1. Let $\beta = \{e_1, e_2\}$ or $\{e_1, e_2, e_3\}$ be the standard basis of $\mathbb{R}^2$ or $\mathbb{R}^3$.

   (a) We have
$$P(e_i) = \begin{cases} 0 & \text{if } i = 1 \\ e_i & \text{if } i = 2, 3. \end{cases}$$
   Thus
$$[P]_\beta = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

   (b) We have
$$Q(e_1) = \begin{pmatrix} \cos\frac{\pi}{3} \\ \sin\frac{\pi}{3} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{pmatrix}, \quad Q(e_2) = \begin{pmatrix} \cos\frac{5\pi}{6} \\ \sin\frac{5\pi}{6} \end{pmatrix} = \begin{pmatrix} -\frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{pmatrix}.$$
   Hence
$$[Q]_\beta = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}.$$

   (c) Let $v_1 = \binom{-3}{1}, v_2 = \binom{1}{3}$ and $\gamma = \{v_1, v_2\}$, which is an ordered basis of $\mathbb{R}^2$. We have
$$[R]_\gamma = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad [\mathbf{1}]_\gamma^\beta = (v_1, v_2) = \begin{pmatrix} -3 & 1 \\ 1 & 3 \end{pmatrix}$$
   and
$$[\mathbf{1}]_\beta^\gamma = \left([\mathbf{1}]_\gamma^\beta\right)^{-1} = \frac{-1}{10}\begin{pmatrix} 3 & -1 \\ -1 & -3 \end{pmatrix}.$$
   Therefore
$$[R]_\beta = [\mathbf{1}]_\gamma^\beta [R]_\gamma [\mathbf{1}]_\beta^\gamma = \frac{1}{5}\begin{pmatrix} 4 & -3 \\ -3 & -4 \end{pmatrix}.$$

2. (a) Assume that S and T are linearly dependent. Since S and T are non-zero elements in a vector space, we have $S = \alpha T$ for some non-zero $\alpha \in F$. Thus $S(v) = \alpha T(v) = T(\alpha v)$, which implies that $\mathrm{Im}(S) \subset \mathrm{Im}(T)$ and $\ker(S) \supset \ker(T)$.
   On the other hand, we have $T = \beta S$ where $\beta = \alpha^{-1}$. Thus we have the other directions $\mathrm{Im}(S) \supset \mathrm{Im}(T)$ and $\ker(S) \subset \ker(T)$.

   (b) Suppose that S and T are linearly dependent. We have
$$\begin{aligned} \{0\} &= \mathrm{Im}(S) \cap \mathrm{Im}(T) \quad \text{(by assumption)} \\ &= \mathrm{Im}(S) \quad \text{(by part (a))}, \end{aligned}$$
   i.e., S is the zero map, which is then a contradiction.

3. We provide 3 (or 9 by various combinations) methods here. (Notice that given A of rank r, the decomposition $A = BC$ in the statement is not unique.)

   Method I (via linear transformations). ($\Rightarrow$) The idea is that the desired equation $A = BC$ means that the map L_A can be decomposed as $L_B \circ L_C$ with $L_B : F^r \to F^m$ injective and $L_C : F^n \to F^r$ surjective. This can be realized as follows.

   Consider the linear map
$$L_A : F^n \to F^m$$
   defined by left multiplication by A on column vectors of F^n. By assumption L_A is of rank r. Thus there exists linearly independent subset $\alpha = \{v_1, \cdots, v_r\}$ of F^n such that $\{L_A(v_1), \cdots, L_A(v_r)\}$ forms a basis of $\mathrm{Im}(L_A)$.

Extend α to a basis $\beta = \{v_1, \cdots, v_n\}$ of F^n. Define the linear maps

$$F^n \xrightarrow{T} F^r \xrightarrow{S} F^m$$

by requiring

$$T(v_i) = \begin{cases} e_i & \text{if } 1 \leq i \leq r \\ 0 & \text{if } r < i \leq n \end{cases} \quad \text{and} \quad S(e_j) = L_A(v_j) \quad (1 \leq j \leq r)$$

where $\{e_i\}$ is the standard basis of F^r. Then by construction,

$$L_A = S \circ T. \tag{1}$$

Notice that $\mathrm{rk}(S) = r$ since S is injective and $\mathrm{rk}(T) = r$ since T is surjective. Now let B and C be the matrix representations of S and T, respectively, using the standard bases. Then $\mathrm{rk}(B) = \mathrm{rk}(S) = r$ and $\mathrm{rk}(C) = \mathrm{rk}(T) = r$. Furthermore $A = BC$ by (1).

($\Leftarrow$) We have

$$F^n \xrightarrow{L_C} F^r \xrightarrow{L_B} F^m$$

with $\mathrm{rk}(L_B) = \mathrm{rk}(B) = r$ and $\mathrm{rk}(L_C) = \mathrm{rk}(C) = r$ by assumption. Thus L_B is injective and L_C is surjective. We have

$$\begin{aligned} L_A(F^n) = L_{BC}(F^n) &= L_B(L_C(F^n)) \\ &= L_B(F^r) \quad (\text{since } L_C \text{ is surjective}). \end{aligned}$$

Since L_B is injective, $\dim L_B(F^r) = \dim F^r = r$. Thus $\mathrm{rk}(A) = \mathrm{rk}(L_A) = \dim L_A(F^n) = r$.

Method II (via row/column operations). ($\Rightarrow$) Since A is of rank r, there exist invertible matrices $P \in M_m(F)$ and $Q \in M_n(F)$ such that

$$PAQ = \begin{pmatrix} I_r & O_1 \\ O_2 & O_3 \end{pmatrix} \tag{2}$$

where $I_r \in M_r(F)$ is the identity matrix and $O_1 \in M_{r,n-r}(F), O_2 \in M_{m-r,r}(F), O_3 \in M_{m-r,n-r}(F)$ are zero matrices. Notice that the right hand side of (2) is equal to the product $B'C'$ where

$$B' = \begin{pmatrix} I_r \\ O_2 \end{pmatrix} \in M_{m \times r}(F) \quad \text{and} \quad C' = (I_r \; O_1) \in M_{r \times n}(F).$$

Moreover B' and C' are both of rank r since they are of RREF and have r pivots.

Now let

$$B = P^{-1}B' \quad \text{and} \quad C = C'Q^{-1}.$$

Then $\mathrm{rk}(B) = \mathrm{rk}(B') = r$ since P^{-1} is invertible and similarly $\mathrm{rk}(C) = r$. We have $A = BC$ by (2).

($\Leftarrow$) Since $B \in M_{m \times r}(F)$ is of rank r, the RREF of B must be

$$\begin{pmatrix} I_r \\ O_2 \end{pmatrix}$$

where $O_2 \in M_{m-r,r}(F)$ is the zero matrix. Thus there exists an invertible $P \in M_m(F)$ such that

$$PB = \begin{pmatrix} I_r \\ O_2 \end{pmatrix}.$$

Similarly there exists an invertible $Q \in M_n(F)$ such that

$$CQ = (I_r \; O_1)$$

where $O_1 \in M_{r,n-r}(F)$ is the zero matrix (which can be seen by considering C^t and doing row operations or performing column operations on C directly). Then we have

$$PAQ = (PB)(CQ) = \begin{pmatrix} I_r & O_1 \\ O_2 & O_3 \end{pmatrix}$$

and

$$\begin{aligned} \mathrm{rk}(A) &= \mathrm{rk}(PAQ) \quad \text{(since } P, Q \text{ invertible)} \\ &= \mathrm{rk} \begin{pmatrix} I_r & O_1 \\ O_2 & O_3 \end{pmatrix} \\ &= r \quad \text{(since } r \text{ pivots)}. \end{aligned}$$

Method III (via matrix operations). ($\Rightarrow$) Write $A = (c_1, \cdots, c_n)$ where $c_i \in F^m$ are columns of A. Since A is of rank r, there exist r linearly independent columns of A, say $c_{k_1}, \cdots, c_{k_r}$ such that all columns of A are linear combinations of them. Thus there exist $\alpha_{ij} \in F$ for $1 \le i \le r, 1 \le j \le n$ such that

$$c_j = \sum_{i=1}^{r} \alpha_{ij} c_{k_i}. \tag{3}$$

Now if we let $B = (c_{k_1}, \cdots, c_{k_r}) \in M_{m \times r}(F)$ and $C = (\alpha_{ij}) \in M_{r \times n}(F)$, then $A = BC$ by (3).

The matrix B has rank r since the columns of B are linearly independent. The matrix C also has rank r since the $k_1, \cdots, k_r$-th columns of C form the identity matrix I_r by (3).

($\Leftarrow$) Since $A = BC$, we have $\mathrm{rk}(A) \le \mathrm{rk}(B) = r$.

On the other hand, there are r independent rows $d_{i_1}, \cdots, d_{i_r}$ of B and r independent columns $c_{j_1}, \cdots, c_{j_r}$ of C. The product

$$M = \begin{pmatrix} d_{i_1} \\ \vdots \\ d_{i_r} \end{pmatrix} (c_{j_1} \cdots c_{j_r})$$

of these rows and columns forms a minor of the product $A = BC$. We have

$$\det M = \det \begin{pmatrix} d_{i_1} \\ \vdots \\ d_{i_r} \end{pmatrix} \cdot \det(c_{j_1} \cdots c_{j_r}),$$

which is non-zero since the two terms in the right hand side are the determinants of two square matrices of full rank. Thus $\mathrm{rk}(M) = r$ and consequently $\mathrm{rk}(A) \ge r$.

4. Since $\mathrm{rk}(T) = r$, there are r linearly independent vectors $v_1, \cdots, v_r$ of V such that $\{T(v_1), \cdots, T(v_r)\}$ forms a basis of $\mathrm{Im}(T)$.

   On one hand, $\nu(T) = n - r$ by dimension formula. Take a basis $\{v_{r+1}, \cdots, v_n\}$ of $\ker(T)$. Then $\beta = \{v_1, \cdots, v_n\}$ is an ordered basis of V (cf. the proof of the dimension formula).

   On the other hand, extend $\{T(v_1), \cdots, T(v_r)\}$ to an ordered basis

   $$\gamma = \{T(v_1), \cdots, T(v_r), w_{r+1}, \cdots, w_m\}.$$

   Then $[T]_\beta^\gamma$ gives the result.

5. (a) Let $A = (a_{ij}), 1 \le i \le m, 1 \le j \le n$ and $B = (b_{pq}), 1 \le p \le n, 1 \le q \le m$. Write $AB = (c_{ij})$ and $BA = (d_{ij})$. Then

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} \quad \text{and} \quad d_{ij} = \sum_{l=1}^{m} b_{il} a_{lj}.$$

We have

$$\begin{aligned} \operatorname{tr}(AB) = \sum_{i=1}^{m} c_{ii} &= \sum_{i=1}^{m} \sum_{k=1}^{n} a_{ik} b_{ki} \\ &= \sum_{k=1}^{n} \sum_{i=1}^{m} b_{ki} a_{ik} \\ &= \sum_{k=1}^{n} d_{kk} = \operatorname{tr}(BA). \end{aligned}$$

(b) Let γ be another ordered basis of V and $Q = [\mathbf{1}]_\beta^\gamma$. Then $[T]_\gamma = Q[T]_\beta Q^{-1}$ and hence

$$\begin{aligned} \operatorname{tr}([T]_\gamma) &= \operatorname{tr}\left((Q[T]_\beta)Q^{-1}\right) \\ &= \operatorname{tr}\left(Q^{-1}(Q[T]_\beta)\right) \quad \text{(by part (a))} \\ &= \operatorname{tr}([T]_\beta). \end{aligned}$$

6. (a) For any $(T+U)(v) \in \operatorname{Im}(T+U)$, we have $(T+U)(v) = T(v)+U(v) \in \operatorname{Im}(T)+\operatorname{Im}(U)$. Thus $\operatorname{Im}(T+U) \subset \operatorname{Im}(T) + \operatorname{Im}(U)$.

(b) We have

$$\begin{aligned} \operatorname{rk}(T+U) &= \dim \operatorname{Im}(T+U) \\ &\le \dim\left(\operatorname{Im}(T) + \operatorname{Im}(U)\right) \quad \text{(by part (a))} \\ &= \dim \operatorname{Im}(T) + \dim \operatorname{Im}(U) - \dim\left(\operatorname{Im}(T) \cap \operatorname{Im}(U)\right) \\ &\le \dim \operatorname{Im}(T) + \dim \operatorname{Im}(U) \\ &= \operatorname{rk}(T) + \operatorname{rk}(U). \end{aligned}$$

(c) For example $U = 0$.

(d) For example T is any non-zero (hence $\operatorname{rk}(T) > 0$) linear map and $U = -T$.

7. The goal of this question is to give basic properties of projections on a vector space. Geometrically speaking, given a *projection* T means to have a decomposition $V = W \oplus W'$ such that T on W is the identity and is zero on W' (as in the textbook). Algebraically a linear $T \in \mathcal{L}(V)$ is a *projection* if

$$T^2 = T. \tag{4}$$

The latter equation actually means a decomposition of the identity map

$$\mathbf{1}_V = T + (\mathbf{1}_V - T)$$

such that the two terms in the right hand side both satisfy (4) (meaning: square = itself). The last equation is part (a) and the equivalence of the two definitions is part (e).

(a) We have $S^2 = (1-T)^2 = 1 - 2T + T^2 \overset{(*)}{=} 1 - T = S$ where $(*)$ follows from $T^2 = T$.

(b) If $v \in \ker(S)$, then $0 = S(v) = v - T(v)$, i.e., $v = T(v) \in \mathrm{Im}(T)$. Therefore $\ker(S) \subset \mathrm{Im}(T)$. On the other hand, for any $T(v) \in \mathrm{Im}(T)$, we have $ST(v) = T(v) - T^2(v) = 0$ since $T^2 = T$ again. Thus $\mathrm{Im}(T) \subset \ker(S)$.

Notice that $S^2 = S$ and $T = 1 - S$. Thus one can replace T by S and S by T in the above argument and obtain the equality $\mathrm{Im}(S) = \ker(T)$.

(c,d) First notice that

$$\mathbf{1}_V = T + S,$$

which immediately shows that $V = \mathrm{Im}(T) + \mathrm{Im}(S)$. On the other hand, we have

$$\begin{aligned}
\dim V &= \dim\big(\mathrm{Im}(T) + \mathrm{Im}(S)\big) \\
&= \dim\big(\mathrm{Im}(T) + \ker(T)\big) \quad \text{(by part (b))} \\
&= \dim \mathrm{Im}(T) + \dim\ker(T) - \dim\big(\mathrm{Im}(T) \cap \ker(T)\big) \\
&= \dim V - \dim\big(\mathrm{Im}(T) \cap \ker(T)\big) \quad \text{(dimension formula)}.
\end{aligned}$$

Thus $\dim\big(\mathrm{Im}(T) \cap \ker(T)\big) = 0$, i.e., $\mathrm{Im}(T) \cap \ker(T) = \{0\}$. We then obtain

$$\begin{aligned}
V &= \mathrm{Im}(T) \oplus \mathrm{Im}(S) \\
&= \ker(S) \oplus \ker(T) \quad \text{(by part (b))}.
\end{aligned}$$

(e) Parts (b) and (c) imply $V = \mathrm{Im}(T) \oplus \ker(T)$. Also if $v \in \mathrm{Im}(T) = \ker(S)$, then $0 = S(v) = v - T(v)$, i.e., $T(v) = v$ for any $v \in \mathrm{Im}(T)$. Thus if we take a basis $\{v_1, \cdots, v_r\}$ of $\mathrm{Im}(T)$ and a basis $\{v_{r+1}, \cdots, v_n\}$ of $\ker(T)$, then $\beta = \{v_1, \cdots, v_n\}$ forms an ordered basis of V and

$$T(v_i) = \begin{cases} v_i & \text{if } 1 \le i \le r \\ 0 & \text{if } r < i \le n. \end{cases}$$

Hence $[T]_\beta$ satisfies the required property.

Remark. Another way to prove part (e) is as follows. The equation $T(T-1) = T^2 - T = 0$ shows that the *minimal polynomial* of T is a factor of $P(x) = x(x-1)$. Since $P(x)$ has different roots 0 and 1, T is diagonalizable with diagonal entries belonging to $\{0, 1\}$. With this result, parts (a) - (d) then follow easily (exercise!).

8. This determinant is the characteristic polynomial (with variable λ) of a so-called *rational canonical form*. We will learn this in next semester. The answer is

$$\det(A) = \lambda^n + a_{n-1}\lambda^{n-1} + a_{n-2}\lambda^{n-2} + \cdots + a_1\lambda + a_0. \tag{5}$$

To obtain this, one can expand it along the last column. Then each minor is just a lower triangular matrix. (Leave to you as an exercise since I am tired of typing.)

Or one can expand it along the first row and use induction. Thus

$$\det(A) = (\spadesuit) + (\heartsuit)$$

where

$$\begin{aligned}
(\spadesuit) &= \lambda \cdot \det \begin{pmatrix} \lambda & 0 & \dots & 0 & a_1 \\ -1 & \lambda & \dots & 0 & a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & \lambda & a_{n-2} \\ 0 & 0 & \dots & -1 & \lambda + a_{n-1} \end{pmatrix} \\
&= \lambda(\lambda^{n-1} + a_{n-1}\lambda^{n-2} + \cdots + a_2\lambda + a_1)
\end{aligned}$$

by induction, while

$$(\heartsuit) = (-1)^{n+1} a_0 \cdot \det \begin{pmatrix} -1 & \lambda & 0 & \dots & 0 \\ 0 & -1 & \lambda & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda \\ 0 & 0 & 0 & \dots & -1 \end{pmatrix} = (-1)^{n+1} a_0 (-1)^{n-1} = a_0.$$

The equation (5) then follows.