Exercises.

Solution to Question 1. We may choose a orthonormal basis $(\mathbf{e}_1, \dots, \mathbf{e}_n)$ for V. And let

$$\mathbf{w} = \sum_{i=1}^{n} \phi(\mathbf{e}_i)\mathbf{e}_i.$$

Then for all $i = 1, \dots, n$,

$$\phi(\mathbf{e}_i) = \langle \mathbf{e}_i, \mathbf{w} \rangle.$$

By the linearity of ϕ, we know that for any $\mathbf{v} \in V$,

$$\phi(\mathbf{v}) = \langle \mathbf{v}, \mathbf{w} \rangle.$$

This $\mathbf{w}$ is unique. If there exists $\mathbf{w}'$ such that

$$\langle \mathbf{v}, \mathbf{w}' \rangle = \phi(\mathbf{v}) = \langle \mathbf{v}, \mathbf{w} \rangle,$$

then for all $\mathbf{v} \in V$,

$$\langle \mathbf{v}, \mathbf{w}' - \mathbf{w} \rangle = 0.$$

So $\mathbf{w}' = \mathbf{w}$.

Solution to Question 2.

(a) Let $(\mathbf{e}_1, \dots, \mathbf{e}_m)$ be a basis for W and define for all $i = 1, \dots, m$,

$$\phi_i(v) := \langle T(v), \mathbf{e}_i \rangle.$$

We may check that ϕ_i is linear. Thus, by problem 1, there exists a $\mathbf{u}_i \in W$ such that for all $\mathbf{v} \in V$

$$\phi_i(\mathbf{v}) = \langle \mathbf{v}, \mathbf{u}_i \rangle.$$

By uniqueness of such a vector $\mathbf{u}_i$, we may define a linear ransformation $T* : W \to V$ by

$$T^*(\mathbf{e}_i) = \mathbf{u}_i.$$

Then for all $\mathbf{v} \in V$ and $\mathbf{w} \in W$,

$$\langle T(\mathbf{v}), \mathbf{w} \rangle = \langle \mathbf{v}, T^*(\mathbf{w}) \rangle.$$

(b) Take $\mathbf{v} = (x, y, z) \in \mathbb{R}^3$ and $\mathbf{w} = (a, b) \in \mathbb{R}^2$. Then

$$\langle T(\mathbf{v}), \mathbf{w} \rangle = (y + 2z)a + 3xb = 3bx + ay + 2az.$$

Therefore,

$$T^*(a, b) = (3b, a, 2a).$$

Solution to Question 3.

(a) For any $\mathbf{u}_1, \mathbf{u}_2 \in V$, we know that $\mathbf{u}_1 = \mathbf{u}_2$ if and only if $\langle \mathbf{v}, \mathbf{u}_1 \rangle = \langle \mathbf{v}, \mathbf{u}_2 \rangle$ for all $\mathbf{v} \in V$. Therefore, to prove T^* is linear, we only need to check for all $\mathbf{v} \in V$ and $\mathbf{w}_1, \mathbf{w}_2 \in W$,

$$\langle \mathbf{v}, T^*(a\mathbf{w}_1 + b\mathbf{w}_2) \rangle = \langle \mathbf{v}, aT^*(\mathbf{w}_1) + bT^*(\mathbf{w}_2) \rangle.$$

This is true, because the LHS is

$$\begin{aligned} \langle \mathbf{v}, T^*(a\mathbf{w}_1 + b\mathbf{w}_2) \rangle &= \langle T(\mathbf{v}), (a\mathbf{w}_1 + b\mathbf{w}_2) \rangle \\ &= \langle T(\mathbf{v}), a\mathbf{w}_1 \rangle + \langle T(\mathbf{v}), b\mathbf{w}_2 \rangle \\ &= a\langle T(\mathbf{v}), \mathbf{w}_1 \rangle + b\langle T(\mathbf{v}), \mathbf{w}_2 \rangle, \end{aligned}$$

and the RHS is

$$\begin{aligned} \langle \mathbf{v}, aT^*(\mathbf{w}_1) + bT^*(\mathbf{w}_2) \rangle &= a\langle \mathbf{v}, T^*(\mathbf{w}_1) \rangle + b\langle \mathbf{v}, T^*(\mathbf{w}_2) \rangle \\ &= a\langle T(\mathbf{v}), \mathbf{w}_1 \rangle + b\langle T(\mathbf{v}), \mathbf{w}_2 \rangle \end{aligned}$$

(b) This is because for any $\mathbf{v} \in V$ and $\mathbf{w} \in W$,

$$\langle (T^*)^*(\mathbf{v}), \mathbf{w} \rangle = \langle \mathbf{v}, T^*(\mathbf{w}) \rangle = \langle T(\mathbf{v}), \mathbf{w} \rangle.$$

This means for all $\mathbf{v} \in V$,

$$(T^*)^*(\mathbf{v}) = T(\mathbf{v}).$$

(c) $\mathbf{w} \in \ker(T^*)$ if and only if $T^*(\mathbf{w}) = 0$. This is equivalent to say, for all $\mathbf{v} \in V$,

$$\langle T(\mathbf{v}), \mathbf{w} \rangle = 0,$$

i.e., $\mathbf{w} \in \operatorname{im}(T)^\perp$. Therefore, $\mathbf{w} \in \ker(T^*) \iff \mathbf{w} \in \operatorname{im}(T)^\perp$.

(d) $\mathbf{u} \in \operatorname{im}(T^*)$ if and only if $T^*(\mathbf{w}) = \mathbf{u}$ for some $\mathbf{w} \in W$. For all $\mathbf{v} \in \ker T$,

$$\langle \mathbf{v}, \mathbf{u} \rangle = \langle \mathbf{v}, T^*(\mathbf{w}) \rangle = \langle T(\mathbf{v}), \mathbf{w} \rangle = 0.$$

So $\operatorname{im}(T^*) \subseteq \ker(T)^\perp$.
Note that $\dim \operatorname{im}(T^*) = \operatorname{rank}(T^*)$ and $\dim \ker(T)^\perp = \operatorname{rank}(T)$. By part (c),

$$\dim \ker(T^*) = \dim \operatorname{im}(T)^\perp.$$

Because

$$\dim V - \operatorname{rank}(T^*) = \dim \ker(T^*) = \dim \operatorname{im}(T)^\perp = \dim V - \operatorname{rank}(T),$$

so

$$\dim \operatorname{im}(T^*) = \operatorname{rank}(T^*) = \operatorname{rank}(T) = \dim \ker(T)^\perp.$$

Hence,

$$\operatorname{im}(T^*) = \ker(T)^\perp.$$

(e) Because for all $\mathbf{v} \in V$ and $\mathbf{w} \in W$,

$$\langle \mathsf{T}(\mathbf{v}), \mathbf{w} \rangle = [\mathsf{T}(\mathbf{v})]_{\mathcal{B}}^{\mathsf{T}} [\mathbf{w}]_{\mathcal{B}} = ([\mathsf{T}]_{\mathcal{A} \leftarrow \mathcal{B}} [\mathbf{v}]_{\mathcal{A}})^{\mathsf{T}} [\mathbf{w}]_{\mathcal{B}} = [\mathbf{v}]_{\mathcal{A}}^{\mathsf{T}} [\mathsf{T}]_{\mathcal{A} \leftarrow \mathcal{B}}^{\mathsf{T}} [\mathbf{w}]_{\mathcal{B}},$$

and

$$\langle \mathsf{T}(\mathbf{v}), \mathbf{w} \rangle = \langle \mathbf{v}, \mathsf{T}^*(\mathbf{w}) \rangle = [\mathbf{v}]_{\mathcal{A}}^{\mathsf{T}} [\mathsf{T}^*]_{\mathcal{A} \leftarrow \mathcal{B}} [\mathbf{w}]_{\mathcal{B}},$$

so

$$[\mathsf{T}]_{\mathcal{A} \leftarrow \mathcal{B}}^{\mathsf{T}} = [\mathsf{T}^*]_{\mathcal{A} \leftarrow \mathcal{B}}.$$

Solution to Question 4.

(a)
- R_θ is a linear transformation on $\mathbb{R}^2$:
  Assume
  $$\begin{aligned}\mathbf{v} &= a(\cos\phi_1, \sin\phi_1),\\ \mathbf{w} &= b(\cos\phi_2, \sin\phi_2).\end{aligned}$$
  Then
  $$\begin{aligned}R_\theta(\mathbf{v}) &= a(\cos(\phi_1+\theta), \sin(\phi_1+\theta)),\\ R_\theta(\mathbf{w}) &= b(\cos(\phi_2+\theta), \sin(\phi_2+\theta)).\end{aligned}$$
  Let $\mathbf{u} = R_\theta(\mathbf{v}) + R_\theta(\mathbf{w})$. Then
  $$\|\mathbf{u}\| = \sqrt{a^2 + 2ab\cos(\phi_1 - \phi_2) + b^2} = \|\mathbf{v}+\mathbf{w}\|.$$
  So we may assume
  $$\mathbf{v}+\mathbf{w} = \|\mathbf{u}\|(\cos\psi, \sin\psi),$$
  where
  $$\begin{aligned}\|\mathbf{u}\|\cos\psi &= a\cos\phi_1 + b\cos\phi_2,\\ \|\mathbf{u}\|\sin\psi &= a\sin\phi_1 + b\sin\phi_2.\end{aligned}$$
  Then
  $$\begin{aligned}R_\theta(\mathbf{v}+\mathbf{w}) &= \|\mathbf{u}\|(\cos(\psi+\theta), \sin(\psi+\theta))\\ &= \|\mathbf{u}\|(\cos\psi\cos\theta - \sin\psi\sin\theta, \cos\psi\sin\theta + \cos\theta\sin\psi)\\ &= ((a\cos\phi_1 + b\cos\phi_2)\cos\theta - (a\sin\phi_1 + b\sin\phi_2)\sin\theta,\\ &\quad (a\cos\phi_1 + b\cos\phi_2)\sin\theta + \cos\theta(a\sin\phi_1 + b\sin\phi_2))\\ &= R_\theta(\mathbf{v}) + R_\theta(\mathbf{w}).\end{aligned}$$
  Also note that the scalar multiple and rotation by θ commutes, so $R_\theta(c\mathbf{v}) = cR_\theta(\mathbf{v})$.
- Let $(\mathbf{e}_1, \mathbf{e}_2)$ be the standard basis. Note that
  $$\begin{aligned}R_\theta(\mathbf{e}_1) &= \cos\theta\mathbf{e}_1 + \sin\theta\mathbf{e}_2,\\ R_\theta(\mathbf{e}_2) &= -\sin\theta\mathbf{e}_1 + \cos\theta\mathbf{e}_2.\end{aligned}$$
  So
  $$R_\theta = \begin{bmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{bmatrix}.$$

(b) $R_\alpha R_\beta = R_{\alpha+\beta}$, because rotation by $\alpha+\beta$ is equivalent to rotation first by α then by β. So

$$\begin{aligned}\begin{bmatrix}\cos(\alpha+\beta) & \sin(\alpha+\beta)\\ -\sin(\alpha+\beta) & \cos(\alpha+\beta)\end{bmatrix} &= \begin{bmatrix}\cos(\alpha) & \sin(\alpha)\\ -\sin(\alpha) & \cos(\alpha)\end{bmatrix}\begin{bmatrix}\cos(\beta) & \sin(\beta)\\ -\sin(\beta) & \cos(\beta)\end{bmatrix}\\ &= \begin{bmatrix}\cos(\alpha)\cos(\beta)-\sin(\alpha)\sin(\beta) & \cos(\alpha)\sin(\beta)+\cos(\beta)\sin(\alpha)\\ -\cos(\alpha)\sin(\beta)-\cos(\beta)\sin(\alpha) & \cos(\alpha)\cos(\beta)-\sin(\alpha)\sin(\beta)\end{bmatrix}\end{aligned}$$

We get trig angle addition formulas

$$\begin{aligned}\cos(\alpha+\beta) &= \cos(\alpha)\cos(\beta)-\sin(\alpha)\sin(\beta),\\ \sin(\alpha+\beta) &= \cos(\alpha)\sin(\beta)+\cos(\beta)\sin(\alpha).\end{aligned}$$

(c) Three Givens rotations in $\mathbb{R}^3$ are

$$\begin{aligned}G(1,2,\theta) &= \begin{bmatrix}\cos\theta & \sin\theta & 0\\ -\sin\theta & \cos\theta & 0\\ 0 & 0 & 1\end{bmatrix}\\ G(1,3,\theta) &= \begin{bmatrix}\cos\theta & 0 & \sin\theta\\ 0 & 1 & 0\\ -\sin\theta & 0 & \cos\theta\end{bmatrix}\\ G(2,3,\theta) &= \begin{bmatrix}1 & 0 & 0\\ 0 & \cos\theta & \sin\theta\\ 0 & -\sin\theta & \cos\theta\end{bmatrix}\end{aligned}$$

(d) Let $G = G(i,j,\theta)$. We need to check that $G^TG = I$.
Let $\langle \mathbf{v}, \mathbf{w}\rangle = \mathbf{v}\cdot\mathbf{w}$. Taking the standard basis, $\langle \mathbf{v}, \mathbf{w}\rangle = \mathbf{v}^T\mathbf{w}$. If we rotation $\mathbf{v}$ and $\mathbf{w}$ by the same angle θ, then their inner product should stay the same. Therefore, for all $\mathbf{v}, \mathbf{w} \in V$,

$$(G(\mathbf{v}))^T(G(\mathbf{w})) = \langle G(\mathbf{v}), G(\mathbf{w})\rangle = \langle \mathbf{v}, \mathbf{w}\rangle = \mathbf{v}^T\mathbf{w}.$$

So $G^TG = I$.

Solution to Question 5.

(a) Check that $H^T = H$

$$H^T = (I - 2\mathbf{u}\mathbf{u}^T)^T = I - 2\mathbf{u}\mathbf{u}^T = H,$$

and that $H^T H = I$

$$\begin{aligned} H^T H &= H^2 \\ &= (I - 2\mathbf{u}\mathbf{u}^T)(I - 2\mathbf{u}\mathbf{u}^T) \\ &= I^2 - 4\mathbf{u}\mathbf{u}^T + 4\mathbf{u}(\mathbf{u}^T\mathbf{u})\mathbf{u}^T \\ &= I^2 - 4\mathbf{u}\mathbf{u}^T + 4\mathbf{u}\mathbf{u}^T \\ &= I. \end{aligned}$$

(b)

$$\begin{aligned} H(\mathbf{v}) &= (I - 2\mathbf{u}\mathbf{u}^T)\mathbf{v} \\ &= \mathbf{v} - 2\mathbf{u}\mathbf{u}^T\mathbf{v} \\ &= \mathbf{v} - 2\mathbf{u}\|\mathbf{v}\| \\ &= \mathbf{v} - 2\mathbf{v} \\ &= -\mathbf{v}. \end{aligned}$$

If $\mathbf{w} \cdot \mathbf{v} = 0$, then

$$\begin{aligned} H(\mathbf{w}) &= (I - 2\mathbf{u}\mathbf{u}^T)\mathbf{w} \\ &= \mathbf{w} - 2\mathbf{u}(\mathbf{u}^T\mathbf{w}) \\ &= \mathbf{w}. \end{aligned}$$

Extended Glossary.

Definition 1. A **permutation matrix** is a square matrix that has exactly one entry of 1 in each row and each column and 0s elsewhere.

Example 1. $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ is a 2×2 permutation matrix.

Non-example 1. $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$ is not a permutation matrix.

Theorem 1. *There are* $n!$ $n \times n$ *permutation matrices.*

Proof. By definition, a $n \times n$ matrix is a permutaion matrix if and only if it has exactly one entry of 1 in each row and column, and 0s elsewhere. We may record the location of 1s by

row	1	2	3	...	n
column	r_1	r_2	r_3	...	r_n

where $(r_1, r_2, \ldots, r_n)$ is a permutation of $(1, 2, \ldots, n)$. Given a permutation matrix, we obtain such a $(r_1, r_2, \ldots, r_n)$. Conversely, each of such a permutation gives a permutation matrix. Hence, there is a one-to-one correspondence between permutations of $(1, 2, \ldots, n)$ and $n \times n$ permutation matrices. So there are $n!$ of them. $\square$