Problem 1.

1. Since $A^2\mathbf{v} - A\mathbf{v} = A\mathbf{v} - \mathbf{v}$, we have $(A^2 - 2A + I)\mathbf{v} = (A - I)^2\mathbf{v} = \mathbf{0}$. This implies that $\mathbf{v}$ is a generalized eigenvector of A with eigenvalue $\lambda = 1$, because $\mathbf{v} \in N_2(A - \lambda I) \subseteq N_\infty(A - \lambda I)$.

2. A counter example is $A = \begin{bmatrix} \dfrac{1+\sqrt{5}}{2} & \\ & \dfrac{1-\sqrt{5}}{2} \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

   Since $A^2 = A + I$, we have $A^2\mathbf{x} = A\mathbf{x} + \mathbf{x}$ for all vectors $\mathbf{x}$, and thus each term in the sequence is the sum of the previous two. However, note that the x-axis and y-aixs are the two eigenspaces of A with distinct eigenvalues, therefore $\mathbf{v}$, in neither of which, is not a generalized eigenvector of A.

3. Suppose $a_i x^i$ is the last (in descending order) non-zero term of $p(x)$. If $p(x)$ is a generalized eigenvector of M, then $(x-1)^k p(x) = 0$ for some $k \in \mathbb{N}^*$. This implies that the last term of $(x-1)^k p(x)$ which is $a_i x^i$ is 0. Contradiction.

4. We claim that $\dim(\mathrm{Ker}(A^{k+1}) - \mathrm{Ker}(A^k)) \le \dim(\mathrm{Ker}(A^k) - \mathrm{Ker}(A^{k-1}))$ for any $k \in \mathbb{N}^*$.

   Proof: Each $\mathbf{v}$ in $\mathrm{Ker}(A^{k+1}) - \mathrm{Ker}(A^k)$ gives a vector $A\mathbf{v}$ in $\mathrm{Ker}(A^k) - \mathrm{Ker}(A^{k-1})$, and since $\mathbf{v} \notin \mathrm{Ker}(A^k)$, $A\mathbf{v} \ne \mathbf{0}$, hence there are as many linear independent $A\mathbf{v}$'s as linear independent $\mathbf{v}$'s. Therefore $\dim(\mathrm{Ker}(A^k) - \mathrm{Ker}(A^{k-1}))$ is no less than $\dim(\mathrm{Ker}(A^{k+1}) - \mathrm{Ker}(A^k))$. $\square$

   Accordingly, we have $\dim \mathrm{Ker}(A^k) \le k \dim \mathrm{Ker}(A)$. If $A^5 = O$, then $\dim \mathrm{Ker}(A^5) = n$, and therefore $\dim \mathrm{Ker}(A) \ge \frac{n}{5}$.

5. We claim that for any $\mathbf{v} \in \mathrm{Ran}(A)$, there exists a unique $\mathbf{w} \in \mathrm{Ran}(A^\mathrm{T})$ s.t. $\mathbf{v} = A\mathbf{w}$. This implies that A is invertible between $\mathrm{Ran}(A^\mathrm{T})$ and $\mathrm{Ran}(A)$.

   Proof: If $\mathbf{v} \in \mathrm{Ran}(A)$, then there does exist a vector $\mathbf{w}' \in \mathbb{C}^n$ s.t. $\mathbf{v} = A\mathbf{w}'$. Since $\mathbb{C}^n = \mathrm{Ran}(A^\mathrm{T}) \oplus \mathrm{Ker}(A)$, we have the respective decomposition $\mathbf{w}' = \mathbf{w} + (\mathbf{w}' - \mathbf{w})$, and thus $\mathbf{v} = A\mathbf{w}$. On the other hand $\mathbf{w}$ is unique, because if $\widetilde{\mathbf{w}}$ does the same thing, then $\widetilde{\mathbf{w}} - \mathbf{w} \in \mathrm{Ran}(A^\mathrm{T})$, and $A(\widetilde{\mathbf{w}} - \mathbf{w}) = \mathbf{v} - \mathbf{v} = \mathbf{0}$ i.e. $\widetilde{\mathbf{w}} - \mathbf{w} \in \mathrm{Ker}(A)$, producing $\widetilde{\mathbf{w}} - \mathbf{w} = \mathbf{0}$, so they are the same. $\square$

   Similarly we conclude that A^T is invertible between $\mathrm{Ran}(A)$ and $\mathrm{Ran}(A^\mathrm{T})$. Hence AA^T is invertible on $\mathrm{Ran}(A)$. Let B be the inverse of AA^T on $\mathrm{Ran}(A)$ and zero otherwise (*Moore–Penrose inverse*), then $AA^\mathrm{T}B = BAA^\mathrm{T} = P$ where P is the projection to $\mathrm{Ran}(A)$.

   Since $AA^\mathrm{T}AAA^\mathrm{T} = O$, for any $\mathbf{v} \in \mathbb{C}^n$,

   $$\begin{aligned} \mathbf{0} &= B(AA^\mathrm{T}AAA^\mathrm{T})BA\mathbf{v} \\ &= (BAA^\mathrm{T})A(AA^\mathrm{T}B)A\mathbf{v} \\ &= PAPA\mathbf{v} \\ &= A^2\mathbf{v}, \end{aligned}$$

   i.e. $A^2 = O$. According to the conclusion above, $2\dim \mathrm{Ker}(A) \ge \dim \mathrm{Ker}(A^2) = n$, i.e. $\dim \mathrm{Ker}(A) \ge \frac{n}{2}$.

Problem 2.

1. When A, B are 2×2 matrices,

   $$A \otimes B = \begin{bmatrix} a_{11}b_{11} & a_{11}b_{12} & a_{12}b_{11} & a_{12}b_{12} \\ a_{11}b_{21} & a_{11}b_{22} & a_{12}b_{21} & a_{12}b_{22} \\ a_{21}b_{11} & a_{21}b_{12} & a_{22}b_{11} & a_{22}b_{12} \\ a_{21}b_{21} & a_{21}b_{22} & a_{22}b_{21} & a_{22}b_{22} \end{bmatrix},$$

   and

   $$B \otimes A = \begin{bmatrix} a_{11}b_{11} & a_{12}b_{11} & a_{11}b_{12} & a_{12}b_{12} \\ a_{21}b_{11} & a_{22}b_{11} & a_{21}b_{12} & a_{22}b_{12} \\ a_{11}b_{21} & a_{12}b_{21} & a_{11}b_{22} & a_{12}b_{22} \\ a_{21}b_{21} & a_{22}b_{21} & a_{21}b_{22} & a_{22}b_{22} \end{bmatrix}.$$

Note that $B \otimes A$ is $A \otimes B$ permuting both the second and the third row and column, hence

$$P = P^{-1} = P_{23} = \begin{bmatrix} 1 & & & \\ & & 1 & \\ & 1 & & \\ & & & 1 \end{bmatrix}.$$

2. Since

$$A \otimes I = \begin{bmatrix} a_{11} & & a_{12} & \\ & a_{11} & & a_{12} \\ a_{21} & & a_{22} & \\ & a_{21} & & a_{22} \end{bmatrix}$$

can be decomposed to two independent submaps $A_1 : \mathbf{v} \mapsto A\mathbf{v}$ where $\mathbf{v} \in \operatorname{Span}\{\mathbf{e}_1, \mathbf{e}_3\}$, and $A_2 : \mathbf{v} \mapsto A\mathbf{v}$ where $\mathbf{v} \in \operatorname{Span}\{\mathbf{e}_2, \mathbf{e}_4\}$, hence $\mathrm{e}^{A\otimes I}$ is the direct sum of e^{A_1} and e^{A_2}, which is $\mathrm{e}^A \otimes I$.

Similarly, $I \otimes B$ can be decomposed to submaps B_1 and B_2 on $\operatorname{Span}\{\mathbf{e}_1, \mathbf{e}_2\}$ and $\operatorname{Span}\{\mathbf{e}_3, \mathbf{e}_4\}$ respectively, and therefore $\mathrm{e}^{I\otimes B} = I \otimes \mathrm{e}^B$.

3. Since

$$(A \otimes I)(I \otimes B) = (AI) \otimes (IB) = A \otimes B,$$

and

$$(I \otimes B)(A \otimes I) = (IA) \otimes (BI) = A \otimes B,$$

$A \otimes I$ and $I \otimes B$ commutes. Therefore

$$\mathrm{e}^A \otimes \mathrm{e}^B = (\mathrm{e}^A I) \otimes (I\mathrm{e}^B) = (\mathrm{e}^A \otimes I)(I \otimes \mathrm{e}^B) = \mathrm{e}^{A\otimes I}\mathrm{e}^{I\otimes B} = \mathrm{e}^{A\otimes I + I\otimes B} = \mathrm{e}^{A\oplus B}.$$

4.

$$\begin{aligned} \operatorname{tr}(A \oplus B) &= \operatorname{tr} \begin{bmatrix} a_{11} + b_{11} & b_{12} & a_{12} & \\ b_{21} & a_{11} + b_{22} & & a_{12} \\ a_{21} & & a_{22} + b_{11} & b_{12} \\ & a_{21} & b_{21} & a_{22} + b_{22} \end{bmatrix} \\ &= (a_{11} + a_{22})(b_{11} + b_{22}) \\ &= \operatorname{tr}(A) + \operatorname{tr}(B). \end{aligned}$$

5. WLOG suppose A is Jordanized. For each Jordan block of A with eigenvalue $\lambda \neq 0$, there exists a $\mu \in \mathbb{C}$ s.t. $\mathrm{e}^\mu = \lambda$. Consider the matrix X by replacing λ with μ for each Jordan block of A, then we have e^X agrees with A on every Jordan block, and thus $\mathrm{e}^X = A$.

6. If both A and B are invertible then $A = \mathrm{e}^X$ and $B = \mathrm{e}^Y$ for some X, Y. Hence $A \otimes B = \mathrm{e}^X \otimes \mathrm{e}^Y = \mathrm{e}^{X\oplus Y}$. Since $\det \mathrm{e}^X = \mathrm{e}^{\operatorname{tr}(X)}$,

$$\det(A \otimes B) = \mathrm{e}^{\operatorname{tr}(X\oplus Y)} = \mathrm{e}^{\operatorname{tr}(X)+\operatorname{tr}(Y)} = \mathrm{e}^{\operatorname{tr}(X)}\mathrm{e}^{\operatorname{tr}(Y)} = \det(A)\det(B).$$

This remains correct when A and/or B is not invertible as a result of continuity.

Problem 3.

1. An example is

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

whose Jordan normal form is

$$J = B^{-1}AB = \begin{bmatrix} \mathrm{i} & 1 & & \\ & \mathrm{i} & & \\ & & -\mathrm{i} & 1 \\ & & & -\mathrm{i} \end{bmatrix},$$

under the basis

$$B = \begin{bmatrix} 1 & 0 & \mathrm{i} & 0 \\ \mathrm{i} & 0 & 1 & 0 \\ 0 & 1 & 0 & \mathrm{i} \\ 0 & \mathrm{i} & 0 & 1 \end{bmatrix}.$$

2. An example is

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

where we have

$$AB = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \neq BA = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

They are both in their JNF and they differ.

3. An example is

$$A = \begin{bmatrix} & \mathrm{i} \\ -\mathrm{i} & \end{bmatrix},$$

whose Jordan normal form is

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

under the basis

$$B = \begin{bmatrix} \mathrm{i} & 1 \\ 1 & \mathrm{i} \end{bmatrix}.$$

4. Suppose λ is an eigenvalue of A, then $\mathrm{e}^{\lambda} = \lambda + 1$ i.e. $\lambda = 0$. Hence all eigenvalues of A are 0 so A is nilpotent, $A^4 = O$. Expand the analytic function e^A at 0 and we have $\sum_{n=0}^{\infty} \frac{1}{n!}A^n = A + I$, i.e.

$$O = \sum_{n=2}^{\infty} \frac{1}{n!}A^n = \frac{1}{2}A^2 + \frac{1}{6}A^3.$$

Since A is Jordan normal, this implies that $A^2 = O$, hence dimensions of Jordan blocks of A do not exceed 2. Therefore, all possible A's are

$$\begin{bmatrix} 0 & & & \\ & 0 & & \\ & & 0 & \\ & & & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & & \\ & 0 & & \\ & & 0 & 1 \\ & & & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & & \\ & 0 & & \\ & & 0 & \\ & & & 0 \end{bmatrix}, \begin{bmatrix} 0 & & & \\ & 0 & 1 & \\ & & 0 & \\ & & & 0 \end{bmatrix}, \begin{bmatrix} 0 & & & \\ & 0 & & \\ & & 0 & 1 \\ & & & 0 \end{bmatrix},$$

though the last three are in fact equivalent.

Problem 4.

1. Suppose $p(x) = (k_0, k_1, k_2, k_3)$ under the basis $(e_0, e_1, e_2, e_3) = (1, x, x^2, x^3)$. The multiple of $q(x)$ will always have its first term the same as $xp(x)$, which is k_3x^4. Therefore, $(Mp)(x) = xp(x) - k_3q(x) = (-dk_3, k_0 - ck_3, k_1 - bk_3, k_2 - ak_3)$, i.e.

$$A = \begin{bmatrix} & & & -d \\ 1 & & & -c \\ & 1 & & -b \\ & & 1 & -a \end{bmatrix}.$$

2. According to the compatibility with addition and multiplication of modular arithmetic, if $a \equiv b \pmod{x}$ then $P(a) \equiv P(b) \pmod{x}$ for any polynomial $P(x)$. Therefore since

$$(Mp)(x) \equiv xp(x) \pmod{q(x)},$$

we have

$$(P(M)p)(x) \equiv P(x)p(x) \pmod{q(x)}$$

for any polynomial $P(x)$. And since $(P(M)p)(x) \in V$, it is the remainder itself. Let $P(x) = \mu(x)$ the minimal polynomial of A, then $\mu(M) = 0$, and we have

$$0 = (\mu(M)p)(x) = \mu(x)p(x) \bmod q(x)$$

for any $p(x) \in V$. This happens if and only if $\mu(x)$ is a multiple of $q(x)$, as well as a monic polynomial and of the lowest degree among all. Therefore $\mu(x) = q(x)$, and since its degree is $4 = n$, it is the characteristic polynomial of A as well.

3. Note that $A^ie_0 = e_i$, thus $A^ie_j = A^{i+j}e_0 = A^je_i$ $(i, j = 0, 1, 2, 3)$. Since B commutes with A and therefore any power of A, the i-th (counting from 0) column of B is

$$Be_i = BA^ie_0 = A^iBe_0 = A^i\sum_{j=0}^{3} B_{j0}e_j = \sum_{j=0}^{3} B_{j0}A^ie_j = \sum_{j=0}^{3} B_{j0}A^je_i,$$

the same as that of $\sum_{j=0}^{3} B_{j0}A^j$. Hence $B = \sum_{j=0}^{3} B_{j0}A^j$, which is a polynomial function of A.

4. Because the minimal polynomial of A coincides with its characteristic polynomial $(\mu(x) = \chi(x) = q(x))$, the Jordan normal form of A has no Jordan blocks sharing a common eigenvalue. Therefore we get the JNF straight from $\chi(x)$, which results in $\operatorname{diag}(0, 1, 2, 3)$.

5. Similarly, the JNF is

$$\begin{bmatrix} 0 & 1 & & \\ & 0 & & \\ & & 1 & \\ & & & 2 \end{bmatrix}.$$

Problem 5.

1. For any orthogonal matrix A, $AA^{\mathrm{T}} = I$, hence $-(A - I)(A^{\mathrm{T}} - I) = -AA^{\mathrm{T}} + A + A^{\mathrm{T}} - I = A + A^{\mathrm{T}} - 2I$.

2. Since $f(0) = f(0)^{\mathrm{T}} = I$, according to the conclusion above,

$$\begin{aligned} \lim_{t\to 0} \frac{-(f(t) - I)(f(t)^{\mathrm{T}} - I)}{t} &= \lim_{t\to 0} \frac{f(t) + f(t)^{\mathrm{T}} - 2I}{t} \\ &= \lim_{t\to 0} \frac{f(t) - f(0)}{t} + \lim_{t\to 0} \frac{f(t)^{\mathrm{T}} - f(0)^{\mathrm{T}}}{t} \\ &= f'(0) + f'(0)^{\mathrm{T}}. \end{aligned}$$

3. Note that

$$\begin{aligned}
\lim_{t\to 0}\frac{-(f(t)-I)(f(t)^{\mathrm{T}}-I)}{t} &= -\lim_{t\to 0}\left(\frac{f(t)-f(0)}{t}\cdot\frac{f(t)^{\mathrm{T}}-f(0)^{\mathrm{T}}}{t}\cdot t\right)\\
&= -f'(0)f'(0)^{\mathrm{T}}\lim_{t\to 0} t\\
&= O,
\end{aligned}$$

hence compare with the equation in the last subproblem and we have $f'(0)+f'(0)^{\mathrm{T}}=O$, i.e. $f'(0)$ is skew-symmetric.

4. Since $f(t)\in \mathrm{SO}_n$,

$$I = f(t)f(t)^{\mathrm{T}}.$$

Take the derivative of t to both sides of the equation and

$$\begin{aligned}
O &= f'(t)f(t)^{\mathrm{T}} + f(t)f'(t)^{\mathrm{T}}\\
&= f'(t)f(t)^{\mathrm{T}} + (f'(t)f(t)^{\mathrm{T}})^{\mathrm{T}}\\
&= f'(t)f(t)^{-1} + (f'(t)f(t)^{-1})^{\mathrm{T}},
\end{aligned}$$

i.e. $f'(t)f(t)^{-1}$ is skew-symmetric.

5. Since $f'(t)f(t)^{-1}=A$, $f'(t)=Af(t)$. Note that this is a first-order homogeneous differential equation and e^{At} is a solution to it, so the general solution is $\mathrm{e}^{At}B$ where B is a constant depending on initial conditions. When $t=0$, $f(0)=B\in\mathrm{SO}_n$.