependent.

3. True.
   Assume
   $$v_i \in V_i, i = 1, 2, 3, 4.$$
   If the four subspaces are dependent, then
   $$v_1 + v_2 + v_3 + v_4 = 0$$
   has non-trivial solutions. Say
   $$v_4 = -(v_1 + v_2 + v_3)$$
   Hence $||v_1|| + ||v_2|| + ||v_3|| > 0$. So at least one of the three vectors is non-zero. If v_3 is non-zero, then it contradicts that "V_3, V_4 are independent". If v_1 or v_2 is non-zero, then

   (a) If $v_1 + v_2 = 0$, then $v_1 = -v_2$, and both of which are non-zero. It violates that "V_1, V_2 are independent".

   (b) Else, since $v_4 \in (V_3 + V_4)$, and $v_1 + v_2 \neq 0$, then $(V_1 + V_2) \cap (V_3 + V_4)$ is a non-trivial space, which is another contradiction.

   Switch v_4 to the vectors from the three other subspaces, and the same conclusion would be drawn.
   ■

Solution 1.2.2.

By definition,

$$T : A \mapsto A^T$$

and

$$T^2(A) = (A^T)^T = A.$$

By observation,

$$A \text{ is symmetric } \Leftrightarrow T(A) = A$$

$$A \text{ is skew-symmetric } \Leftrightarrow T(A) = -A.$$

T is bijective, so if a matrix A_0 is both symmetric and anti-symmetric, then

$$T(A_0) + T(A_0) = A_0 + (-A_0) = 0 \Rightarrow A_0 = 0$$

Thus, the space of symmetric and skew-symmetric matrices are independent.
Denote the former as Sym, and the latter as Skew. Consider

$$A = A_1 + A_2,$$

where

$$A_1 = \frac{1}{2}(A + T(A)), A_2 = \frac{1}{2}(A - T(A))$$

$$T(A_1) = \frac{1}{2}T(A + T(A)) = \frac{1}{2}(T(A) + T^2(A)) = \frac{1}{2}(T(A) + A) \in \text{Sym}$$

$$T(A_2) = \frac{1}{2}T(A - T(A)) = \frac{1}{2}(T(A) - T^2(A)) = \frac{1}{2}(T(A) - A) \in \text{Skew}$$

So every matrix can be expressed as a linear combination of two matrices, one from Sym, and one from Skew. Hence

$$V = \text{Sym} \bigoplus \text{Skew}$$

In particular, the former has dimension $\frac{1}{2}n(n+1)$, and the latter has $\frac{1}{2}n(n-1)$.

With regards to block form of T,

$$T = \begin{bmatrix} e_1 & \\ & -e_2 \end{bmatrix},$$

where e_1 is the block representing the transformation:

$$T_1 : \text{Sym} \to \text{Sym}, T_1 \text{ is identity transformation}$$

e_2 is the block representing the transformation:

$$T_2 : \text{Skew} \to \text{Skew}, T_2 \text{ is identity transformation}$$

Solution 1.2.3.

1. Ker(B):
   $\forall v \in \text{Ker}(B)$,

   $$B(Av) = BAv = ABv = A(Bv) = A \cdot 0 = 0,$$

   so $v \in \text{Ker}(B) \Rightarrow Av \in \text{Ker}(B)$ $\square$
   Ran(B) :
   $\forall w \in \text{Ran}(B)$, $\exists v$ s.t. $Bv = w$.

   $$A(w) = A(Bv) = ABv = BAv = B(Av),$$

   so $w \in \text{Ran}(B) \Rightarrow Aw \in \text{Ran}(B)$.

2. $$Ap(A) = A \cdot \sum_{i=0}^{n-1} p_i A^i = \sum_{i=0}^{n-1} p_i A^{i+1} = (\sum_{i=0}^{n-1} p_i A^i)A = p(A)A$$

3. Let k be the smallest positive integer satisfies $\mathrm{Dim}(\mathrm{Ker}((A-\lambda I)^k)) = n$, where $k \le n$. Then $\mathrm{Ker}((A-\lambda I)^k) = N_\infty(A-\lambda I)$, $\mathrm{Ran}((A-\lambda I)^k) = R_\infty(A-\lambda I)$. As

$$(A-\lambda I)^k = \sum_{i=0}^{k} \begin{pmatrix} k \\ i \end{pmatrix} A^i(-\lambda)^{k-i}$$

is a polynomial of A, by (2) $A(A-\lambda I)^k = (A-\lambda I)^k A$, and by (1), and by the fact that $\lambda \in \mathbf{C}$ has no restriction, the conclusion is drawn.

Solution 1.2.4.

1. Suppose W is an A-invariant subspace, but none of the eigenvectors of A belongs to W. Then $\forall \lambda \in \mathbf{C}, \forall v \in W$, $v \neq 0$

$$Av \neq \lambda v \Leftrightarrow (A-\lambda I)v \neq 0$$
$$\Leftrightarrow |A|_W - \lambda I| \neq 0$$

Hence

$$p_{A|_W}(\lambda)$$

has no roots. By the fundamental theorem of algebra, $\mathrm{deg} p_{A|_W} = 0$. Since $p_A(\lambda) = p_{A|_w}(\lambda) \cdot p_{A|_{W^\perp}}(\lambda)$,
$\dim(W) = \mathrm{deg} p_{A|_W} = 0 \Rightarrow W = \{0\}$, which contradicts the assumption that $v \neq 0$. □

2. $$B(A-\lambda I) = BA - \lambda B = AB - \lambda B = (A-\lambda I)B,$$

so by 1.2.3.1. $\forall \lambda \in \mathbf{C} : \mathrm{Ker}(A-\lambda I)$ is B-invariant. □

3. In (1), let $W = \mathrm{Ker}(A-\lambda_1 I)$, and fix λ_1 as an eigenvalue of A, then B has an eigenvector in $\mathrm{Ker}(A-\lambda_1 I)$, which is the λ_1-eigenspace of A, hence that eigenvector of B is also an eigenvector of A. ■

Solution 1.3.1.

1. $$A : \begin{bmatrix} \Re(x) \\ \Im(x) \\ \Re(y) \\ \Im(y) \end{bmatrix} \mapsto \begin{bmatrix} \Re(x) - \Re(y) \\ -\Im(x) \\ \Im(x) - \Re(y) \\ \Im(x) - \Im(y) \end{bmatrix},$$

So

$$A = \begin{bmatrix} 1 & & -1 & \\ & -1 & & \\ & 1 & -1 & \\ & 1 & & -1 \end{bmatrix}, A - \lambda I = \begin{bmatrix} 1-\lambda & & -1 & \\ & -1-\lambda & & \\ & 1 & -1-\lambda & \\ & 1 & & -1-\lambda \end{bmatrix},$$

Then $\lambda = 1$ (multiplicity 1), $\lambda = -1$ (multiplicity 3).

W.R.T. $\lambda = 1$, eigenvector $v = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$.

W.R.T. $\lambda = -1$,

$$A + I = \begin{bmatrix} 2 & & -1 & \\ & 0 & & \\ & 1 & 0 & \\ & 1 & & 0 \end{bmatrix}, (A+I)^2 = \begin{bmatrix} 4 & -1 & -2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

By observation, $\mathrm{Dim}(\mathrm{Ker}(A+I)) = 2, \mathrm{Dim}(\mathrm{Ker}(A+I)^2) = 3$, so

$$\begin{bmatrix} 0 \\ 2 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix},$$

are the basis of the generalized eigenspace associated with $\lambda = -1$. If we let

$$B = \begin{bmatrix} 1 & & & 1 \\ & 2 & & \\ 2 & -1 & & \\ 2 & & 1 & \end{bmatrix}, J = \begin{bmatrix} -1 & 1 & & \\ & -1 & & \\ & & -1 & \\ & & & 1 \end{bmatrix}$$

Then,

$$A = \begin{bmatrix} 1 & & & 1 \\ & 2 & & \\ 2 & -1 & & \\ 2 & & 1 & \end{bmatrix} \begin{bmatrix} -1 & 1 & & \\ & -1 & & \\ & & -1 & \\ & & & 1 \end{bmatrix} \cdot \frac{1}{4} \begin{bmatrix} & 1 & 1 & \\ & 2 & & \\ & -2 & -4 & 4 \\ 4 & -1 & -2 & \end{bmatrix}$$

The column vectors of the leftmost matrix is the basis in which the mapping is in its JCF.

2.

$$A : \begin{bmatrix} p_0 \\ p_1 \\ p_2 \\ p_3 \\ p_4 \end{bmatrix} \mapsto \begin{bmatrix} p_0 + p_1 \\ 2p_2 \\ p_1 + 3p_3 \\ 4p_4 \\ 0 \end{bmatrix},$$

So

$$A = \begin{bmatrix} 1 & 1 & & & \\ & & 2 & & \\ & 1 & & 3 & \\ & & & & 4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, A - \lambda I = \begin{bmatrix} 1-\lambda & 1 & & & \\ & -\lambda & 2 & & \\ & 1 & -\lambda & 3 & \\ & & & -\lambda & 4 \\ & & & & -\lambda \end{bmatrix},$$

$$\det(A - \lambda I) = (1-\lambda)(\lambda^2)(\lambda^2 - 2),$$

so $\lambda = 1$, $\lambda = \pm\sqrt{2}$, $\lambda = 0$(multiplicity 2).

W.R.T. $\lambda = 1$, eigenvector $v = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$.

W.R.T. $\lambda = \sqrt{2}$, eigenvector $v = \begin{bmatrix} 2+\sqrt{2} \\ \sqrt{2} \\ 1 \\ 0 \\ 0 \end{bmatrix}$.

W.R.T. $\lambda = -\sqrt{2}$, eigenvector $v = \begin{bmatrix} 2+\sqrt{2} \\ -\sqrt{2} \\ 1 \\ 0 \\ 0 \end{bmatrix}$.

W.R.T. $\lambda = 0$, observe that

$$A^2 = \begin{bmatrix} 1 & 1 & 2 & & \\ & 2 & & 6 & \\ & & 2 & & 12 \\ & & & 0 & \\ & & & & 0 \end{bmatrix}$$

$$\begin{bmatrix} 12 \\ 0 \\ -6 \\ 0 \\ 1 \end{bmatrix} \in \mathrm{Ker}(A^2), \text{ but } A \begin{bmatrix} 12 \\ 0 \\ -6 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 12 \\ -12 \\ 0 \\ 4 \\ 0 \end{bmatrix} \notin \mathrm{Ker}(A),$$

so

$$\left\{\begin{bmatrix}12\\-12\\0\\4\\0\end{bmatrix},\begin{bmatrix}12\\0\\-6\\0\\1\end{bmatrix},\begin{bmatrix}2+\sqrt{2}\\\sqrt{2}\\1\\0\\0\end{bmatrix},\begin{bmatrix}1\\0\\0\\0\\0\end{bmatrix},\begin{bmatrix}2+\sqrt{2}\\-\sqrt{2}\\1\\0\\0\end{bmatrix}\right\}$$

is the basis in which the transformation would be expressed in JCF, in which

$$J=\begin{bmatrix}0&1&&&\\&0&&&\\&&\sqrt{2}&&\\&&&1&\\&&&&-\sqrt{2}\end{bmatrix}$$

3. Since

$$\begin{bmatrix}&&&a_1\\&&a_2&\\&a_3&&\\a_4&&&\end{bmatrix}\begin{bmatrix}1&&&\\&&1&\\&&&1\\&1&&\end{bmatrix}=\begin{bmatrix}&a_1&&\\&&&a_2\\&&a_3&\\a_4&&&\end{bmatrix},$$

$$\begin{bmatrix}1&&&\\&&1&\\&&&1\\&1&&\end{bmatrix}\begin{bmatrix}&a_1&&\\a_4&&&\\&&&a_2\\&&a_3&\end{bmatrix}=\begin{bmatrix}&a_1&&\\&&&a_2\\&&a_3&\\a_4&&&\end{bmatrix},$$

$\begin{bmatrix}&a_1&&\\a_4&&&\\&&&a_2\\&&a_3&\end{bmatrix}$ and $\begin{bmatrix}&&&a_1\\&&a_2&\\&a_3&&\\a_4&&&\end{bmatrix}$ shares the same JCF.

Further decompose $\begin{bmatrix}&a_1&&\\a_4&&&\\&&&a_2\\&&a_3&\end{bmatrix}$ into two parts: $\begin{bmatrix}&a_1\\a_4&\end{bmatrix}$ and $\begin{bmatrix}&a_2\\a_3&\end{bmatrix}$.

Then the problem boils down to finding the Jordan decomposition of the two blocks, in which both in the form of $\begin{bmatrix}&a\\b&\end{bmatrix}$.

(a) $a=b=0$, then the basis in this block relative to the matrix we had been decomposed are standard basis: $\begin{bmatrix}1\\0\end{bmatrix}$ and $\begin{bmatrix}0\\1\end{bmatrix}$, and the Jordan block is zero matrix.

(b) $a=0, b\neq 0$, then it is $\begin{bmatrix}0\\1\end{bmatrix}$ and $\begin{bmatrix}1\\0\end{bmatrix}$, corresponds to the Jordan block of $\begin{bmatrix}0&1\\0&0\end{bmatrix}$

(c) $a\neq 0, b=0$, then it is $\begin{bmatrix}1\\0\end{bmatrix}$ and $\begin{bmatrix}0\\1\end{bmatrix}$, corresponds to the Jordan block of $\begin{bmatrix}0&1\\0&0\end{bmatrix}$.

(d) Otherwise, the block is diagonalizable, and $\begin{bmatrix}\sqrt{a}\\\sqrt{b}\end{bmatrix}$ is the eigenvector of $\lambda=-\sqrt{ab}$, $\begin{bmatrix}\sqrt{a}\\-\sqrt{b}\end{bmatrix}$ is the eigenvector of $\lambda=\sqrt{ab}$

Use the basis generated above to form a matrix that preserves the order of decomposing the matrix, and use $\begin{bmatrix}1&&&\\&&1&\\&&&1\\&1&&\end{bmatrix}$ to multiply that matrix, then the column vectors of the product would be the basis, and putting the Jordan blocks correspond to each block back in order will be the Jordan block of that anti-diagonal matrix.

Solution 1.3.2.

1. The number of dots in the k-th row, is $\mathrm{Dim}(\mathrm{Ker}(A^k) - \mathrm{Ker}(A^{k-1}) + \{0\})$, which is the dimension gap between the kernels of adjacent matrix raised to power k and k-1, also representing number of "basis vectors" that span the gap between the kernels of the matrix powers, ignoring zero element.

   Meanwhile, the number of dots in the k-th column, is the length of the Jordan chain corresponding to the k-th generalized eigenvector. Moreover, the number of columns is the dimension of the kernel of A.

2. It is given that the summands are non-increasing.
   So every self-conjugating partition has one-to-one correspondence to layers of L, in which number of dots in each L are strictly decreasing from top to bottom.
   This claim can be proved by contradiction that if it is not decreasing, then one "arm" of the L shaped dots, which is the row or the column of it would be longer than its preceding row or column, which is a contradiction to the assumption.
   The above prove distinctness.

   Below proves bijection exists.
   We treat summands of the distinct odd partitions as the number of dots in each "L".
   Every "L" has $2(k-1)+1 = 2k-1$ dots, where k is the number of dots in the row (or the column), which indicates that "L" has odd number of dots.
   That also implies for any partition,
   the distinct odd partition represented by the number of dots in each "L" for two self-conjugating partitions would be equal iff the two self-conjugating partitions are itself equal.

3. Suppose $A = \begin{bmatrix} 0 & a_{12} & a_{13} & a_{14} \\ & 0 & a_{23} & a_{24} \\ & & 0 & a_{34} \\ & & & 0 \end{bmatrix}$.
   Then $A^2 = \begin{bmatrix} 0 & 0 & a_{12}a_{23} & a_{12}a_{24} + a_{13}a_{34} \\ & 0 & 0 & a_{23}a_{34} \\ & & 0 & 0 \\ & & & 0 \end{bmatrix}$, $A^3 = \begin{bmatrix} 0 & 0 & 0 & a_{12}a_{23}a_{24} \\ & 0 & 0 & 0 \\ & & 0 & 0 \\ & & & 0 \end{bmatrix}$.
   Let k be the smallest integer satisfies $\mathrm{Ker}(A^k) = N_\infty(A)$
   Then $k = 4$ iff none of the elements in the upper diagonal is zero.
   $k = 3$ iff some of the elements in upper diagonal is zero. That requires a_{12} and a_{34} cannot both be zero.
   $k = 2$ iff $a_{12} = a_{34} = 0$, or $a_{13} = a_{23} = a_{24} = 0$.
   $k = 1$ iff all elements are zero.
   As $X \sim \mathcal{U}(-1, 1)$ is a continuous distribution if X is an entry of the upper triangular matrix, $P(X = x) = 0$, so $P(k = 4) = 1$, $P(k = 3) = P(k = 2) = P(k = 1) = 0$, means that it is almost certain that it would be in the case of $4 = 4$, and nothing else. So for case $4 = 4$, $P = 1$, and $P = 0$ otherwise.

Solution 1.4.1.

1. $B = \begin{bmatrix} I & X \\ & I \end{bmatrix}$, where X satisfies

$$\begin{bmatrix} 1 & 2 \\ & 1 \end{bmatrix} X - X \begin{bmatrix} 3 & 5 \\ & 4 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

   Let $X = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, then

$$\begin{cases} -2a + 2c & = & 1 \\ -5a - 3b + 2d & = & 2 \\ -2c & = & 3 \\ -5c - 3d & = & 4 \end{cases}$$

So $B = \begin{bmatrix} 1 & 0 & -2 & 31/9 \\ 0 & 1 & -3/2 & 7/6 \\ & & 1 & \\ & & & 1 \end{bmatrix}$

2.

$$\mathcal{B} = \left\{ \begin{bmatrix} -4 \\ -3 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 62 \\ 21 \\ 0 \\ 18 \end{bmatrix} \right\}$$

is a basis that span $V_3 + V_4$.

Solution 1.4.2.

1.

$$\begin{bmatrix} I & & & \\ & & I & \\ & I & & \\ & & & I \end{bmatrix} \begin{bmatrix} X & & I & \\ & Y & & I \\ & & X & \\ & & & Y \end{bmatrix} = \begin{bmatrix} X & & I & \\ & & X & \\ & Y & & I \\ & & & Y \end{bmatrix}$$

$$\begin{bmatrix} X & I & & \\ & X & & \\ & & Y & I \\ & & & Y \end{bmatrix} \begin{bmatrix} I & & & \\ & & I & \\ & I & & \\ & & & I \end{bmatrix} = \begin{bmatrix} X & & I & \\ & & X & \\ & Y & & I \\ & & & Y \end{bmatrix}$$

Hence $\begin{bmatrix} X & & I & \\ & Y & & I \\ & & X & \\ & & & Y \end{bmatrix}$ and $\begin{bmatrix} X & I & & \\ & X & & \\ & & Y & I \\ & & & Y \end{bmatrix}$ are similar.

2.

$$p_B(x) = (3-x)(4-x) = 12 - 7x + x^2,$$
$$p_A(x) = (3-x)^2(4-x)^2 = 144 - 168x + 73x^2 - 14x^3 + x^4$$

3.

$$A = \begin{bmatrix} & 1 & 1 & \\ & & & 1 \\ & & & 1 \\ & & & 0 \end{bmatrix}$$

A has two pivot columns, $\mathrm{rank}(A) = 2$.

$$A^2 = \begin{bmatrix} 0 & 0 & 0 & 2 \\ & & & 0 \\ & & & 0 \\ & & & 0 \end{bmatrix},$$

So $\mathrm{rank}(A^2) = 1$, $\mathrm{rank}(A^3) = \mathrm{rank}(A^4) = 0$.

$$J = \begin{bmatrix} 0 & 1 & & \\ & 0 & 1 & \\ & & 0 & \\ & & & 0 \end{bmatrix}$$

4. $p_B(x) = x^2, p_A(x) = x^3$

5. If

$$p_B(x) = \prod_{i=1}^{s} (x - \lambda_i I)^{m_i},$$

where m_i is the algebraic multiplicity of λ_i, and s is number of distinct eigenvalues of B, then

$$p_A(x) = p_B(x) \cdot \prod_{i=1}^{s} (x - \lambda_i I)^{\mathrm{Dim}(\mathrm{Ker}(A - \lambda_i I))}$$

Solution 1.4.3.

1. $\forall a : a \in R$,

$$L(aI) = aX - Xa = aX - aX = 0$$

2.

$$\begin{aligned} L(X)Y + XL(Y) &= (AX - XA)Y + X(AY - YA) = AXY - XAY + XAY - XYA \\ &= A(XY) - (XY)A = L(XY) \end{aligned}$$

3. Here we prove the following lemma first:

$$S(n) : L(X^n) = n \cdot L(X) \cdot X^{n-1}$$

Showing that $S(n)$ holds $\forall n \in Z^+$.
$S(1)$ is trivial.
Suppose $S(k)$ holds, then for $S(k+1)$,

$$\begin{aligned} L(X^{k+1}) &= XL(X^k) + L(X)X^k \\ &= X(kL(X)X^{k-1}) + L(X)X^k = kL(X)X^k + L(X)X^k = (k+1)L(X)X^k, \end{aligned}$$

So $S(k) \Rightarrow S(k+1)$. By first principle of induction, $(\forall n)(n \in Z^+ \Rightarrow S(n))$.
$p(x)$ is a polynomial, i.e. it is analytic, and has finite terms in its Taylor's expansion, so W.L.O.G. let $p(x) = \sum_{i=0}^{n} a_i x^i$

$$L(p(X)) = L(\sum_{i=0}^{n} a_i \cdot X^i) = \sum_{i=0}^{n} a_i \cdot L(X^i) = \sum_{i=1}^{n} a_i \cdot L(X^i) = \sum_{i=1}^{n} a_i \cdot i \cdot L(X) \cdot X^{i-1} = L(X)p'(X)$$

4. Given $L(X) = I$, then for any polynomial p,

$$L(p(X)) = L(X)p'(X) = p'(X),$$

By choosing p as the minimal polynomial of X, deg $p > 0$, then

$$p'(X) = L(p(X)) = L(0) = 0,$$

So X is one of the zeros of p', that deg $p' <$ deg p, contradicting the minimality of p, hence $L(X) = I$ is not possible.

5. Given that A is diagonalizable with distinct eigenvalues, so W.L.O.G. let $A = P\Lambda P^{-1}$, then

$$L(X) = 0 \Leftrightarrow AX = XA \Leftrightarrow P\Lambda P^{-1}X = XP\Lambda P^{-1} \Leftrightarrow \Lambda P^{-1}XP = P^{-1}XP\Lambda$$

So $L(X) = 0$ iff $[P^{-1}XP, \Lambda] = 0$ iff $P^{-1}XP$ is diagonal iff A and X shares the same eigenvectors, and hence $\mathrm{Dim}(\mathrm{Ker}(L)) = n$.

6. Let t_1, t_2, t_3 are all distinct. Then $\forall X : X \in M_n(R)$,

$$L(X) = \begin{bmatrix} t_1 & & \\ & t_2 & \\ & & t_3 \end{bmatrix} X - X \begin{bmatrix} t_1 & & \\ & t_2 & \\ & & t_3 \end{bmatrix}$$

$$= \begin{bmatrix} t_1 & & \\ & t_2 & \\ & & t_3 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} - \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} t_1 & & \\ & t_2 & \\ & & t_3 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & (t_1 - t_2)a_{12} & (t_1 - t_3)a_{13} \\ (t_2 - t_1)a_{21} & 0 & (t_2 - t_3)a_{23} \\ (t_3 - t_1)a_{31} & (t_3 - t_2)a_{32} & 0 \end{bmatrix}$$

$\mathrm{Ran}(L)$ are hollow matrices. For instance,

$$A = \begin{bmatrix} 1 & & \\ & 2 & \\ & & 3 \end{bmatrix}$$

Solution 1.5.1.

First, examine the differentiability of f.
For $f : C \to C$, $f = u + iv$ is differentiable at $z_0 = x_0 + iy_0$ iff

$$\frac{\partial u}{\partial x}\bigg|_{x=x_0} = \frac{\partial v}{\partial y}\bigg|_{y=y_0}, \frac{\partial u}{\partial y}\bigg|_{y=y_0} = -\frac{\partial v}{\partial x}\bigg|_{x=x_0}$$

For $f(z) = z|z|$, $u(x,y) = x\sqrt{x^2+y^2}$, $v(x,y) = y\sqrt{x^2+y^2}$, then

$$u_x = \sqrt{x^2+y^2} + \frac{x^2}{\sqrt{x^2+y^2}}, u_y = \frac{xy}{\sqrt{x^2+y^2}}$$

$$v_x = \frac{xy}{\sqrt{x^2+y^2}}, v_y = \sqrt{x^2+y^2} + \frac{y^2}{\sqrt{x^2+y^2}}.$$

Solving $u_x = v_y$, $v_x = -u_y$ yields

$$\begin{cases} x^2 = y^2 \\ 2xy = 0 \\ x^2 + y^2 > 0 \end{cases}$$

However, no such tuple (x_0, y_0) satisfies all three equations above, and hence f is nowhere complex differentiable.
But if it is the case $f|_R : R \to R$, then it is real differentiable everywhere.

$f(x) = x|x| = x\sqrt{x^2}$, then $f'(x) = \sqrt{x^2} + x \cdot \frac{x}{\sqrt{x^2}} = 2|x|$

1. $$A_t = \begin{bmatrix} 1 & 1 \\ & 1+t \end{bmatrix} = \begin{bmatrix} 1 & -x \\ & 1 \end{bmatrix}\begin{bmatrix} 1 & \\ & 1+t \end{bmatrix}\begin{bmatrix} 1 & -x \\ & 1 \end{bmatrix}$$

   So $1 = -xt$, $x = -1/t$.

   $$f(A_t) = \begin{bmatrix} 1 & 1/t \\ & 1 \end{bmatrix}\begin{bmatrix} 1 & \\ & (1+t)^2 \end{bmatrix}\begin{bmatrix} 1 & -1/t \\ & 1 \end{bmatrix} = \begin{bmatrix} 1 & t+2 \\ & (1+t)^2 \end{bmatrix}, \forall t : t \geq -1$$

   So

   $$\lim_{t\to 0} f(A_t) = \lim_{t\to 0}\begin{bmatrix} 1 & t+2 \\ & (1+t)^2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ & 1 \end{bmatrix}$$

2. f is real differentiable, but not complex differentiable, which is obviously not analytic, so f cannot equal to its Taylor series, hence f is not defined for matrices with complex eigenvalues.
   Moreover,

   $$p_{A_t}(\lambda) = \det\left(\begin{bmatrix} 1-\lambda & 1 \\ -t^2 & 1-\lambda \end{bmatrix}\right) = (1-\lambda)^2 + t^2$$

   so for A_t, $\lambda = 1 \pm it$.
   As the limit is taking values from the punctured neighborhood of zero, the limit does not exist, so f is not defined.

Solution 1.5.2.

1. $\sin(z) = \frac{e^{iz}-e^{-iz}}{2i}$, so

   $$\sin(tA) = \frac{e^{itA} - e^{-itA}}{2i}, \frac{d}{dt}\sin(tA) = \frac{iAe^{itA} - (-iA)e^{-itA}}{2i} = \frac{A(e^{itA} + e^{-itA})}{2} = A\cos(tA)$$

2. W.L.O.G., assume f is analytic at $z = 0$ for $f : \mathbf{C} \to \mathbf{C}$.
   Then,

   $$f(z) = \sum_{k=0}^{+\infty} a_k z^k,$$

where

$$a_k = \frac{1}{2\pi i}\oint_{|z|=\epsilon}\frac{f(z)dz}{z^{n+1}}$$

Before seeking for the solution, first prove a lemma:

$$S(n):\begin{bmatrix} 2A & A \\ & 2A \end{bmatrix}^n = \begin{bmatrix} (2A)^n & n\cdot 2^{n-1}A^n \\ & (2A)^n \end{bmatrix}, \forall n : n\in \mathbf{Z}^+$$

$S(1)$ is trivial.
Assume $S(k)$ holds, then for $S(k+1)$,

$$\begin{bmatrix} 2A & A \\ & 2A \end{bmatrix}^{k+1} = \begin{bmatrix} 2A & A \\ & 2A \end{bmatrix}^{k}\begin{bmatrix} 2A & A \\ & 2A \end{bmatrix} = \begin{bmatrix} (2A)^k & k\cdot 2^{k-1}A^k \\ & (2A)^k \end{bmatrix}\begin{bmatrix} 2A & A \\ & 2A \end{bmatrix}$$

$$= \begin{bmatrix} (2A)^{k+1} & (2^k + k\cdot 2^k)A^{k+1} \\ & (2A)^{k+1} \end{bmatrix} = \begin{bmatrix} (2A)^{k+1} & (k+1)\cdot 2^k A^{k+1} \\ & (2A)^{k+1} \end{bmatrix}, S(k)\Rightarrow S(k+1)$$

By first principle of induction, $\forall n : n \in \mathbf{Z}^+ \Rightarrow S(n)$. ■
So

$$f\left(\begin{bmatrix} 2A & A \\ & 2A \end{bmatrix}\right) = \sum_{k=0}^{+\infty} a_k \begin{bmatrix} 2A & A \\ & 2A \end{bmatrix}^k = \sum_{k=0}^{+\infty}\begin{bmatrix} a_k(2A)^k & a_k\cdot k\cdot 2^{k-1}A^k \\ & a_k(2A)^k \end{bmatrix}$$

Let $X = BJB^{-1}$, then for $p(x) = x^n$, $p'(X) = Bp'(J)B^{-1}$, $p'(J) = nJ^{n-1}$, so $p'(X) = nX^{n-1}$

$$\Rightarrow f'(X) = \sum_{k=1}^{+\infty} a_k\cdot k\cdot X^{k-1}$$

$$\Rightarrow f\left(\begin{bmatrix} 2A & A \\ & 2A \end{bmatrix}\right) = \begin{bmatrix} f(2A) & f'(A)A \\ & f(2A) \end{bmatrix} \Rightarrow B = f'(A)\cdot A.$$

3. It can be disproved by letting $f(x) = x^2$, $A = \begin{bmatrix} 1 & 1 \\ & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & \\ & 3 \end{bmatrix}$

$$f'(A)B = \left(\begin{bmatrix} f'(1) & f''(1) \\ & f'(1) \end{bmatrix}\right)\begin{bmatrix} 2 & \\ & 3 \end{bmatrix} = \begin{bmatrix} 4 & 6 \\ & 6 \end{bmatrix}$$

$$f'\left(\begin{bmatrix} 1+2t & 1 \\ & 1+3t \end{bmatrix}\right) = \begin{bmatrix} 1 & -1/t \\ & 1 \end{bmatrix}\begin{bmatrix} 2+4t & \\ & 2+6t \end{bmatrix}\begin{bmatrix} 1 & 1/t \\ & 1 \end{bmatrix} = \begin{bmatrix} 2+4t & -2 \\ & 2+6t \end{bmatrix}$$

Plugging in $t = 0$ yields $\begin{bmatrix} 2 & -2 \\ & 2 \end{bmatrix}$.
Actually, the proposition holds only when $[A, B] = 0$. In this case, binomial theorem is suitable for matrices A and B.

$$[f(A+tB)]' = \left(\sum_{n=0}^{+\infty}\sum_{k=0}^{n} a_n C_n^k A^{n-k}(Bt)^k\right)' = \sum_{n=0}^{+\infty}\sum_{k=0}^{n} a_n C_n^k A^{n-k} k B^k t^{k-1}$$

Plugging in $t = 0$,

$$[f(A+tB)]'|_{t=0} = \sum_{n=1}^{+\infty} a_n C_n^1 A^{n-1} B = f'(A)\cdot B$$

Solution 1.5.3.

1. Let v be the common eigenvector of A and B, λ_A and λ_B are the two eigenvalues associated with v for A and B. V_2 is a matrix such that the column vectors are orthonormal basis of orthogonal complement of span of v. Then for A,

$$\begin{bmatrix} v & V_2 \end{bmatrix}^H A \begin{bmatrix} v & V_2 \end{bmatrix} = \begin{bmatrix} v & V_2 \end{bmatrix}^H \begin{bmatrix} \lambda_A v & AV_2 \end{bmatrix} = \begin{bmatrix} \lambda_A & v^H AV_2 \\ 0 & V_2^H AV_2 \end{bmatrix}$$

Similarly, for B,

$$\begin{bmatrix} v & V_2 \end{bmatrix}^H B \begin{bmatrix} v & V_2 \end{bmatrix} = \begin{bmatrix} \lambda_B & v^H B V_2 \\ 0 & V_2^H B V_2 \end{bmatrix}$$

So $A_1 = V_2^H A V_2$, $B_1 = V_2^H B V_2$.

$$A_1 B_1 = V_2^H A V_2 V_2^H B V_2 = V_2^H A B V_2$$
$$B_1 A_1 = V_2^H B V_2 V_2^H A V_2 = V_2^H B A V_2$$

The above two equations are equal as A commutes with B, i.e.

$$[A, B] = 0 \Rightarrow [A_1, B_1] = 0$$

2. From (1), $[A_1, B_1] = 0$, so A_1, B_1 should have a common eigenvector. We may assume that it is one of the column vectors of V_2, otherwise V_2 can be reconstructed in the previous step to contain that common eigenvector.
Moveover, when that common eigenvector is placed on the first column of V_2 in (1), then

$$A_1 = \begin{bmatrix} \lambda_{A_1} & v'^H B V_{22} \\ 0 & V_{22}^H A_1 V_{22} \end{bmatrix},$$

where v' is the common eigenvector of A_1 and B_1. Same for B_1.
Hence we can iteratively apply that transformation to $(A_i)_{22}$ and $(B_i)_{22}$ for all i.
In addition, for each iteration, the size of A_i, B_i strictly decreases.
Given that A and B are finite dimensional, the process of iteration eventually terminates.
The iterations terminate when i is such that A_i and B_i has dimension one.
That gives

$$\begin{bmatrix} v & V_2 \end{bmatrix}^H A \begin{bmatrix} v & V_2 \end{bmatrix}$$

an upper triangular matrix. Same for B. ■

Solution 1.6.1.

1.

$$L = \begin{bmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{bmatrix}$$

L is commonly known as Vandermonde matrix.

2. For any square Vandermonde matrix,

$$\det(L) = \prod_{1 \le i < j \le n} (a_j - a_i)$$

$det(L) \neq 0$ iff none of the factors above is zero iff $a_i \neq a_j$, $\forall i, j : 1 \le i < j \le n$

3. $a_i \neq a_j$, $\forall i, j : 1 \le i < j \le n \Leftrightarrow det(L) \neq 0 \Leftrightarrow$ all row vectors of $L (= \mathrm{ev}_{a_i})$ are linearly independent.
n independent vectors are required to span V^*, so $\mathrm{ev}_{a_i}, 1 \le i \le n$ are linearly independent $\Leftrightarrow$ $\mathrm{ev}_{a_1}, \cdots, \mathrm{ev}_{a_n}$ form a basis of V^*.

4. The three vectors should satisfy

$$\begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} p_1 & p_2 & p_3 \end{bmatrix} = I_3$$

So inverting that Vandermonde matrix solves the three polynomials.

$$\begin{bmatrix} 1 & -1 & 1 & 1 & & \\ 1 & 0 & 0 & 0 & 1 & \\ 1 & 1 & 1 & 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 1 & 1 & & \\ 0 & 1 & -1 & -1 & 1 & \\ 0 & 0 & 1 & 0.5 & -1 & 0.5 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & -0.5 & 0 & 0.5 \\ 0 & 0 & 1 & 0.5 & -1 & 0.5 \end{bmatrix}$$

Therefore,

$$p_{-1}(x) = -\frac{1}{2}x + \frac{1}{2}x^2, p_0(x) = 1 - x^2, p_1(x) = \frac{1}{2}x + \frac{1}{2}x^2$$

5. This is equivalent to finding the cokernel of the Vandermonde matrix

$$x^T L = \begin{bmatrix} 1 & -2 & 4 & -8 \\ 1 & -1 & 1 & -1 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \end{bmatrix} = 0$$

So

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ -2 & -1 & 0 & 1 & 2 \\ 4 & 1 & 0 & 1 & 4 \\ -8 & -1 & 0 & 1 & 8 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 & 4 \\ 0 & -3 & -4 & -3 & 0 \\ 0 & 7 & 8 & 9 & 16 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 2 & 6 & 12 \\ 0 & 0 & -6 & -12 & -12 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 3 & 6 \\ 0 & 0 & 0 & 1 & 4 \end{bmatrix}, x = x_5 \cdot \begin{bmatrix} 1 \\ -4 \\ 6 \\ -4 \\ 1 \end{bmatrix}$$

Therefore,

$$\mathrm{ev}_{-2} - 4\mathrm{ev}_{-1} + 6\mathrm{ev}_0 - 4\mathrm{ev}_1 + \mathrm{ev}_2 = 0$$

Solution 1.6.2.

Dual vectors are linear maps, so linearity should be checked.

1. It is.

$$0 \mapsto \mathrm{ev}_5((x+1)\cdot 0) = 0$$
$$\forall p, q : p, q \in V, \forall c_1, c_2 : c_1, c_2 \in F$$
$$\begin{aligned} \mathrm{ev}_5((x+1)\cdot(c_1 p + c_2 q)(x)) &= 6(c_1 p + c_2 q)(5) = c_1 \cdot 6p(5) + c_2 \cdot 6q(5) \\ &= c_1 \mathrm{ev}_5((x+1)p(x)) + c_2 \mathrm{ev}_5((x+1)q(x)) \end{aligned}$$

2. It is not. Suppose $p(x) = 1 - x^2$, $q(x) = 1 + x^2$,

$$(p+q) \mapsto \lim_{x\to+\infty} \frac{(p+q)(x)}{x} = 0,$$

But

$$\lim_{x\to+\infty} \frac{p(x)}{x} + \lim_{x\to+\infty} \frac{q(x)}{x} = -\infty + \infty$$

is not defined.

3. It is. Suppose $p(x) = a_1 + b_1 x + c_1 x^2$, $q(x) = a_2 + b_2 x + c_2 x^2$,

$$\lim_{x\to+\infty} \frac{(Ap+Bq)(x)}{x^2} = \lim_{x\to+\infty} \left(\frac{Aa_1 + Ba_2}{x^2} + \frac{Ab_1 + Bb_2}{x} + (Ac_1 + Bc_2) \right) = Ac_1 + Bc_2,$$
$$\lim_{x\to+\infty} \frac{Ap(x)}{x^2} = \lim_{x\to+\infty} \left(\frac{Aa_1}{x^2} + \frac{Ab_1}{x} + (Ac_1) \right) = Ac_1,$$
$$\lim_{x\to+\infty} \frac{Bq(x)}{x^2} = \lim_{x\to+\infty} \left(\frac{Ba_2}{x^2} + \frac{Bb_2}{x} + (Bc_2) \right) = Bc_2$$

Sum up the last two equations yields the first. Moreover,

$$0 \mapsto \lim_{x\to+\infty} \frac{0}{x^2} = 0$$

4. It is not.

$$[(x^2 + x)] \mapsto (3^2 + 3)(2(4) + 1) = 108$$
$$[x^2] \mapsto (3^2)(2(4)) = 72$$
$$[x] \mapsto (3)(1) = 3$$

Sum up the last 2 equations, it does not equal to the first.

5. It is not. Suppose $p(x) = x$, then
$$p \mapsto deg(p) = 1,$$
But
$$2p \mapsto deg(2p) = 1,$$
So double of the first equation cannot yield the second, it is not linear.

Solution 1.6.3.

Since $\nabla f : \mathbf{R}^2 \to \mathbf{R}$, if the map is linear, it is a dual vector.
The following lemma should be proved first:
Lemma: Given differentiable function f, its directional derivative with direction $\mathbf{u}$ is

$$\nabla_{\mathbf{u}}(f) = \nabla(f) \cdot \mathbf{u}$$

where $\nabla(f) = \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{bmatrix}$.
Proof:
Directional derivative of f at $\mathbf{p}$ with direction u is

$$\nabla_u(f) = \lim_{t \to 0} \frac{f(p_x + tu_x, p_y + tu_y) - f(p_x, p_y)}{t}$$

Since f is differentiable, at (x_0, y_0), then

$$\Delta f = f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0) = f'_x(x_0, y_0)\Delta x + f'_y(x_0, y_0)\Delta y + \epsilon_1 \Delta x + \epsilon_2 \Delta y,$$

where $\epsilon_1, \epsilon_2 \to 0$ when $\Delta x, \Delta y \to 0$ (as $\frac{\epsilon_1 \Delta x + \epsilon_2 \Delta y}{\sqrt{(\Delta x)^2 + (\Delta y)^2}} \le |\epsilon_1| + |\epsilon_2|$). So,

$$\lim_{t \to 0^+} \frac{f(p_x + tu_x, p_y + tu_y) - f(p_x, p_y)}{t}$$
$$= \lim_{t \to 0^+} \left(f'_x(p_x, p_y) \frac{tu_x}{t} + f'_y(p_x, p_y) \frac{tu_y}{t} \right) = f'_x(p_x, p_y)u_x + f'_y(p_x, p_y)u_y \quad \square$$

With regards to the zero element,
$$\nabla_0(f) = \nabla(f) \cdot 0 = 0$$
With regards to linearity,

$$\nabla_{au+bv}(f) = \nabla(f) \cdot (au + bv) = a\nabla(f) \cdot u + b\nabla(f) \cdot v = a\nabla_u(f) + b\nabla_v(f)$$

It is trivial from the derivation above that the coordinates of the dual vector, is exactly the gradient of f.

Solution 1.6.4.

Denote U^0 as the annihilator of U.

1. Let $\alpha \in W^*$.

$$\alpha \in (\mathrm{Ran}(L))^0 \Leftrightarrow \alpha(L(v)) = 0 (\forall v \in V) \Leftrightarrow \alpha \circ L = 0 \Leftrightarrow L^*(\alpha) = 0 \Leftrightarrow \alpha \in \mathrm{Ker}(L^*)$$

   Hence $(\mathrm{Ran}(L))^0 = \mathrm{Ker}(L^*)$.

2. Let $\beta \in V^*$.
   On one hand,

$$\beta \in \mathrm{Ran}(L^*) \Leftrightarrow \exists \omega \in W^* : \beta = L^*(\omega) \Leftrightarrow \exists \omega \in W^* : \beta(v) = \omega \circ L(v)$$

   In particular, when $v \in \mathrm{Ker}(L), \beta(v) = \omega \circ L(v) = 0$.
   Therefore $\mathrm{Ran}(L^*) \subseteq (\mathrm{Ker}(L))^0$.
   Before proving the backward direction, define $\sim$ as an equivalence relation that:

$$x \sim y \Leftrightarrow L(x) = L(y)$$

It is trivial that ~ is reflexive, symmetric and transitive. In particular,

$$[x] = [y] \Leftrightarrow (x - y) \in \mathrm{Ker}(L)$$

Let $f : V \to V/\sim$ be a surjective function such that $f : v \mapsto [v]$. Then $L = L_1 \circ f$, where $L_1 : V/\sim \to \mathrm{Ran}(L)$ is bijective. Immediately we have $f = L_1^{-1} \circ L$.
So on the other hand, if $\beta \in (\mathrm{Ker}(L))^0$, $\mathrm{Ker}(L) \subseteq \mathrm{Ker}(\beta)$. Moreover β can be factorized as $\beta = \xi \circ f$, where $\xi : V/\sim \to \mathbf{R}$.

$$\beta = \xi \circ f = \xi \circ (T_1^{-1} \circ T) = (\xi \circ T_1^{-1}) \circ T = T^*(\xi \circ T_1^{-1})$$

So, $(\mathrm{Ker}(L))^0 \subseteq \mathrm{Ran}(L^*)$. Hence $(\mathrm{Ker}(L))^0 = \mathrm{Ran}(L^*)$.

Solution 1.7.1.

1. To show $(\cdot, \cdot)$ is inner product, bilinearity, symmetry and positive definitiveness of it should be shown. It is trivial that $A \in M_n(R)$.

   i) Bilinearity:
   $\forall u, u_1, u_2, v, v_1, v_2 \in R^n,\ a, b \in R$,

   $$(au_1 + bu_2, v) = (au_1 + bu_2)^T Av = au_1^T Av + bu_2^T Av = a(u_1, v) + b(u_2, v)$$

   $$(u, av_1 + bv_2) = u^T A(av_1 + bv_2) = a(u^T Av_1) + b(u^T Av_2) = a(u, v_1) + b(u, v_2)$$

   ii) Symmetry:

   $$(u, v) = u^T Av = (u^T Av)^T = v^T A^T u = v^T Au = (v, u)$$

   iii) Positive definitiveness:
   By definition, A is positive definite, therefore $\forall x \in R^n$,

   $$x^T Ax \geq 0$$

   In addition, $x^T Ax = 0 \Leftrightarrow x = 0$.

2. Suppose u is the pre-image of v^T of the bra map. Then,

   $$v^T = u^T A \Leftrightarrow v^T A^{-1} = u^T \Leftrightarrow u = (A^{-1})^T v$$

   So the Riesz map is

   $$v^T \mapsto A^{-1} v$$

3. Immediately from (2),

   $$v \mapsto v^T A$$

4. Riesz is $V^* \to V$, so dual of Riesz is $V^* \to V^{**}$, i.e. v^T would be sent to an image such that the image is a linear functional: $V^* \to R$.

   $$[\mathrm{Riesz}^*(\alpha)](\beta)] = \alpha \circ \mathrm{Riesz}(\beta) = \beta A^{-1} \alpha^T$$

   For $\alpha, \beta \in V^*$. So the dual will send $\alpha \in V^*$ in the following way:

   $$\alpha \mapsto (\beta \mapsto \beta A^{-1} \alpha^T)$$

Solution 1.7.2.

1. v is a linear functional, hence showing v sends constant to zero requires v sending one to zero.
   Let $f(x) = a, g(x) = b$, then $v(fg) = abv(1)$, so we may let $f(x) \equiv g(x) \equiv 1$,

   $$v(1) = f(p)v(g) + v(f)g(p) = v(1) + v(1)$$

   The only possible solution for the above equation is $v(1) = 0$.

2. Let $x_1 = x, x_2 = y, x_3 = z$, then

$$v((x_i - p_i)f) = (x_i - p_i)(\mathbf{p})v(f) + v(x_i - p_i)f(\mathbf{p}) = (p_i - p_i)v(f) + v(x_i)f(\mathbf{p}) - v(p_i)f(\mathbf{p}) = v(x_i)f(\mathbf{p})$$

$(i = 1, 2, 3)$

3. By applying Leibniz's formula repeatedly,

$$v(fgh) = v(f)g(\mathbf{p})h(\mathbf{p}) + f(\mathbf{p})v(g)h(\mathbf{p}) + f(\mathbf{p})g(\mathbf{p})v(h)$$

a, b, c are non-negative integers, and $a + b + c > 1$, so W.L.O.G. let

$$(w_1 - p_i)(w_2 - p_j)Q = (x - p_1)^a(y - p_2)^b(z - p_3)^c,$$

where Q is a function of x, y, z and the two factors prior to it comes from the factor of the non-zero powers of $(x - p_1)^a(y - p_2)^b(z - p_3)^c$.
For instance, when $a = 2$, then $w_1 = w_2 = x, p_i = p_j = p_1$, when $a = b = 1$, $w_1 = x, w_2 = y, p_i = p_1, p_j = p_2$. Then

$$v((x - p_1)^a(y - p_2)^b(z - p_3)^c) = v((w_1 - p_i)(w_2 - p_j)Q)$$
$$= v(w_1 - p_i)(w_2 - p_j)(\mathbf{p})Q(\mathbf{p}) + (w_1 - p_i)(\mathbf{p})v(w_2 - p_j)Q(\mathbf{p}) + (w_1 - p_i)(\mathbf{p})(w_2 - p_j)(\mathbf{p})v(Q)$$

By setting $f \equiv 1$ in (2), the last expression immediately yields zero.

4. Taylor expansion of f at p is

$$f(x, y, z) = \sum_{n=0}^{+\infty} \sum_{\substack{i,j,k \geq 0 \\ i+j+k=n}} \left. \frac{\partial^{(i+j+k)} f}{\partial x^i \partial y^j \partial z^k} \right|_{\mathbf{p}} (x - p_1)^i (y - p_2)^j (z - p_3)^k$$

Since applying v to constant yields zero, so does those polynomial with degree ≥ 2, therefore

$$v(f) = v(f'_x(\mathbf{p})(x - p_1)) + v(f'_y(\mathbf{p})(y - p_2)) + v(f'_z(\mathbf{p})(z - p_3))$$
$$= f'_x(\mathbf{p})v(x) + f'_y(\mathbf{p})v(y) + f'_z(\mathbf{p})v(z)$$

5. It is proved in solution 1.6.3. that, for any differentiable f (which is done by generalizing the proof to higher dimension), $\nabla_u(f) = \nabla f \cdot u$, so for any $v = v(x)\hat{\mathbf{i}} + v(y)\hat{\mathbf{j}} + v(z)\hat{\mathbf{k}}$, and for any analytic f,

$$\nabla_v(f) = \nabla f \cdot v = f'_x(\mathbf{p})v(x) + f'_y(\mathbf{p})v(y) + f'_z(\mathbf{p})v(z) = v(f)$$

where f is arbitrary analytic function, so $\nabla_v = v$. ■

Solution 1.7.3.

1.
$$X(fg) = X(f)g + fX(g),$$

So $\forall p : p \in M$,

$$X(fg)(p) = X(f)(p) \cdot g(p) + f(p) \cdot X(g)(p)$$
$$\Leftrightarrow X_p(fg) = X_p(f) \cdot g(p) + f(p) \cdot X_p(g)$$

The last equation satisfies Leibniz's rule. □

2. $\forall p \in M$,

$$[X(f)](p) = X_p(f) = \nabla f|_p(X_p) = df|_p(X_p) = [df(X)](p)$$

Hence, $X(f) = df(X)$. ■

3. Let f and g be two analytic functions, then the following should be proved:

$$(X \circ Y - Y \circ X)(fg) = [(X \circ Y)(f)]g + f[(X \circ Y - Y \circ X)(g)]$$

L.H.S.

$$= (X \circ Y - Y \circ X)(fg) = (X \circ Y)(fg) - (Y \circ X)(fg)$$

$$= X(Y(f)g + fY(g)) - Y(X(f)g + fX(g)) = X(Y(f)g) + X(fY(g)) - Y(X(f)g) - Y(fX(g))$$
$$= X(Y(f))g + Y(f)X(g) + X(f)Y(g) + fX(Y(g)) - Y(X(f))g - X(f)Y(g) - Y(f)X(g) - fY(X(g))$$
$$= X(Y(f))g + fX(Y(g)) - Y(X(f))g - fY(X(g))$$

R.H.S.

$$= [(X \circ Y - Y \circ X)(f)]g + f[(X \circ Y - Y \circ X)(g)]$$
$$= X(Y(f))g - Y(X(f))g + fX(Y(g)) - fY(X(g))$$

Therefore L.H.S. = R.H.S. ■

4. A and B are both skew-symmetric, so

$$(AB-BA)^T = (AB)^T - (BA)^T = B^T A^T - A^T B^T = (-B)(-A) - (-A)(-B) = BA - AB = -(AB-BA) ■$$

* Credit to Fan Sunqi for some parts in Sol. 1.8.2.

Solution 1.8.1.

1. $(\alpha \otimes \beta \otimes \gamma)_{ijk} = \alpha_i \beta_j \gamma_k$, where subscript n is the index of component of the row vector.

2. **Lemma1 :**
Let $U = \alpha \otimes \beta \otimes \gamma$, $\mathcal{K}(u, v, w) = u \otimes v \otimes w$, L be a 1×8 row vector that its entries are in the form of $\alpha_i \beta_j \gamma_k$, where i, j, k are put in lexicographic order. Then, $U = L \circ \mathcal{K}$.
Proof 1:
$U(u, v, w) = \alpha(u)\beta(v)\gamma(v)$. Expand the R.H.S. directly follows.
Lemma 2:
$M_E(U) = M'_E(L) \circ \mathcal{K}$, where M'_E is just a matrix, whenever E is an elementary matrix.
Proof 2:
Case 1: $E = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, then $M'_E(L) = L \cdot \begin{bmatrix} 0 & I_4 \\ I_4 & 0 \end{bmatrix}$.
Case 2: $E = \begin{bmatrix} s_1 & 0 \\ 0 & s_2 \end{bmatrix}$, then $M'_E(L) = L \cdot \begin{bmatrix} s_1 I_4 & 0 \\ 0 & s_2 I_4 \end{bmatrix}$
Case 3: $E = \begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}$, then $M'_E(L) = L \cdot \begin{bmatrix} I_4 & 0 \\ kI_4 & I_4 \end{bmatrix}$
Let $U_1 = L_1 \circ \mathcal{K}$, $U_2 = L_2 \circ \mathcal{K}$, then $aU_1 + bU_2 = aL_1 \circ \mathcal{K} + bL_2 \circ \mathcal{K} = (aL_1 + bL_2) \circ \mathcal{K}$,

$$M_E(aU_1 + bU_2) = M'_E(aL_1 + bL_2) \circ \mathcal{K} = aM'_E(L_1) \circ \mathcal{K} + bM'_E(L_2) \circ \mathcal{K} = aM_E(U_1) + bM_E(U_2)$$

Other matrices in $((R^2)^*)^{\otimes 3}$ are just linear combinations of the U 's, the proof directly follows.

3. Suppose $M = \sum_{i=1}^{r} a_i U_i$, where $U_i =$, $a_i \neq 0$ and $\sum_{i=1}^{r} t_i U_i = 0$ yields $t_1 = t_2 = \cdots = t_r = 0$, then

$$M_E(M) = M_E\left(\sum_{i=1}^{r} a_i U_i\right) = \sum_{i=1}^{r} a_i M_E(U_i)$$

By the way, M_E is bijective, since elementary operations E are always invertible, so $M_E(M)$ is a rank r tensor.

4. Assume the 3D matrix M to have rank less than r, then $M = \sum_{i=1}^{r'} a_i U_i$, $r' < r$ and U_i are simple tensors. Then for each $a_i U_i$, that layer has rank 1, and the sum yields that, that layer can only have rank less than r, which is a contradiction.

5. It is rank two, since $M = ([1, -1] \otimes [1, -1] \otimes [1, -1] + [1, 1] \otimes [1, 1] \otimes [1, 1])/2$.

Solution 1.8.2.

1. $$M(v,v,v) = v^T \left(x \begin{bmatrix} 3 & 4 & 5 \\ 4 & 5 & 6 \\ 5 & 6 & 7 \end{bmatrix} + y \begin{bmatrix} 4 & 5 & 6 \\ 5 & 6 & 7 \\ 6 & 7 & 8 \end{bmatrix} + z \begin{bmatrix} 5 & 6 & 7 \\ 6 & 7 & 8 \\ 7 & 8 & 9 \end{bmatrix} \right) v$$
   It can be treated as a "cubic form", analogous to quadratic forms, so it is
   $$M(v,v,v) = 3x^3 + 6y^3 + 9z^3 + 12x^2y + 15xy^2 + 15x^2z + 21xz^2 + 21y^2z + 24yz^2 + 36xyz$$

2. Similar to 1.8.1, convert M to be a composition of a 1×27 row vector and the Kronecker product of the three input vectors, $M = L \circ \mathcal{K}$. Then, for each entry in $\mathcal{K}(v_1, v_2, v_3)$, it must be in the form of $v_1^i v_2^j v_3^k$, where superscript denotes the position of entry in its own vector.
   Let σ' be an element in a group that isomorphic to S_6, that $\sigma'(v_1, v_2, v_3) = (v_{\sigma(1)}, v_{\sigma(2)}, v_{\sigma(3)})$, then σ' induces σ'' which is another element that comes from another group isomorphic to S_6 and σ'' correspond to σ',
   $$\mathcal{K}(\sigma'(v_1, v_2, v_3)) = \sigma'' \circ \mathcal{K}(v_1, v_2, v_3)$$
   $L \circ \sigma''$ gives a row vector that permutates around the entries of L. As entries of L implicitly refers to the entries of M, $L = [3, 4, 5, 4, 5, 6, \cdots]$, that $L_1 = (1,1,1)$ entry, $L_2 = (1,1,2)$ entry and so on, $L \circ \sigma''$ permutates the subscript of the entry that the entry of L refers to in M. However, in M, $M_{ijk} = i + j + k$ is invariant over permutation, so $L \circ \sigma'' = L$.
   Intuitively, M can be thought as a cube that is invariant over "rotations of triangles that results rotational symmetry", on the axis that pass through $(1,1,1)$ and $(3,3,3)$. In addition, straighten up the axis. Watch the cube from the top, then it has two-fold reflectional symmetry.

3. On one hand, rank of M is not less than 2 by 1.8.1.4.
   On the other hand, the upper layer and the lower layer can be treated as adding or subtracting a layer of matrix with all ones.
   Further crack down the layer of $\begin{bmatrix} 4 & 5 & 6 \\ 5 & 6 & 7 \\ 6 & 7 & 8 \end{bmatrix}$: it is a rank two matrix,
   $$\begin{bmatrix} 4 & 5 & 6 \\ 5 & 6 & 7 \\ 6 & 7 & 8 \end{bmatrix} = [5,6,7]^T[1,1,1] + [1,1,1]^T[-1,0,1]$$
   Moreover,
   $$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} = [1,1,1]^T[1,1,1]$$
   So,
   $$M = [5,6,7] \otimes [1,1,1] \otimes [-1,0,1] + [1,1,1] \otimes [-1,0,1] \otimes [-1,0,1] + [1,1,1] \otimes [1,1,1] \otimes [-1,0,1]$$
   M is expressed as sum of three simple tensors, so its rank is at most 3.

Solution 1.9.1.

1. Proposition 10.5.16. yields $Au \otimes Bv = (A \otimes B)(u \otimes v)$.
   $u \otimes v$ is Kronecker product, and by proposition 10.5.14, Kronecker product is bilinear.
   Since $A \otimes B$ is linear, the map is the product of a linear map, and a bilinear map, which is automatically bilinear.

2. Since it is finite dimensional, the space should be isomorphic to some vector space.
   Let X and Y in 1 to be the matrix in a vector space in which X and Y has Jordan blocks on its diagonal and elsewhere zero.
   Then $X \otimes Y$ is an upper diagonal matrix, in which the diagonal elements are $\lambda_i \mu_j$, $i = 1, 2, \cdots, \mathrm{Dim}(V)$, $j = 1, 2, \cdots, \mathrm{Dim}(W)$, in which λ, μ are eigenvalues of X and T.
   So
   $$\mathrm{trace}(X \otimes Y) = \sum_{i=1}^{\mathrm{Dim}(V)} \sum_{j=1}^{\mathrm{Dim}(W)} \lambda_i \mu_j = \sum_{i=1}^{\mathrm{Dim}(V)} \lambda_i \sum_{j=1}^{\mathrm{Dim}(W)} \mu_j = \sum_{i=1}^{\mathrm{Dim}(V)} \lambda_i \mathrm{trace}(Y) = \mathrm{trace}(X)\mathrm{trace}(Y)$$

Solution 1.9.2.

$$\text{trace} \in \mathcal{L}(R^n \otimes (R^n)^*, R) = (R^n)^* \otimes (R^n)^{**} \otimes R = (R^n)^* \otimes R^n$$

Let r be the rank of a particular matrix A, then

$$\text{trace}(A) = \sum_{i=1}^{r} t_i \text{trace}(u_i v_i^T) = \sum_{i=1}^{r}\sum_{j=1}^{n} t_i u_i^j v_i^j$$

Therefore, it is just a dot product of the rightmost vectors.
Dot product is a special kind of inner product, in which it is the case for solution 1.7.1. when $A = I$, and hence

$$\text{trace}_{ij} = \delta_{ij}$$

Solution 1.9.3.

1. Symmetry:

$$(v_1 \otimes w_1, v_2 \otimes w_2) = (v_1, v_2)(w_1, w_2) = (v_2, v_1)(w_2, w_1) = (v_2 \otimes w_2, v_1 \otimes w_1)$$

   Positive definiteness:

$$(v \otimes w, v \otimes w) = (v, v)(w, w)$$

   It is zero iff $v = 0$ or $w = 0$, iff $v \otimes w = 0$

2.

$$(e_1 \otimes e_1, e_2 \otimes e_2) = (e_1, e_2)(e_1, e_2)$$

   Since e_1, e_2 are two basis vectors, their inner product must not be ± 1. (or else they lie on the same line and cannot be basis vectors). There are two components, so the rank must not be greater than 2. These two simple tensors are not proportional, so rank must not be less than 2.

3. $\omega = v \otimes w$

$$(\omega, L \otimes I_B(\omega)) = (v \otimes w, Lv \otimes w) = (v_1^2 - v_2^2)(w_1^2 + w_2^2)$$
$$(\omega, I_A \otimes L(\omega)) = (v \otimes w, v \otimes Lw) = (v_1^2 + v_2^2)(w_1^2 - w_2^2)$$

   Let the former to be x, and the latter to be y. Partition R^2 into four regions by the lines $x + y = 0$ and $x - y = 0$.
   Case 1: $x + y \le 0 \wedge x - y \le 0$. Then let

$$v_1 = 0, v_2 = \pm 1 \Rightarrow w_1 = \pm\sqrt{-\frac{x-y}{2}}, w_2 = \pm\sqrt{-\frac{x+y}{2}}$$

   Case 2: $x + y \ge 0 \wedge x - y \ge 0$. Then let

$$v_1 = \pm 1, v_2 = 0 \Rightarrow w_1 = \pm\sqrt{\frac{x+y}{2}}, w_2 = \pm\sqrt{\frac{x-y}{2}}$$

   Case 3: $x + y \ge 0 \wedge x - y \le 0$. Then let

$$w_1 = \pm 1, w_2 = 0 \Rightarrow v_1 = \pm\sqrt{\frac{x+y}{2}}, v_2 = \pm\sqrt{-\frac{x-y}{2}}$$

   Case 3: $x + y \le 0 \wedge x - y \ge 0$. Then let

$$w_1 = 0, w_2 = \pm 1 \Rightarrow w_1 = \pm\sqrt{-\frac{x+y}{2}}, w_2 = \pm\sqrt{\frac{x-y}{2}}$$

   Hence it is possible to take any point of $(x, y) \in R^2$.

4.

$$L \otimes I_B(\omega) = L \otimes I_B(ae_1 \otimes e_1) + L \otimes I_B(be_2 \otimes e_2) = aLe_1 \otimes e_1 + bLe_2 \otimes e_2 = ae_1 \otimes e_1 - be_2 \otimes e_2$$

$$I_A \otimes L(\omega) = I_A \otimes L(ae_1 \otimes e_1) + I_A \otimes L(be_2 \otimes e_2) = ae_1 \otimes Le_1 + be_2 \otimes Le_2 = ae_1 \otimes e_1 - be_2 \otimes e_2$$

   So they are indeed identical.

Solution 1.10.1.

1.
$$T^{11} = T^{22} = -\frac{1}{\sqrt{2}} - 1, T^{12} = T^{21} = -1$$

2. Indeed, by observation, the two columns in $\mathrm{Rot}(\pi/4)$ are possible to be x and y, so let

$$\mathbf{x} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{y} = \frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Then,

$$\mathbf{x} \otimes \mathbf{x}$$

has all entries 1/2, and

$$\mathbf{y} \otimes \mathbf{y}$$

has entry 1/2 when having equal indices, and $-1/2$ when having distinct indices.
Immediately

$$T = -(\sqrt{2}+1)\mathbf{x} \otimes \mathbf{x} - \mathbf{y} \otimes \mathbf{y}$$

3. By observing the result of (2) and the definiton of T, the resultant magnitude of force in $\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ direction is larger than that of $\frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ direction, so the direction of major axis is $\frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 1 \end{bmatrix}$, while the direction of minor is $\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$

4. As the force are "normal" to the circle,

$$T = \sum_{k=1}^{r} c_k \mathbf{x}_k \otimes \mathbf{x}_k$$

where $\mathbf{x}_k \in R^2$, $c_k \in R$, $\|\mathbf{x}_k\| = 1$(in different directions), $c_k < 0$, $k = 1, 2, \cdots, r$.
(Here only only consider forces on upper half of the circle, and the force on lower part are moved to the position π rad apart of the original position of the circle, with direcction reversed). Let $\mathbf{x_k} = a_k i + b_k j$, then Matrix form of T_{ij} is T^i_j

$$T^i_j = \begin{bmatrix} \sum_{k=1}^r c_k(a_k)^2 & \sum_{k=1}^r c_k a_k b_k \\ \sum_{k=1}^r c_k a_k b_k & \sum_{k=1}^r c_k (b_k)^2 \end{bmatrix}$$

By Cauchy-Schwarz inequality,

$$\det(T^i_j) = \left(\sum_{k=1}^{r} c_k(a_k)^2\right)\left(\sum_{k=1}^{r} c_k(b_k)^2\right) - \left(\sum_{k=1}^{r} c_k a_k b_k\right)^2$$

$$= \left(\sum_{k=1}^{r} (\sqrt{-c_k}a_k)^2\right)\left(\sum_{k=1}^{r} (\sqrt{-c_k}b_k)^2\right) - \left(\sum_{k=1}^{r} (\sqrt{-c_k}a_k)(\sqrt{-c_k}b_k)\right)^2 \geq 0$$

Equality holds iff

$$\frac{\sqrt{-c_k}a_k}{\sqrt{-c_k}b_k} = C$$

However, it is not possible, unless $r = 1$.
And trivially, $\mathrm{tr}(T^i_j) < 0$, so the matrix is negative semi-definite. It is negative definite iff $r > 1$.
Moreover, the matrix is symmetric, when the matrix representation is not symmetric, or when it is not negative semi-definite, there must be force that is not perpendicularly towards the center of circle.
For instance, $T^{ij} = \delta^{ij}$ is not possible.

Solution 1.10.2.

1. $\forall \alpha \in V^*, v \in V,$

$$\alpha_{\mathcal{B}}(v_{\mathcal{B}}) = \alpha_{\mathcal{C}}(v_{\mathcal{C}}) = \alpha_{\mathcal{C}}(Mv_{\mathcal{B}})$$
$$\Leftrightarrow \alpha_{\mathcal{B}} = \alpha_{\mathcal{C}} M \Leftrightarrow \alpha_{\mathcal{C}} = \alpha_{\mathcal{B}} M^{-1}$$

2.
$$T(v,w) = (\sum_{i,j} x_{ij} b_i^* \otimes b_j^*)(u,v) = \sum_{i,j} x_{ij} b_i^*(v) b_j^*(w) = x_{ij} v_{\mathcal{B}}^i w_{\mathcal{B}}^j = v_{\mathcal{B}}^i \delta_{ik} x_j^k w_{\mathcal{B}}^j = v_{\mathcal{B}}^T T_{\mathcal{B}} w_{\mathcal{B}}$$

3. Replacing x and b to y and c in the proof above immediately yields

$$v_{\mathcal{C}}^T T_{\mathcal{C}} w_{\mathcal{C}} = T(u,v) = v_{\mathcal{B}}^T T_{\mathcal{B}} w_{\mathcal{B}}$$

Moreover, $v_{\mathcal{C}} = Mv_{\mathcal{B}}$, therefore

$$v_{\mathcal{B}}^T M^T T_{\mathcal{C}} M w_{\mathcal{B}} = v_{\mathcal{B}}^T T_{\mathcal{B}} w_{\mathcal{B}}$$

for any v, w, and hence

$$M^T T_{\mathcal{C}} M = T_{\mathcal{B}} \Rightarrow T_{\mathcal{C}} = (M^T)^{-1} T_{\mathcal{B}} M^{-1}$$

4. From 1, $\alpha_{\mathcal{B}} = \alpha_{\mathcal{C}} M$, so

$$\alpha_{\mathcal{C}} T_{\mathcal{C}} \beta_{\mathcal{C}}^T = T(\alpha, \beta) = \alpha_{\mathcal{B}} T_{\mathcal{B}} \beta_{\mathcal{B}}^T = \alpha_{\mathcal{C}} M T_{\mathcal{B}} M^T \beta_{\mathcal{C}}^T$$

for any α, β, and similarly

$$T_{\mathcal{C}} = M T_{\mathcal{B}} M^T$$

5. Similar to above,

$$\alpha_{\mathcal{C}} T_{\mathcal{C}} v_{\mathcal{C}} = T(\alpha, v) = \alpha_{\mathcal{B}} T_{\mathcal{B}} v_{\mathcal{B}} = \alpha_{\mathcal{C}} M T_{\mathcal{B}} M^{-1} v_{\mathcal{C}}$$

So

$$T_{\mathcal{C}} = M T_{\mathcal{B}} M^{-1}$$

Solution 1.10.3.

Gradient:

$$\nabla f = 2xi + 2yj + 2zk$$

New function:

$$f_{\text{new}}(x+y, y+z, z) = x^2 + y^2 + z^2$$

Let $u = x + y, v = y + z$, then

$$f_{\text{new}}(u,v,z) = (u - v + z)^2 + (v - z)^2 + z^2$$
$$= u^2 + v^2 + z^2 - 2uv - 2vz + 2uz + v^2 + z^2 - 2vz + z^2$$
$$= u^2 + 2v^2 + 3z^2 - 2uv - 4vz + 2zu$$
$$f_{\text{new}}(x,y,z) = x^2 + 2y^2 + 3z^2 - 2xy - 4yz + 2zx$$

Gradient of new function:

$$\nabla f_{\text{new}} = (2x - 2y + 2z)i + (4y - 2x - 4z)j + (6z - 4y + 2x)k$$

$$(M^{-1})^T = ((I+N)^{-1})^T = (I - N + N^2)^T = \begin{bmatrix} 1 & & \\ -1 & 1 & \\ 1 & -1 & 1 \end{bmatrix}$$

Verification:

$$\text{R.H.S.} = (M^{-1})^T(\nabla f(x,y,z) = 2xi + (-2x + 2y)j + (2x - 2y + 2z)k$$

$$\text{L.H.S.} = \nabla f(x+y, y+z, z) = (2(x+y) - 2(y+z) + 2z)i + (4(y+z) - 2(x+y) - 4z)j + (6z - 4(y+z) + 2(x+y))k$$
$$= 2xi + (2y - 2x)j + (2z - 2y + 2x)k$$

So L.H.S = R.H.S. Note: sgn of a permutation is $(-1)^{\{\text{number of inversions of the permutation}\}}$.

Solution 1.11.1.

1.
$$e^1 \otimes e^2(e_1 \otimes e_2) = e^1(e_1)e^2(e_2) = 1$$

2.
$$\mathrm{Alt}(e^1 \otimes e^2)(e_1 \otimes e_2) = \frac{1}{2}(e^1 \otimes e^2 - e^2 \otimes e^1)(e_1 \otimes e_2) = \frac{1}{2}$$

3.
$$e^1 \otimes e^2(\mathrm{Alt}(e_1 \otimes e_2)) = e^1 \otimes e^2(\frac{1}{2}(e_1 \otimes e_2 - e_2 \otimes e_1)) = \frac{1}{2}$$

4.
$$\mathrm{Alt}(e^1 \otimes e^2)\mathrm{Alt}(e_1 \otimes e_2) = \frac{1}{4}(e^1 \otimes e^2 - e^2 \otimes e^1)(e_1 \otimes e_2 - e_2 \otimes e_1) = \frac{1}{2}$$

5.
$$e^1 \wedge e^2(\mathrm{Alt}(e_1 \otimes e_2)) = 2\mathrm{Alt}(e^1 \otimes e^2)\mathrm{Alt}(e_1 \otimes e_2) = 1$$

6.
$$e^1 \wedge e^2(e_1 \otimes e_2) = 2\mathrm{Alt}(e^1 \otimes e^2)(e_1 \otimes e_2) = 1$$

7.
$$e^1 \otimes e^2(e_1 \wedge e_2) = 2e^1 \otimes e^2(\mathrm{Alt}(e_1 \otimes e_2)) = 1$$

8.
$$e^1 \wedge e^2(e_1 \wedge e_2) = 4\mathrm{Alt}(e^1 \otimes e^2)\mathrm{Alt}(e_1 \otimes e_2) = 2$$

9.
$$\langle e_1 \otimes e_2, e_1 \otimes e_3 \rangle = \langle e_1, e_1 \rangle \langle e_2, e_3 \rangle = 0$$
$$\langle u_1 \wedge v_1, u_2 \wedge v_2 \rangle = \langle u_1 \otimes v_1 - v_1 \otimes u_1, u_2 \otimes v_2 - v_2 \otimes u_2 \rangle$$
$$= \langle u_1 \otimes v_1, u_2 \otimes v_2 \rangle - \langle u_1 \otimes v_1, v_2 \otimes u_2 \rangle - \langle v_1 \otimes u_1, u_2 \otimes v_2 \rangle + \langle v_1 \otimes u_1, v_2 \otimes u_2 \rangle$$
$$= \langle u_1, u_2 \rangle \langle v_1, v_2 \rangle - \langle u_1, v_2 \rangle \langle v_1, u_2 \rangle - \langle v_1, u_2 \rangle \langle u_1, v_2 \rangle + \langle v_1, v_2 \rangle \langle u_1, u_2 \rangle$$
$$= 2(\langle u_1, u_2 \rangle \langle v_1, v_2 \rangle - \langle u_1, v_2 \rangle \langle v_1, u_2 \rangle)$$
$$\langle e_1 \wedge e_2, e_1 \wedge e_3 \rangle = 2(\langle e_1, e_1 \rangle \langle e_2, e_3 \rangle - \langle e_1, e_3 \rangle \langle e_1, e_2 \rangle) = 0$$

10.
$$\langle e_1 \wedge e_2 \wedge e_3, e_1 \wedge e_2 \wedge e_3 \rangle$$
$$= \sum_{\sigma, \tau \in S_3} \mathrm{sgn}(\sigma)\mathrm{sgn}(\tau) \langle \sigma(e_{123}), \tau(e_{123}) \rangle$$
$$= \sum_{\sigma, \tau \in S_3} \mathrm{sgn}(\sigma)\mathrm{sgn}(\tau) \delta_{il}\delta_{jm}\delta_{ln} e^i_{\sigma(1)} e^j_{\sigma(2)} e^k_{\sigma(3)} e^l_{\tau(1)} e^m_{\tau(2)} e^n_{\tau(3)}$$

Any term in this sum is 1 iff $\sigma = \tau$, and 0 otherwise. So,
$$= \sum_{\sigma, \tau \in S_3} \mathrm{sgn}(\sigma)\mathrm{sgn}(\tau) \langle \sigma(e_{123}), \tau(e_{123}) \rangle$$
$$= \sum_{\sigma \in S_3} \delta_{il}\delta_{jm}\delta_{ln} e^i_{\sigma(1)} e^j_{\sigma(2)} e^k_{\sigma(3)} e^l_{\sigma(1)} e^m_{\sigma(2)} e^n_{\sigma(3)} = 6$$
$$\Rightarrow \|e_1 \wedge e_2 \wedge e_3\| = \sqrt{6}$$
$$\|\mathrm{Alt}(e_1 \otimes e_2 \otimes e_3)\| = \|\frac{1}{3!} e_1 \wedge e_2 \wedge e_3\| = \frac{\sqrt{6}}{6}$$
$$\|\mathrm{Alt}(e_1 \otimes e_2) \wedge e_3\| = \frac{1}{2!}\|e_1 \wedge e_2 \wedge e_3\| = \frac{\sqrt{6}}{2}$$

Solution 1.11.2.

1. In $\bigwedge_3 R^3$, $\forall u, v, w \in R^3$
$$\det(u, v, w) = \sum_{\sigma \in S_3} \mathrm{sgn}(\sigma) u^{\sigma(1)} v^{\sigma(2)} w^{\sigma(3)}$$
By definition of Levi-Civita notation,
$$e_{ijk} = \begin{cases} 0, & (i-j)(j-k)(k-i) = 0 \\ \mathrm{sgn}(\sigma), & \text{otherwise} \end{cases}$$
where for the latter case, in particular, $\sigma(1) = i, \sigma(2) = j, \sigma(3) = k$. That immediately yields
$$\det(u, v, w) = e_{ijk} u^i v^j w^k = e_{ijk} e^i(u) e^j(v) e^k(w) = (e_{ijk} e^i \otimes e^j \otimes e^k)(u, v, w)$$

2. $\forall u, v \in R^2, \alpha, \beta \in (R^2)^*$,
$$e_{ij} e^{mn} u^i v^j \alpha_m \beta_n = e_{ij} u^i v^j e^{mn} \alpha_m \beta_n = \det\left(\begin{bmatrix} \alpha(u) & \alpha(v) \\ \beta(u) & \beta(v) \end{bmatrix}\right) = \alpha(u)\beta(v) - \alpha(v)\beta(u)$$
$$(\delta_i^m \delta_j^n - \delta_i^n \delta_j^m)(u^i v^j \alpha_m \beta_n) = \delta_i^m u^i \alpha_m \delta_j^n v^j \beta_n - \delta_i^n u^i \beta_n \delta_j^m v^j \alpha_m = \alpha(u)\beta(v) - \beta(u)\alpha(v)$$
$$\Rightarrow \delta_i^m \delta_j^n - \delta_i^n \delta_j^m = e_{ij} e^{mn}$$

3. For a component of $e_{ijk} e^{imn}$ to be nonzero, only $j = m, k = n$ or $j = n, k = m$ is possible.
For the former case, the parities of the two permutations are the same, which yields 1.
For the latter case, the parities of the two permutations are different, which yields -1.
Therefore, $e_{ijk} e^{imn} = \delta_j^m \delta_k^n - \delta_j^n \delta_k^m$ $\square$.

4. Set $m = j$, then $e_{ijk} e^{ijn} = \delta_j^j \delta_k^n - \delta_j^n \delta_k^j = 3\delta_k^n - \delta_k^n = 2\delta_k^n$
Set $n = k$, then $e_{ijk} e^{ijk} = 2\delta_k^k = 6$

Solution 1.11.3.

1. $\tau^2 = \frac{1}{2}(e_{1122} - e_{1221} - e_{2112} + e_{2211})$,
$$\langle a \otimes b \otimes c \otimes d, e \otimes f \otimes g \otimes h \rangle = \langle a \otimes b, e \otimes f \rangle \langle c \otimes d, g \otimes h \rangle = \langle a, e \rangle \langle b, f \rangle \langle c, g \rangle \langle d, h \rangle$$
So expansion of the inner product by enumerating all combinations (through distributing properties) ignores non-identical terms.
$$\langle \tau^2, \tau^2 \rangle = \frac{1}{4}\left(\langle e_{1122}, e_{1122} \rangle + \langle e_{1221}, e_{1221} \rangle + \langle e_{2112}, e_{2112} \rangle + \langle e_{2211}, e_{2211} \rangle\right) = 1 \Rightarrow \|\tau^2\| = 1$$
Suppose $\tau = u \otimes v$. Then W.L.O.G. let $u = u^1 e_1 + u^2 e_2, v = v^1 e_1 + v^2 e_2$.
$$\tau = u \otimes v = (u^1 e_1 + u^2 e_2) \otimes (v^1 e_1 + v^2 e_2) = u^1 v^1 e_1 \otimes e_1 + u^1 v^2 e_1 \otimes e_2 + u^2 v^1 e_2 \otimes e_1 + u^2 v^2 e_2 \otimes e_2$$
Hence, $u^1 v^1 = u^2 v^2 = 0, u^1 v^2 = \frac{1}{\sqrt{2}}, u^2 v^1 = -\frac{1}{\sqrt{2}}$, and $0 = u^1 v^1 u^2 v^2 = -\frac{1}{2}$, which is impossible. $\square$

2. Trivially its matrix representation is $\begin{bmatrix} 1 & \\ & -1 \end{bmatrix}$, so $M_A = e_1^T \otimes e_1^T - e_2^T \otimes e_2^T$.

3. Its matrix representation is $\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & \\ & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}^{-1} = \begin{bmatrix} & 1 \\ 1 & \end{bmatrix}$,
$N_A = e_1^T \otimes e_2^T + e_2^T \otimes e_1^T$.

4. Matrix multiplication is associative, and is distributive over addition, so the matrix representation of the sum of the two bilinear map is just the sum of the matrix:
For M_B, its representation is
$$\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

That matrix has determinant -1 and trace 0, so ±1 are the eigenvalues.
For N_B, its representation is

$$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ -1 & -1 \end{bmatrix}$$

That matrix has determinant -1 and trace 0, so ±1 are the eigenvalues.

5.
$$M_A \otimes M_B + M_A \otimes N_B + N_A \otimes M_B - N_A \otimes N_B$$
$$= \frac{1}{\sqrt{2}}(M_A \otimes M_A + M_A \otimes N_A + M_A \otimes M_A - M_A \otimes N_A + N_A \otimes M_A + N_A \otimes N_A - N_A \otimes M_A + N_A \otimes N_A)$$
$$= \sqrt{2}(M_A \otimes M_A + N_A \otimes N_A) = \sqrt{2}((e^{11} - e^{22}) \otimes (e^{11} - e^{22}) + (e^{12} + e^{21}) \otimes (e^{12} + e^{21}))$$
$$= \sqrt{2}(e^{1111} - e^{1122} - e^{2211} + e^{2222} + e^{1212} + e^{1221} + e^{2112} + e^{2121}),$$
$$(M_A \otimes M_B + M_A \otimes N_B + N_A \otimes M_B - N_A \otimes N_B)(\tau^2)$$
$$= \frac{\sqrt{2}}{2}(-e^{1122}(e_{1122}) + e^{1221}(-e_{1221}) + e^{2112}(-e_{2112}) - e^{2211}(e_{2211})) = -2\sqrt{2}$$

6. Let $v = \cos\theta e_1 + \sin\theta e_2$, $w = \cos\phi e_1 + \sin\phi e_2$.
$$v \otimes v = \cos^2\theta e_{11} + \cos\theta\sin\theta(e_{12} + e_{21}) + \sin^2\theta e_{22}$$
$$v \otimes v \otimes w \otimes w = \cos^2\theta\cos^2\phi e_{1111} + \cos^2\theta\sin^2\phi e_{1122}$$
$$+ \cos\theta\sin\theta\cos\phi\sin\phi(e_{1212} + e_{1221} + e_{2112} + e_{2121}) + \sin^2\theta\cos^2\phi e_{2211} + \sin^2\theta\sin^2\phi e_{2222} + \cdots$$

The 8 other terms are neglected as they are irrelevant in calculation below.

$$(M_A \otimes M_B + M_A \otimes N_B + N_A \otimes M_B - N_A \otimes N_B)(v \otimes v \otimes w \otimes w)$$
$$= \sqrt{2}(\cos^2\theta\cos^2\phi - \cos^2\theta\sin^2\phi + 4\cos\theta\sin\theta\cos\phi\sin\phi - \sin^2\theta\cos^2\phi + \sin^2\theta\sin^2\phi)$$
$$= \sqrt{2}((\cos^2\theta - \sin^2\theta)(\cos^2\phi - \sin^2\phi) + \sin(2\theta)\sin(2\phi))$$
$$= \sqrt{2}(\cos(2\theta)\cos(2\phi) + \sin(2\theta)\sin(2\phi)) = \sqrt{2}\cos(2\theta - 2\phi),$$

It is trivial that $-\sqrt{2} \le \sqrt{2}\cos(2\theta - 2\phi) \le \sqrt{2}$. ■