Math 4310 Name: Rami Pellumbi (rp534)

Homework 10

Please print out these pages. I encourage you to work with your classmates on this homework. Please list your collaborators on this cover sheet. (Your grade will not be affected.) Even if you work in a group, you should write up your solutions yourself! You should include all computational details, and proofs should be carefully written with full details. As always, please write neatly and legibly.

Please follow the instructions for the "extended glossary" on separate paper (LaTeX it if you can!) Hand in your final draft, including full explanations and write your glossary in complete, mathematically and grammatically correct sentences. Your answers will be assessed for style and accuracy.

Please **staple** *this cover sheet, your exercise solutions, and your glossary together, in that order, and hand in your homework in class.*

GRADES

Exercises ______________________ / 50

Extended Glossary

Component	Correct?	Well-written?
Definition	/6	/6
Example	/4	/4
Non-example	/4	/4
Theorem	/5	/5
Proof	/6	/6
Total	/25	/25

Exercises.

1. Suppose that V is a finite dimensional vector space over $\mathbb{R}$, equipped with an inner product. Suppose that $\phi \in V^*$ (dual vector space), (i.e. $\phi : V \to \mathbb{R}$ is a linear transformation). Show that there is a unique vector $\vec{w} \in V$ such that for every $\vec{v} \in V$,

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

Proof.
Existence.
First we show that there exists a vector $\vec{w} \in V$ as desired.

Let $\dim(V) = n < \infty$ and suppose that $(\vec{e_1}, ..., \vec{e_n})$ is an orthonormal basis for V. Then we have

$$\begin{aligned}\phi(\vec{v}) &= \phi(\langle \vec{v_1}, \vec{e_1} \rangle \vec{e_1} + ... + \langle \vec{v_n}, \vec{e_n} \rangle \vec{e_n}) \\ &= \langle \vec{v_1}, \vec{e_1} \rangle \phi(\vec{e_1}) + ... + \langle \vec{v_n}, \vec{e_n} \rangle \phi(\vec{e_n}) \\ &= \langle \vec{v_1}, \overline{\phi(\vec{e_1})} \vec{e_1} \rangle + ... + \langle \vec{v_n}, \overline{\phi(\vec{e_n})} \vec{e_n} \rangle \\ &= \langle \vec{v}, \overline{\phi(\vec{e_1})} \vec{e_1} + ... + \overline{\phi(\vec{e_n})} \vec{e_n} \rangle\end{aligned}$$

The first equality comes from representing $\vec{v}$ as a combination of its orthonormal basis, the second comes from distributing ϕ, and the last two comes from properties of inner products.

Setting

$$\vec{w} = \overline{\phi(\vec{e_1})} \vec{e_1} + ... + \overline{\phi(\vec{e_n})} \vec{e_n}$$

completes the proof.

Uniqueness.

For every $\vec{v} \in V$ take $\vec{w_1}, \vec{w_2} \in V$ such that

$$\phi(\vec{v}) = \langle \vec{v}, \vec{w_1} \rangle = \langle \vec{v}, \vec{w_2} \rangle.$$

Then we have

$$\begin{aligned} 0 &= \langle \vec{v}, \vec{w_1} \rangle - \langle \vec{v}, \vec{w_2} \rangle \\ 0 &= \langle \vec{v}, \vec{w_1} - \vec{w_2} \rangle \end{aligned}$$

Letting $\vec{v} = \vec{w_1} - \vec{w_2}$, then $\langle \vec{v}, \vec{w_1} - \vec{w_2} \rangle = 0$ if and only if $\vec{v} = \vec{w_1} - \vec{w_2} = 0$. It follows that $\vec{w_1} = \vec{w_2}$, completing the proof. □

2. Suppose that $T : V \to W$ is a linear transformation where V and W are finite dimensional vector spaces over $\mathbb{R}$, each equipped with an inner product.

   (a) Prove that there is a function $T^* : W \to V$ which satisfied for all $\vec{v} \in V$ and $\vec{w} \in W$, that

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

   Note that the first inner product in this formula refers to the inner product in W, and the second inner product refers to the inner product in V. T^* is called the ***adjoint*** *of T.*

   Proof. We use the fact proved in (1) above.
   For $T \in \mathcal{L}(V, W)$, and $\vec{w} \in W$, consider the linear functional

   $$\phi : V \to \mathbb{R}$$

   defined by

   $$\vec{v} \mapsto \langle T\vec{v}, \vec{w} \rangle.$$

   This ϕ as defined depends on both T and $\vec{w}$.
   By part (1), there exists a unique vector in $\vec{v}^* \in V$ such that this linear functional is given by taking the inner product with $\vec{v}^*$. If we let this $\vec{v}^* = T^*\vec{w}$, it follows that $T^*\vec{w}$ is the unique vector in V such that

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

   □

   (b) Define

   $$T : \mathbb{R}^3 \to \mathbb{R}^2$$

   by

   $$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} y + 2z \\ 3x \end{pmatrix}.$$

   Define a formula for

   $$T^* : \mathbb{R}^2 \to \mathbb{R}^3.$$

Proof. Fix a point $(x', y') \in \mathbb{R}^2$. Then for every $(x, y, z) \in \mathbb{R}^3$ we have

$$\begin{aligned}\langle (x,y,z), T^*(x',y') \rangle &= \langle T(x,y,z), (x',y') \rangle \\ &= \langle (y+2z, 3x), (x',y') \rangle \\ &= (y+2z)x' + (3x)y' \\ &= \langle (x,y,z), (3y', x', 2x') \rangle\end{aligned}$$

Thus

$$T^*(x', y') = (3y', x', 2x').$$

$\square$

3. Suppose that $T : V \to W$ is a linear transformation where V and W are finite dimensional vector spaces over $\mathbb{R}$, each equipped with an inner product.

   (a) Show that T^* is a linear transformation.

   Proof.

   Fix $\vec{w_1}, \vec{w_2} \in W$ and $\lambda_1, \lambda_2 \in \mathbb{R}$. Then $\forall \vec{v} \in V$,

$$\begin{aligned}\langle \vec{v}, T^*(\lambda_1 \vec{w_1} + \lambda_2 \vec{w_2}) \rangle &= \langle T\vec{v}, \lambda_1 \vec{w_1} + \lambda_2 \vec{w_2} \rangle \\ &= \langle T\vec{v}, \lambda_1 \vec{w_1} \rangle + \langle T\vec{v}, \lambda_2 \vec{w_2} \rangle \\ &= \overline{\lambda_1} \langle T\vec{v}, \vec{w_1} \rangle + \overline{\lambda_2} \langle T\vec{v}, \vec{w_2} \rangle \\ &= \overline{\lambda_1} \langle \vec{v}, T^*(\vec{w_1}) \rangle + \overline{\lambda_2} \langle \vec{v}, T^*(\vec{w_2}) \rangle \\ &= \langle \vec{v}, \lambda_1 T^*(\vec{w_1}) \rangle + \langle \vec{v}, \lambda_2 T^*(\vec{w_2}) \rangle \\ &= \langle \vec{v}, \lambda_1 T^*(\vec{w_1}) + \lambda_2 T^*(\vec{w_2}) \rangle\end{aligned}$$

   which shows that

$$T^*(\lambda_1 \vec{w_1} + \lambda_2 \vec{w_2}) = \lambda_1 T^*(\vec{w_1}) + \lambda_2 T^*(\vec{w_2})$$

   giving T^* is a linear transformation as desired. $\square$

   (b) Show that $(T^*)^* = T$.

   Proof. We use the that the $\overline{\langle u, v \rangle} = \langle v, u \rangle$.
   Suppose that $T \in \mathcal{L}(V, W)$ and that $T^* \in \mathcal{L}(W, V)$ such that

$$\langle T\vec{v}, \vec{w} \rangle = \langle \vec{v}, T^*\vec{w} \rangle.$$

   Taking the conjugate of both sides yields

$$\langle \vec{w}, T\vec{v} \rangle = \langle T^*\vec{w}, \vec{v} \rangle.$$

   By definition,

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

   so that

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

   giving $T = (T^*)^*$ as desired. $\square$

(c) Show that $\ker(T^*) = \mathrm{im}(T)^\perp$.

Proof. Begin by noting

$$\ker(T^*) = \{\vec{w} \in W \mid T\vec{w} = \vec{0}\} \subseteq W$$
$$\mathrm{im}(T)^\perp = \{\vec{w} \in W \mid \langle \vec{w}, \vec{v}\rangle = 0 \; \forall \, \vec{v} \in \mathrm{im}(T)\} \subseteq W$$

so that this equality can make sense.
First, suppose that $\vec{w} \in \ker(T^*)$. Then $T^*\vec{w} = \vec{0}$ and for $\vec{v} \neq \vec{0} \in V$,

$$\langle T\vec{v}, \vec{w}\rangle = \langle \vec{w}, T^*\vec{w}\rangle = \langle \vec{w}, \vec{0}\rangle = 0$$

giving that $w \in \mathrm{im}(T)^\perp$ (since $T\vec{v} \in \mathrm{im}(T)$) and thus

$$\ker(T^*) \subseteq \mathrm{im}(T)^\perp.$$

Next, suppose that $w \in \mathrm{im}(T)^\perp$. Then for $v \neq \vec{0} \in V$

$$\langle T\vec{v}, \vec{w}\rangle = \langle \vec{v}, T^*w\rangle = 0$$
$$\implies T^*\vec{w} = 0$$

by properties of inner products. Thus $w \in \ker(T^*)$ and

$$\mathrm{im}(T)^\perp \subseteq \ker(T^*)$$

and thus

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

□

(d) Show that $\mathrm{im}(T^*) = \ker(T)^\perp$.

Proof. We use the fact $(U^\perp)^\perp = U$.
The equality asking to be proved is the same as

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

Begin by noting

$$\ker(T) = \{\vec{v} \in V \mid T\vec{v} = \vec{0}\} \subseteq V$$
$$\mathrm{im}(T^*)^\perp = \{\vec{v} \in V \mid \langle \vec{v}, \vec{w}\rangle = 0 \; \forall \, \vec{w} \in \mathrm{im}(T^*)\} \subseteq V$$

so that this equality can make sense.
Suppose first that $\vec{v} \in \ker(T)$. Then for $\vec{w} \in W$

$$\langle T\vec{v}, \vec{w}\rangle = \langle \vec{0}, \vec{w}\rangle = \langle \vec{v}, T^*\vec{w}\rangle = 0$$

giving that $\vec{v} \in \mathrm{im}(T^*)^\perp$ since $T^*\vec{w} \in \mathrm{im}(T^*)$. Thus, we have that

$$\ker(T) \subseteq \mathrm{im}(T^*)^\perp.$$

Next suppose that $\vec{v} \in \mathrm{im}(T^*)^{\perp}$. Then for $\vec{w} \neq 0 \in W$,

$$\langle T\vec{v}, \vec{w}\rangle = \langle \vec{v}, T^*\vec{w}\rangle = 0$$

$$\implies T\vec{v} = \vec{0} \to \vec{v} \in \ker(T).$$

Thus,

$$\mathrm{im}(T^*)^{\perp} \subseteq \ker(T)$$

and

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

Taking the orthogonal complement of both sides gives the desired equality

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

□

(e) Suppose that $\mathcal{A}$ is an orthonormal basis for V and $\mathcal{B}$ is an orthonormal basis for W, and that

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

Show that

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

Proof. First we state a few assumptions.
Let $A = [T]_{\mathcal{B}\leftarrow\mathcal{A}}$, $B = [T^*]_{\mathcal{A}\leftarrow\mathcal{B}}$ and

$$\dim(V) = n < \infty$$

$$\dim(W) = m < \infty.$$

Let $\mathcal{A} = (\vec{v}_1, ..., \vec{v}_n)$ be an orthonormal basis for V and $\mathcal{B} = (\vec{w}_1, ..., \vec{w}_n)$ be an orthonormal basis for W.
Finally, suppose that $1 \leq j \leq m$ and $1 \leq k \leq n$.
We know that the j^{th} column of B consists of the scalars needed to write $T^*(w_j)$ as a linear combination of $\vec{v}_1, ..., \vec{v}_n$.
Similarly, the k^{th} column of A consists of the scalars needed to write $T(\vec{v}_k)$ as a linear combination of $\vec{w}_1, ...\vec{w}_m$
We thus express $T^*(\vec{w}_j)$ as

$$T^*(\vec{w}_j) = \sum_{r=1}^{n} B_{r,j}\vec{v}_r,$$

and $T(\vec{v}_k)$ as

$$T(\vec{v}_k) = \sum_{r=1}^{m} A_{r,k}\vec{w}_r.$$

From the definition of the adjoint,

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

Evaluating the left hand side first, we have

$$\langle \vec{v}_k, T^*(\vec{w}_j)\rangle = \langle \vec{v}_k, \sum_{r=1}^{n} B_{r,j}\vec{v}_r\rangle = B_{k,j}$$

since $\mathcal{A}$ is an orthonormal basis.
Similarly, the right hand side becomes,

$$\langle T\vec{v}_k, \vec{w}_j\rangle = \langle \sum_{r=1}^{m} A_{r,k}w_r, \vec{w}_j\rangle = A_{j,k}$$

since $\mathcal{B}$ is an orthonormal basis.
Since $B_{k,j} = A_{j,k}$ for every possible j, k, $B = A^T$ by definition.

□

4. **Rotation Matrices**

(a) Show that in the plane $\mathbb{R}^2$, counterclockwise rotation by angle θ is a linear transformation with matrix

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

Proof. Consider the standard basis $(\vec{e_1}, \vec{e_2})$ of $\mathbb{R}^2$. Given $\theta \in [0, 2\pi]$ and $\lambda_1, \lambda_2 \in \mathbb{R}$,

$$R_\theta\left(\lambda_1\begin{pmatrix}\vec{e_1} & \vec{e_2}\end{pmatrix} + \lambda_2\begin{pmatrix}\vec{e_1} & \vec{e_2}\end{pmatrix}\right)$$

$$= R_\theta\left((\lambda_1+\lambda_2)\begin{pmatrix}\vec{e_1} & \vec{e_2}\end{pmatrix}\right)$$

$$= (\lambda_1+\lambda_2)\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

$$= \lambda_1\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} + \lambda_2\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

$$= \lambda_1 R_\theta\begin{pmatrix}\vec{e_1} & \vec{e_2}\end{pmatrix} + \lambda_2 R_\theta\begin{pmatrix}\vec{e_1} & \vec{e_2}\end{pmatrix}$$

showing R_θ is a linear transformation as desired. □

(b) Show that $R_\alpha R_\beta = R_{\alpha+\beta}$. Argue geometrically. What does this say about trig angle addition formulas?

Proof.

$$R_\alpha R_\beta = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \begin{pmatrix} \cos(\beta) & -\sin(\beta) \\ \sin(\beta) & \cos(\beta) \end{pmatrix}$$

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

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

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

by the trig angle formulas. Thus

$$R_\alpha R_\beta = R_{\alpha+\beta}$$

This makes intuitive geometric sense.
If we rotate a pair of vectors in $\mathbb{R}^2$ counterclockwise by β and then rotate them again counterclockwise by α it would be the same as just rotating them counterclockwise by $\alpha + \beta$.

□

(c) Given $1 \leq i < j \leq n$ and an angle θ, a **Givens** rotation $G(i, j, \theta)$ on $\mathbb{R}^n$ is the linear transformation which fixed e_l for $l \neq i, j$, and in the (x_i, x_j)-plane, is given by rotation angle θ.
Write down the matrices of all three Givens rotations in $\mathbb{R}^3$.
Need to rotate the two planes not being kept constant.
Let us first rotate the $x - y$ plane.

$$G(x, y, \theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Next, the $x - z$ plane

$$G(x, z, \theta) = \begin{pmatrix} \cos(\theta) & 0 & -\sin(\theta) \\ 0 & 1 & 0 \\ \sin(\theta) & 0 & \cos(\theta) \end{pmatrix}$$

Finally, the $y - z$ plane

$$G(y, z, \theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{pmatrix}$$

(d) Show that $G(i, j, \theta)$ is orthogonal.

Proof. Based off the above, it can be determined (and verified via Wikipedia) that the general Givens rotation $G(i, j, \theta)$in $\mathbb{R}^n$ is the $n \times n$ matrix where

$$\begin{cases} G_{k,k} = 1 & k \neq i, j \\ G_{k,k} = \cos(\theta) & k = i, j \\ G_{j,i} = -G_{i,j} = -\sin(\theta) & \\ \text{everything else } 0 & \end{cases}$$

for $i > j \in 1, ..., n$.
From this it is clear that $G(i, j, \theta)$ takes a vector $\vec{v} \in \mathbb{R}^n$ and rotates it by θ in the i, j-plane. (Fun fact, there are $\binom{n}{2}$ Givens rotations on $\mathbb{R}^n$).
To show that $G(i, j, \theta)$ is orthogonal, we multiply by its transpose and show it is the identity.
But this is clear to see from the definition. For diagonal entries not at $G_{i,i}, G_{j,j}$ these will be 1 since all of those diagonal entries of G are 1 with zeros in the rest of the column.
Then

$$G_{i,i} = G_{j,j} = \cos^2(\theta) + \sin^2(\theta) = 1.$$

The rest of the entries will be zero. Thus we have the identity in $\mathbb{R}^n$ and $G(i, j, \theta)$ is orthogonal. □

5. **Reflection Matrices**

Let $V = \mathbb{R}^n$, equipped with the standard inner product. Given a nonzero vector $\vec{v}$ and a unit vector $\vec{u} = \frac{\vec{v}}{\|v\|}$, let $H \in \mathbb{R}^{n \times n}$ be the square matrix defined by

$$H = I - 2\frac{\vec{v}\vec{v}^T}{\vec{v}^T\vec{v}} = I - 2\vec{u}\vec{u}^T.$$

(a) Show that H is a symmetric and orthogonal $n \times n$ matrix.

Proof. The matrix is clearly in $\mathbb{R}^{n \times n}$ since $V = \mathbb{R}^n$ (also given).
Symmetric.

$$H = I - 2\vec{u}\vec{u}^T = I - 2\frac{\|v\|^2}{\|v\|^2} = I - 2$$

which is a diagonal matrix and thus symmetric.
Orthogonal.

$$\begin{aligned} HH^T &= (I - 2\vec{u}\vec{u}^T)(I - 2\vec{u}\vec{u}^T)^T = (I - 2\vec{u}\vec{u}^T)(I^T - 2\vec{u}^T\vec{u}) \\ &= II^T - 2\vec{u}\vec{u}^T - 2\vec{u}^T\vec{u} + 4\vec{u}\vec{u}^T\vec{u}^T\vec{u} = I - 2 - 2 + 4 = I \end{aligned}$$

□

(b) Show that H is a reflection. That is, $H(\vec{v}) = -\vec{v}$, and if $\vec{w} \in V^{\perp}$, then $H(\vec{w}) = \vec{w}$.

$$H(\vec{v}) = (I - 2uu^{T})\vec{v} = I\vec{v} - 2\vec{v} = -\vec{v}$$

$$H(\vec{w}) = (I - 2uu^{T})\vec{w} = I\vec{w} - 2\frac{\vec{v}\vec{v}^{T}\vec{w}}{\|v\|^{2}} = I\vec{w} - 0 = \vec{w}$$

Extended Glossary.

Define the notion of a **permutation matrix**. Give an example, a non-example,and state and prove a theorem about permutation matrices.
In accumulating the knowledge for this glossary I stumbled upon this. It is a very interesting read and some cool math (but much to long to be included as the glossary)! https://webpages.uncc.edu/ghetyei/courses/old/F07.3116/birkhofft.pdf

Definition 1. An $n \times n$ matrix P is a **permutation matrix** if and only if it can be obtained from the $n \times n$ identity matrix I by performing one or more interchanges of the rows and columns of I.

The following proposition states a common application of permutation matrices.

Proposition 1. Each such matrix P represents a permutation of elements and, when used to multiply another $n \times m$ matrix, A, results in permuting the rows (when pre-multiplying, to form PA) or columns (when post-multiplying, to form AP) of the matrix A.

We now state an example and non-example of permutation matrices.

Example 1. An example of a 4×4 permutation matrix is

$$P = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

which is obtained by swapping the 1^{st} and 4^{th} columns of the 4×4 identity matrix (could also be seen as swapping the first and fourth rows)

$$I_{4\times 4} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Per the proposition above, for the matrix

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ a & b & c & d \\ e & f & g & h \end{pmatrix}$$

we see that

$$PA = \begin{pmatrix} e & f & g & h \\ 5 & 6 & 7 & 8 \\ a & b & c & d \\ 1 & 2 & 3 & 4 \end{pmatrix}$$

interchanges the first and fourth rows of A, while

$$AP = \begin{pmatrix} 4 & 2 & 3 & 1 \\ 8 & 6 & 7 & 5 \\ d & b & c & a \\ e & f & g & h \end{pmatrix}$$

interchanges the first and fourth columns of A.

Example 2. An example of a 4×4 matrix which is not a permutation matrix is

$$A = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

since A can not be reproduced from interchanged the rows and columns of $I_{4\times4}$.

What follows is a series of propositions used in the proof of the theorem.

Proposition 2. Consider the identity matrix in $\mathbb{R}^{n\times n}$ and suppose $1 \leq i < j \leq n$. Then interchanging row i with row j in $I_{n\times n}$ is equivalent to interchanging column i with column j.

Proof. Each element of $I_{n\times n}$ is 1 in the diagonal and zero elsewhere.
Interchanging row i with row j results in the matrix where

$$p_{i,i} = 0, p_{j,j} = 0,$$

$$p_{i,j} = 1, p_{j,i} = 1$$

and the rest of the diagonals are one with the other elements zero.
Likewise, interchanging column i with column j results in the matrix where

$$p_{i,i} = 0, p_{j,j} = 0,$$

$$p_{i,j} = 1, p_{j,i} = 1$$

and the rest of the diagonals are one with the other elements zero. □

Proposition 3. The identity matrix $I_{n\times n}$ has $n!$ possible permutations.

Proof. There are n rows possible of permuting, and treating each row as an object, n objects can be permuted in $n!$ ways. It is only $n!$ since interchanging two rows is equivalent to interchanging two columns per proposition 2. □

Proposition 4. A permutation matrix, P, is a full rank matrix.

Proof. The rows are the standard basis of the space of $1 \times n$ vectors, and its columns are the standard basis of the space of $n \times 1$ vectors. Thus the columns of an $n \times n$ permutation matrix, P, are comprised of n linearly independent vectors and so P is full rank. □

Proposition 5. A permutation matrix, P, is an orthogonal matrix.

Proof. Suppose that P is an $n \times n$ permuation matrix.
P is invertible since P by the Invertible Matrix Theorem since P is a full rank matrix. Thus $\exists P^{-1}$ such that

$$PP^{-1} = I$$

Need to show that $P^T = P^{-1}$.
From the definition of P,

$$(PP^T)_{i,j} = \sum_{r=1}^{n} P_{i,r}P^T_{r,j} = \sum_{r=1}^{n} P_{i,r}P_{j,r} \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

$\square$

so that

$$PP^T = I_{n\times n}$$

Theorem 1. *Applying the permutation P finitely many times will result back in the identity. That is,*

$$P^k = I$$

for some $k > 0$.

Proof. By proposition 3, there are only finitely many possible permutations. So the application of P will eventually repeat itself. That is, for some $i > j \geq 0$

$$P^i = P^j$$

By proposition 4, P is an invertible matrix. Taking

$$P^{i-j} = P^iP^{-j} = P^i(P^j)^{-1} = P^j(P^j)^{-1} = I$$

completes the proof. $\square$