Linear Algebra 2

Lecture #17

Permutation matrices. Orthogonal matrices

Irena Penev

March 20, 2024

- This lecture has three parts:

- This lecture has three parts:
  1. Permutation matrices

- This lecture has three parts:
    1. Permutation matrices
    2. Orthogonal matrices

- This lecture has three parts:
  1. Permutation matrices
  2. Orthogonal matrices
  3. Scalar product, coordinate vectors, and matrices of linear functions

1. Permutation matrices

1. Permutation matrices

Definition

A *permutation matrix* is a square matrix that has exactly one 1 in each row and each column, and has 0's everywhere else.

- Examples:

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

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

Definition

A *permutation matrix* is a square matrix that has exactly one 1 in each row and each column, and has 0's everywhere else.

- The 0's and 1's in permutation matrices may belong to any field $\mathbb{F}$ of our choice.

Definition

A *permutation matrix* is a square matrix that has exactly one 1 in each row and each column, and has 0's everywhere else.

- The 0's and 1's in permutation matrices may belong to any field $\mathbb{F}$ of our choice.
    - In our study of permutation matrices, we will never need to add two non-zero numbers, and whenever we multiply two numbers, at least one of the two numbers will be 0 or 1.

Definition

A *permutation matrix* is a square matrix that has exactly one 1 in each row and each column, and has 0's everywhere else.

- The 0's and 1's in permutation matrices may belong to any field $\mathbb{F}$ of our choice.
  - In our study of permutation matrices, we will never need to add two non-zero numbers, and whenever we multiply two numbers, at least one of the two numbers will be 0 or 1.
  - So, it does not matter which particular field we are working in, and therefore, we will not emphasize this.

Definition

A *permutation matrix* is a square matrix that has exactly one 1 in each row and each column, and has 0's everywhere else.

- Obviously, identity matrices are permutation matrices.

Definition

A *permutation matrix* is a square matrix that has exactly one 1 in each row and each column, and has 0's everywhere else.

- Obviously, identity matrices are permutation matrices.
- Moreover, $n \times n$ permutation matrices are precisely the matrices that can be obtained from the identity matrix I_n by reordering (i.e. permuting) rows, or alternatively, by reordering (i.e. permuting) columns.

Definition

A *permutation matrix* is a square matrix that has exactly one 1 in each row and each column, and has 0's everywhere else.

- Obviously, identity matrices are permutation matrices.
- Moreover, $n \times n$ permutation matrices are precisely the matrices that can be obtained from the identity matrix I_n by reordering (i.e. permuting) rows, or alternatively, by reordering (i.e. permuting) columns.
- So, the columns of an $n \times n$ permutation matrix are the standard basis vectors $\mathbf{e}_1, \dots, \mathbf{e}_n$ (appearing in some order in that matrix), whereas the rows are $\mathbf{e}_1^T, \dots, \mathbf{e}_n^T$ (again, appearing in some order in that matrix).

Definition

For a positive integer n and a permutation $\pi \in S_n$, we define the *matrix of the permutation* π, denoted by P_π, to be the $n \times n$ matrix that has 1 in the $\big(i, \pi(i)\big)$-th entry for each each index $i \in \{1, \dots, n\}$, and has 0 in all other entries. In other words, for each index $i \in \{1, \dots, n\}$, the i-th row of the matrix P_π is $\mathbf{e}^T_{\pi(i)}$.

- For example, for the permutation

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 4 & 1 & 6 & 5 & 3 \end{pmatrix},$$

in S_6, we obtain the 6×6 permutation matrix

$$P_\pi = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}.$$

Definition

For a positive integer n and a permutation $\pi \in S_n$, we define the *matrix of the permutation* π, denoted by P_π, to be the $n \times n$ matrix that has 1 in the $\big(i, \pi(i)\big)$-th entry for each each index $i \in \{1, \dots, n\}$, and has 0 in all other entries. In other words, for each index $i \in \{1, \dots, n\}$, the i-th row of the matrix P_π is $\mathbf{e}_{\pi(i)}^T$.

- Obviously, for a positive integer n, the matrix of the identity permutation 1_n in S_n is precisely the identity matrix I_n, i.e. $P_{1_n} = I_n$.

Proposition 2.3.10

Let n be a positive integer, and let $\pi \in S_n$. Then P_π is a permutation matrix.

Proof.

Proposition 2.3.10

Let n be a positive integer, and let $\pi \in S_n$. Then P_π is a permutation matrix.

Proof. Obviously, P_π is an $n \times n$ matrix, all of whose entries are 0's and 1's.

Proposition 2.3.10

Let n be a positive integer, and let $\pi \in S_n$. Then P_π is a permutation matrix.

Proof. Obviously, P_π is an $n \times n$ matrix, all of whose entries are 0's and 1's. Moreover, by the definition of P_π, we have that for each index $i \in \{1, \dots, n\}$, the i-th row of P_π is the row vector $\mathbf{e}_{\pi(i)}^T$.

Proposition 2.3.10

Let n be a positive integer, and let $\pi \in S_n$. Then P_π is a permutation matrix.

Proof. Obviously, P_π is an $n \times n$ matrix, all of whose entries are 0's and 1's. Moreover, by the definition of P_π, we have that for each index $i \in \{1, \dots, n\}$, the i-th row of P_π is the row vector $\mathbf{e}_{\pi(i)}^T$. So, P_π has exactly one 1 in each row.

Proposition 2.3.10

Let n be a positive integer, and let $\pi \in S_n$. Then P_π is a permutation matrix.

Proof. Obviously, P_π is an $n \times n$ matrix, all of whose entries are 0's and 1's. Moreover, by the definition of P_π, we have that for each index $i \in \{1, \dots, n\}$, the i-th row of P_π is the row vector $\mathbf{e}^T_{\pi(i)}$. So, P_π has exactly one 1 in each row. Note that this means that the matrix P_π has exactly n entries that are 1, whereas all the other entries are 0's.

Proposition 2.3.10

Let n be a positive integer, and let $\pi \in S_n$. Then P_π is a permutation matrix.

Proof. Obviously, P_π is an $n \times n$ matrix, all of whose entries are 0's and 1's. Moreover, by the definition of P_π, we have that for each index $i \in \{1, \dots, n\}$, the i-th row of P_π is the row vector $\mathbf{e}_{\pi(i)}^T$. So, P_π has exactly one 1 in each row. Note that this means that the matrix P_π has exactly n entries that are 1, whereas all the other entries are 0's.

It remains to show that the matrix P_π has exactly one 1 in each column.

Proposition 2.3.10

Let n be a positive integer, and let $\pi \in S_n$. Then P_π is a permutation matrix.

Proof. Obviously, P_π is an $n \times n$ matrix, all of whose entries are 0's and 1's. Moreover, by the definition of P_π, we have that for each index $i \in \{1, \dots, n\}$, the i-th row of P_π is the row vector $\mathbf{e}_{\pi(i)}^T$. So, P_π has exactly one 1 in each row. Note that this means that the matrix P_π has exactly n entries that are 1, whereas all the other entries are 0's.

It remains to show that the matrix P_π has exactly one 1 in each column. Since P_π has exactly n many 1's, it is enough to show that no column has more than one 1.

Proposition 2.3.10

Let n be a positive integer, and let $\pi \in S_n$. Then P_π is a permutation matrix.

Proof. Obviously, P_π is an $n \times n$ matrix, all of whose entries are 0's and 1's. Moreover, by the definition of P_π, we have that for each index $i \in \{1, \dots, n\}$, the i-th row of P_π is the row vector $\mathbf{e}^T_{\pi(i)}$. So, P_π has exactly one 1 in each row. Note that this means that the matrix P_π has exactly n entries that are 1, whereas all the other entries are 0's.

It remains to show that the matrix P_π has exactly one 1 in each column. Since P_π has exactly n many 1's, it is enough to show that no column has more than one 1. Since the rows of P_π (from top to bottom) are $\mathbf{e}^T_{\pi(1)}, \dots, \mathbf{e}^T_{\pi(n)}$, and since all those row vectors are pairwise distinct (because π is a permutation), we see that no two rows of P_π have a 1 in the same position.

Proposition 2.3.10

Let n be a positive integer, and let $\pi \in S_n$. Then P_π is a permutation matrix.

Proof. Obviously, P_π is an $n \times n$ matrix, all of whose entries are 0's and 1's. Moreover, by the definition of P_π, we have that for each index $i \in \{1, \dots, n\}$, the i-th row of P_π is the row vector $\mathbf{e}^T_{\pi(i)}$. So, P_π has exactly one 1 in each row. Note that this means that the matrix P_π has exactly n entries that are 1, whereas all the other entries are 0's.

It remains to show that the matrix P_π has exactly one 1 in each column. Since P_π has exactly n many 1's, it is enough to show that no column has more than one 1. Since the rows of P_π (from top to bottom) are $\mathbf{e}^T_{\pi(1)}, \dots, \mathbf{e}^T_{\pi(n)}$, and since all those row vectors are pairwise distinct (because π is a permutation), we see that no two rows of P_π have a 1 in the same position. So, no column of P_π has more than one 1, and we are done. □

- By Proposition 2.3.10, the matrix of a permutation is a permutation matrix.

- By Proposition 2.3.10, the matrix of a permutation is a permutation matrix.
- What about the converse: is every permutation matrix the matrix of some permutation?

- By Proposition 2.3.10, the matrix of a permutation is a permutation matrix.
- What about the converse: is every permutation matrix the matrix of some permutation?
  - The answer to this question is “yes,” and it follows from a simple counting argument, as follows.

- By Proposition 2.3.10, the matrix of a permutation is a permutation matrix.
- What about the converse: is every permutation matrix the matrix of some permutation?
  - The answer to this question is "yes," and it follows from a simple counting argument, as follows.
  - Let n be a positive integer.

- By Proposition 2.3.10, the matrix of a permutation is a permutation matrix.
- What about the converse: is every permutation matrix the matrix of some permutation?
  - The answer to this question is "yes," and it follows from a simple counting argument, as follows.
  - Let n be a positive integer.
  - The $n \times n$ permutation matrices are precisely those $n \times n$ matrices whose columns are the standard basis vectors $\mathbf{e}_1, \dots, \mathbf{e}_n$, appearing in some order. There are $n!$ many ways to order the vectors $\mathbf{e}_1, \dots, \mathbf{e}_n$, and consequently, there are $n!$ many $n \times n$ permutation matrices.

- By Proposition 2.3.10, the matrix of a permutation is a permutation matrix.
- What about the converse: is every permutation matrix the matrix of some permutation?
  - The answer to this question is "yes," and it follows from a simple counting argument, as follows.
  - Let n be a positive integer.
  - The $n \times n$ permutation matrices are precisely those $n \times n$ matrices whose columns are the standard basis vectors $\mathbf{e}_1, \dots, \mathbf{e}_n$, appearing in some order. There are $n!$ many ways to order the vectors $\mathbf{e}_1, \dots, \mathbf{e}_n$, and consequently, there are $n!$ many $n \times n$ permutation matrices.
  - On the other hand, $|S_n| = n!$, and consequently, there are $n!$ many matrices of permutations in S_n.

- By Proposition 2.3.10, the matrix of a permutation is a permutation matrix.
- What about the converse: is every permutation matrix the matrix of some permutation?
    - The answer to this question is "yes," and it follows from a simple counting argument, as follows.
    - Let n be a positive integer.
    - The $n \times n$ permutation matrices are precisely those $n \times n$ matrices whose columns are the standard basis vectors $\mathbf{e}_1, \dots, \mathbf{e}_n$, appearing in some order. There are $n!$ many ways to order the vectors $\mathbf{e}_1, \dots, \mathbf{e}_n$, and consequently, there are $n!$ many $n \times n$ permutation matrices.
    - On the other hand, $|S_n| = n!$, and consequently, there are $n!$ many matrices of permutations in S_n.
        - We are using the fact that different permutations have different matrices.

- By Proposition 2.3.10, the matrix of a permutation is a permutation matrix.
- What about the converse: is every permutation matrix the matrix of some permutation?
  - The answer to this question is "yes," and it follows from a simple counting argument, as follows.
  - Let n be a positive integer.
  - The $n \times n$ permutation matrices are precisely those $n \times n$ matrices whose columns are the standard basis vectors $\mathbf{e}_1, \dots, \mathbf{e}_n$, appearing in some order. There are $n!$ many ways to order the vectors $\mathbf{e}_1, \dots, \mathbf{e}_n$, and consequently, there are $n!$ many $n \times n$ permutation matrices.
  - On the other hand, $|S_n| = n!$, and consequently, there are $n!$ many matrices of permutations in S_n.
    - We are using the fact that different permutations have different matrices.
  - So, the number of $n \times n$ permutation matrices is the same as the number of matrices of permutations in S_n.

- By Proposition 2.3.10, the matrix of a permutation is a permutation matrix.
- What about the converse: is every permutation matrix the matrix of some permutation?
  - The answer to this question is "yes," and it follows from a simple counting argument, as follows.
  - Let n be a positive integer.
  - The $n \times n$ permutation matrices are precisely those $n \times n$ matrices whose columns are the standard basis vectors $\mathbf{e}_1, \dots, \mathbf{e}_n$, appearing in some order. There are $n!$ many ways to order the vectors $\mathbf{e}_1, \dots, \mathbf{e}_n$, and consequently, there are $n!$ many $n \times n$ permutation matrices.
  - On the other hand, $|S_n| = n!$, and consequently, there are $n!$ many matrices of permutations in S_n.
    - We are using the fact that different permutations have different matrices.
  - So, the number of $n \times n$ permutation matrices is the same as the number of matrices of permutations in S_n.
  - It now follows from Proposition 2.3.10 that $n \times n$ permutation matrices are precisely the matrices of permutations in S_n.

Proposition 2.3.11

Let n be a positive integer, and let $\pi \in S_n$ be a permutation. Then both the following hold:

(a) $\forall i \in \{1, \dots, n\}$: $\mathbf{e}_i^T P_\pi = \mathbf{e}_{\pi(i)}$, i.e. the i-th row of P_π is $\mathbf{e}_{\pi(i)}^T$;

(b) $\forall j \in \{1, \dots, n\}$: $P_\pi \mathbf{e}_j = \mathbf{e}_{\pi^{-1}(j)}$, i.e. the j-th column of P_π is $\mathbf{e}_{\pi^{-1}(j)}$.

Consequently, in terms of its rows and columns, P_π can be written as follows:

$$P_\pi = \begin{bmatrix} \mathbf{e}_{\pi(1)}^T \\ \vdots \\ \mathbf{e}_{\pi(n)}^T \end{bmatrix} = \begin{bmatrix} \mathbf{e}_{\pi^{-1}(1)} & \dots & \mathbf{e}_{\pi^{-1}(n)} \end{bmatrix}.$$

Proof. The last statement of the proposition follows immediately from (a) and (b). So, it is enough to prove (a) and (b).

Proposition 2.3.11

(a) $\forall i \in \{1, \dots, n\}$: $\mathbf{e}_i^T P_\pi = \mathbf{e}_{\pi(i)}$, i.e. the i-th row of P_π is $\mathbf{e}_{\pi(i)}^T$;

(b) $\forall j \in \{1, \dots, n\}$: $P_\pi \mathbf{e}_j = \mathbf{e}_{\pi^{-1}(j)}$, i.e. the j-th column of P_π is $\mathbf{e}_{\pi^{-1}(j)}$.

Proof (continued). (a) Fix an index $i \in \{1, \dots, n\}$. By Proposition 1.8.2, $\mathbf{e}_i^T P_\pi$ is precisely the i-th row of the matrix P_π, and by the definition of the matrix P_π, its i-th row is precisely $\mathbf{e}_{\pi(i)}$.

Proposition 2.3.11

- (a) $\forall i \in \{1, \dots, n\}$: $\mathbf{e}_i^T P_\pi = \mathbf{e}_{\pi(i)}$, i.e. the i-th row of P_π is $\mathbf{e}_{\pi(i)}^T$;
- (b) $\forall j \in \{1, \dots, n\}$: $P_\pi \mathbf{e}_j = \mathbf{e}_{\pi^{-1}(j)}$, i.e. the j-th column of P_π is $\mathbf{e}_{\pi^{-1}(j)}$.

Proof (continued). (a) Fix an index $i \in \{1, \dots, n\}$. By Proposition 1.8.2, $\mathbf{e}_i^T P_\pi$ is precisely the i-th row of the matrix P_π, and by the definition of the matrix P_π, its i-th row is precisely $\mathbf{e}_{\pi(i)}$.

(b) Fix an index $j \in \{1, \dots, n\}$.

Proposition 2.3.11

- (a) $\forall i \in \{1, \dots, n\}$: $\mathbf{e}_i^T P_\pi = \mathbf{e}_{\pi(i)}$, i.e. the i-th row of P_π is $\mathbf{e}_{\pi(i)}^T$;
- (b) $\forall j \in \{1, \dots, n\}$: $P_\pi \mathbf{e}_j = \mathbf{e}_{\pi^{-1}(j)}$, i.e. the j-th column of P_π is $\mathbf{e}_{\pi^{-1}(j)}$.

Proof (continued). (a) Fix an index $i \in \{1, \dots, n\}$. By Proposition 1.8.2, $\mathbf{e}_i^T P_\pi$ is precisely the i-th row of the matrix P_π, and by the definition of the matrix P_π, its i-th row is precisely $\mathbf{e}_{\pi(i)}$.

(b) Fix an index $j \in \{1, \dots, n\}$. By Proposition 1.4.4, $P_\pi \mathbf{e}_j$ is precisely the j-th column of the matrix P_π.

Proposition 2.3.11

(a) $\forall i \in \{1, \dots, n\}$: $\mathbf{e}_i^T P_\pi = \mathbf{e}_{\pi(i)}$, i.e. the i-th row of P_π is $\mathbf{e}_{\pi(i)}^T$;

(b) $\forall j \in \{1, \dots, n\}$: $P_\pi \mathbf{e}_j = \mathbf{e}_{\pi^{-1}(j)}$, i.e. the j-th column of P_π is $\mathbf{e}_{\pi^{-1}(j)}$.

Proof (continued). (a) Fix an index $i \in \{1, \dots, n\}$. By Proposition 1.8.2, $\mathbf{e}_i^T P_\pi$ is precisely the i-th row of the matrix P_π, and by the definition of the matrix P_π, its i-th row is precisely $\mathbf{e}_{\pi(i)}$.

(b) Fix an index $j \in \{1, \dots, n\}$. By Proposition 1.4.4, $P_\pi \mathbf{e}_j$ is precisely the j-th column of the matrix P_π. Set $i := \pi^{-1}(j)$, so that $j = \pi(i)$.

Proposition 2.3.11

- (a) $\forall i \in \{1, \dots, n\}$: $\mathbf{e}_i^T P_\pi = \mathbf{e}_{\pi(i)}$, i.e. the i-th row of P_π is $\mathbf{e}_{\pi(i)}^T$;
- (b) $\forall j \in \{1, \dots, n\}$: $P_\pi \mathbf{e}_j = \mathbf{e}_{\pi^{-1}(j)}$, i.e. the j-th column of P_π is $\mathbf{e}_{\pi^{-1}(j)}$.

Proof (continued). (a) Fix an index $i \in \{1, \dots, n\}$. By Proposition 1.8.2, $\mathbf{e}_i^T P_\pi$ is precisely the i-th row of the matrix P_π, and by the definition of the matrix P_π, its i-th row is precisely $\mathbf{e}_{\pi(i)}$.

(b) Fix an index $j \in \{1, \dots, n\}$. By Proposition 1.4.4, $P_\pi \mathbf{e}_j$ is precisely the j-th column of the matrix P_π. Set $i := \pi^{-1}(j)$, so that $j = \pi(i)$. By (a), the i-th row of P_π is the row vector $\mathbf{e}_{\pi(i)}^T = \mathbf{e}_j^T$.

Proposition 2.3.11

- (a) $\forall i \in \{1, \dots, n\}$: $\mathbf{e}_i^T P_\pi = \mathbf{e}_{\pi(i)}$, i.e. the i-th row of P_π is $\mathbf{e}_{\pi(i)}^T$;
- (b) $\forall j \in \{1, \dots, n\}$: $P_\pi \mathbf{e}_j = \mathbf{e}_{\pi^{-1}(j)}$, i.e. the j-th column of P_π is $\mathbf{e}_{\pi^{-1}(j)}$.

Proof (continued). (a) Fix an index $i \in \{1, \dots, n\}$. By Proposition 1.8.2, $\mathbf{e}_i^T P_\pi$ is precisely the i-th row of the matrix P_π, and by the definition of the matrix P_π, its i-th row is precisely $\mathbf{e}_{\pi(i)}$.

(b) Fix an index $j \in \{1, \dots, n\}$. By Proposition 1.4.4, $P_\pi \mathbf{e}_j$ is precisely the j-th column of the matrix P_π. Set $i := \pi^{-1}(j)$, so that $j = \pi(i)$. By (a), the i-th row of P_π is the row vector $\mathbf{e}_{\pi(i)}^T = \mathbf{e}_j^T$. So, P_π has 1 in its (i, j)-th entry.

Proposition 2.3.11

- (a) $\forall i \in \{1, \dots, n\}$: $\mathbf{e}_i^T P_\pi = \mathbf{e}_{\pi(i)}$, i.e. the i-th row of P_π is $\mathbf{e}_{\pi(i)}^T$;
- (b) $\forall j \in \{1, \dots, n\}$: $P_\pi \mathbf{e}_j = \mathbf{e}_{\pi^{-1}(j)}$, i.e. the j-th column of P_π is $\mathbf{e}_{\pi^{-1}(j)}$.

Proof (continued). (a) Fix an index $i \in \{1, \dots, n\}$. By Proposition 1.8.2, $\mathbf{e}_i^T P_\pi$ is precisely the i-th row of the matrix P_π, and by the definition of the matrix P_π, its i-th row is precisely $\mathbf{e}_{\pi(i)}$.

(b) Fix an index $j \in \{1, \dots, n\}$. By Proposition 1.4.4, $P_\pi \mathbf{e}_j$ is precisely the j-th column of the matrix P_π. Set $i := \pi^{-1}(j)$, so that $j = \pi(i)$. By (a), the i-th row of P_π is the row vector $\mathbf{e}_{\pi(i)}^T = \mathbf{e}_j^T$. So, P_π has 1 in its (i, j)-th entry. Since P_π is a permutation matrix (by Proposition 2.3.10), and therefore has exactly one 1 in each column, it follows that the j-th column of P_π is $\mathbf{e}_i = \mathbf{e}_{\pi^{-1}(j)}$. $\square$

Proposition 2.3.11

Let n be a positive integer, and let $\pi \in S_n$ be a permutation. Then both the following hold:

(a) $\forall i \in \{1, \dots, n\}$: $\mathbf{e}_i^T P_\pi = \mathbf{e}_{\pi(i)}$, i.e. the i-th row of P_π is $\mathbf{e}_{\pi(i)}^T$;

(b) $\forall j \in \{1, \dots, n\}$: $P_\pi \mathbf{e}_j = \mathbf{e}_{\pi^{-1}(j)}$, i.e. the j-th column of P_π is $\mathbf{e}_{\pi^{-1}(j)}$.

Consequently, in terms of its rows and columns, P_π can be written as follows:

$$P_\pi = \begin{bmatrix} \mathbf{e}_{\pi(1)}^T \\ \vdots \\ \mathbf{e}_{\pi(n)}^T \end{bmatrix} = \begin{bmatrix} \mathbf{e}_{\pi^{-1}(1)} & \dots & \mathbf{e}_{\pi^{-1}(n)} \end{bmatrix}.$$

Proof. The last statement of the proposition follows immediately from (a) and (b). So, it is enough to prove (a) and (b).

Proposition 2.3.12

Let n be a positive integer, and let $\pi \in S_n$. Then

$$P_{\pi^{-1}} \;=\; P_\pi^T.$$

Proof.

Proposition 2.3.12

Let n be a positive integer, and let $\pi \in S_n$. Then

$$P_{\pi^{-1}} = P_\pi^T.$$

Proof. We have that

$$P_\pi^T \overset{(*)}{=} \left(\begin{bmatrix} \mathbf{e}_{\pi^{-1}(1)} & \dots & \mathbf{e}_{\pi^{-1}(n)} \end{bmatrix} \right)^T = \begin{bmatrix} \mathbf{e}_{\pi^{-1}(1)}^T \\ \vdots \\ \mathbf{e}_{\pi^{-1}(n)}^T \end{bmatrix} \overset{(*)}{=} P_{\pi^{-1}},$$

where both instances of (*) follow from Proposition 2.3.11. □

Proposition 2.3.11

(a) $\forall i \in \{1, \ldots, n\}$: $\mathbf{e}_i^T P_\pi = \mathbf{e}_{\pi(i)}$, i.e. the i-th row of P_π is $\mathbf{e}_{\pi(i)}^T$;

Proposition 2.3.13

Let n be a positive integer, and let σ and π be permutations in S_n. Then $P_{\sigma \circ \pi} = P_\pi P_\sigma$.

Proof.

Proposition 2.3.11

(a) $\forall i \in \{1, \ldots, n\}$: $\mathbf{e}_i^T P_\pi = \mathbf{e}_{\pi(i)}$, i.e. the i-th row of P_π is $\mathbf{e}_{\pi(i)}^T$;

Proposition 2.3.13

Let n be a positive integer, and let σ and π be permutations in S_n. Then $P_{\sigma\circ\pi} = P_\pi P_\sigma$.

Proof. It suffices to show that matrices $P_{\sigma\circ\pi}$ and $P_\pi P_\sigma$ have the same corresponding rows.

Proposition 2.3.11

(a) $\forall i \in \{1, \dots, n\}$: $\mathbf{e}_i^T P_\pi = \mathbf{e}_{\pi(i)}$, i.e. the i-th row of P_π is $\mathbf{e}_{\pi(i)}^T$;

Proposition 2.3.13

Let n be a positive integer, and let σ and π be permutations in S_n. Then $P_{\sigma \circ \pi} = P_\pi P_\sigma$.

Proof. It suffices to show that matrices $P_{\sigma \circ \pi}$ and $P_\pi P_\sigma$ have the same corresponding rows. Fix an index $i \in \{1, \dots, n\}$.

Proposition 2.3.11

(a) $\forall i \in \{1, \ldots, n\}$: $\mathbf{e}_i^T P_\pi = \mathbf{e}_{\pi(i)}$, i.e. the i-th row of P_π is $\mathbf{e}_{\pi(i)}^T$;

Proposition 2.3.13

Let n be a positive integer, and let σ and π be permutations in S_n. Then $P_{\sigma \circ \pi} = P_\pi P_\sigma$.

Proof. It suffices to show that matrices $P_{\sigma \circ \pi}$ and $P_\pi P_\sigma$ have the same corresponding rows. Fix an index $i \in \{1, \ldots, n\}$. By Proposition 1.8.2, the i-th row of the matrix $P_{\sigma \circ \pi}$ is $\mathbf{e}_i^T P_{\sigma \circ \pi}$, and the i-th row of the matrix $P_\pi P_\sigma$ is $\mathbf{e}_i^T (P_\pi P_\sigma)$.

Proposition 2.3.11

(a) $\forall i \in \{1,\dots,n\}$: $\mathbf{e}_i^T P_\pi = \mathbf{e}_{\pi(i)}$, i.e. the i-th row of P_π is $\mathbf{e}_{\pi(i)}^T$;

Proposition 2.3.13

Let n be a positive integer, and let σ and π be permutations in S_n. Then $P_{\sigma\circ\pi} = P_\pi P_\sigma$.

Proof. It suffices to show that matrices $P_{\sigma\circ\pi}$ and $P_\pi P_\sigma$ have the same corresponding rows. Fix an index $i \in \{1,\dots,n\}$. By Proposition 1.8.2, the i-th row of the matrix $P_{\sigma\circ\pi}$ is $\mathbf{e}_i^T P_{\sigma\circ\pi}$, and the i-th row of the matrix $P_\pi P_\sigma$ is $\mathbf{e}_i^T(P_\pi P_\sigma)$. So, we just need to show that $\mathbf{e}_i^T P_{\sigma\circ\pi} = \mathbf{e}_i^T(P_\pi P_\sigma)$.

Proposition 2.3.11

(a) $\forall i \in \{1, \ldots, n\}$: $\mathbf{e}_i^T P_\pi = \mathbf{e}_{\pi(i)}$, i.e. the i-th row of P_π is $\mathbf{e}_{\pi(i)}^T$;

Proposition 2.3.13

Let n be a positive integer, and let σ and π be permutations in S_n. Then $P_{\sigma \circ \pi} = P_\pi P_\sigma$.

Proof. It suffices to show that matrices $P_{\sigma \circ \pi}$ and $P_\pi P_\sigma$ have the same corresponding rows. Fix an index $i \in \{1, \ldots, n\}$. By Proposition 1.8.2, the i-th row of the matrix $P_{\sigma \circ \pi}$ is $\mathbf{e}_i^T P_{\sigma \circ \pi}$, and the i-th row of the matrix $P_\pi P_\sigma$ is $\mathbf{e}_i^T (P_\pi P_\sigma)$. So, we just need to show that $\mathbf{e}_i^T P_{\sigma \circ \pi} = \mathbf{e}_i^T (P_\pi P_\sigma)$. But follows easily via repeated application of Proposition 2.3.11(a).

Proposition 2.3.11

(a) $\forall i \in \{1, \dots, n\}$: $\mathbf{e}_i^T P_\pi = \mathbf{e}_{\pi(i)}$, i.e. the i-th row of P_π is $\mathbf{e}_{\pi(i)}^T$;

Proposition 2.3.13

Let n be a positive integer, and let σ and π be permutations in S_n. Then $P_{\sigma\circ\pi} = P_\pi P_\sigma$.

Proof. It suffices to show that matrices $P_{\sigma\circ\pi}$ and $P_\pi P_\sigma$ have the same corresponding rows. Fix an index $i \in \{1, \dots, n\}$. By Proposition 1.8.2, the i-th row of the matrix $P_{\sigma\circ\pi}$ is $\mathbf{e}_i^T P_{\sigma\circ\pi}$, and the i-th row of the matrix $P_\pi P_\sigma$ is $\mathbf{e}_i^T(P_\pi P_\sigma)$. So, we just need to show that $\mathbf{e}_i^T P_{\sigma\circ\pi} = \mathbf{e}_i^T(P_\pi P_\sigma)$. But follows easily via repeated application of Proposition 2.3.11(a). Indeed, we have that

$$\begin{aligned} \mathbf{e}_i^T(P_\pi P_\sigma) &= (\mathbf{e}_i^T P_\pi)P_\sigma \overset{(*)}{=} \mathbf{e}_{\pi(i)}^T P_\sigma \overset{(*)}{=} \mathbf{e}_{\sigma\left(\pi(i)\right)} \\ &= \mathbf{e}_{(\sigma\circ\pi)(i)}^T \overset{(*)}{=} \mathbf{e}_i^T P_{\sigma\circ\pi}, \end{aligned}$$

where all three instances of (*) follow from Prop. 2.3.11(a). □

Theorem 2.3.14

Let n be a positive integer, and let $\pi \in S_n$. Then P_π is invertible, and moreover,

$$P_\pi^{-1} \;=\; P_{\pi^{-1}} \;=\; P_\pi^T.$$

Proof.

Theorem 2.3.14

Let n be a positive integer, and let $\pi \in S_n$. Then P_π is invertible, and moreover,

$$P_\pi^{-1} \;=\; P_{\pi^{-1}} \;=\; P_\pi^T.$$

Proof. The fact that $P_{\pi^{-1}} = P_\pi^T$ follows immediately from Proposition 2.3.12.

Theorem 2.3.14

Let n be a positive integer, and let $\pi \in S_n$. Then P_π is invertible, and moreover,

$$P_\pi^{-1} \;=\; P_{\pi^{-1}} \;=\; P_\pi^T.$$

Proof. The fact that $P_{\pi^{-1}} = P_\pi^T$ follows immediately from Proposition 2.3.12. It remains to show that P_π is invertible, and that its inverse is $P_{\pi^{-1}}$.

Theorem 2.3.14

Let n be a positive integer, and let $\pi \in S_n$. Then P_π is invertible, and moreover,

$$P_\pi^{-1} \;=\; P_{\pi^{-1}} \;=\; P_\pi^T.$$

Proof. The fact that $P_{\pi^{-1}} = P_\pi^T$ follows immediately from Proposition 2.3.12. It remains to show that P_π is invertible, and that its inverse is $P_{\pi^{-1}}$.

We now compute:

$$P_\pi P_{\pi^{-1}} \;\overset{(*)}{=}\; P_{\pi^{-1}\circ\pi} \;=\; P_{1_n} \;=\; I_n,$$

where (*) follows immediately from Proposition 2.3.13.

Theorem 2.3.14

Let n be a positive integer, and let $\pi \in S_n$. Then P_π is invertible, and moreover,

$$P_\pi^{-1} \;=\; P_{\pi^{-1}} \;=\; P_\pi^T.$$

Proof. The fact that $P_{\pi^{-1}} = P_\pi^T$ follows immediately from Proposition 2.3.12. It remains to show that P_π is invertible, and that its inverse is $P_{\pi^{-1}}$.

We now compute:

$$P_\pi P_{\pi^{-1}} \;\overset{(*)}{=}\; P_{\pi^{-1}\circ\pi} \;=\; P_{1_n} \;=\; I_n,$$

where (*) follows immediately from Proposition 2.3.13. Analogously, $P_{\pi^{-1}} P_\pi = I_n$.

Theorem 2.3.14

Let n be a positive integer, and let $\pi \in S_n$. Then P_π is invertible, and moreover,

$$P_\pi^{-1} \;=\; P_{\pi^{-1}} \;=\; P_\pi^T.$$

Proof. The fact that $P_{\pi^{-1}} = P_\pi^T$ follows immediately from Proposition 2.3.12. It remains to show that P_π is invertible, and that its inverse is $P_{\pi^{-1}}$.

We now compute:

$$P_\pi P_{\pi^{-1}} \;\overset{(*)}{=}\; P_{\pi^{-1}\circ\pi} \;=\; P_{1_n} \;=\; I_n,$$

where (*) follows immediately from Proposition 2.3.13. Analogously, $P_{\pi^{-1}} P_\pi = I_n$. So, P_π and $P_{\pi^{-1}}$ are invertible and are each other's inverses. This completes the argument. □

Theorem 2.3.14

Let n be a positive integer, and let $\pi \in S_n$. Then P_π is invertible, and moreover,

$$P_\pi^{-1} = P_{\pi^{-1}} = P_\pi^T.$$

- **Remark:** A matrix $Q \in \mathbb{R}^{n\times n}$ is *orthogonal* if it satisfies $Q^T Q = I_n$.
  - Theorem 2.3.14 guarantees that permutation matrices are orthogonal (as long as we consider the 0's and 1's in those matrices as belonging to $\mathbb{R}$, rather than to some other field).

- As our next theorem (Theorem 2.3.15, next slide) shows, multiplying a matrix by a permutation matrix on the left permutes the rows of the original matrix.

- As our next theorem (Theorem 2.3.15, next slide) shows, multiplying a matrix by a permutation matrix on the left permutes the rows of the original matrix.
- On the other hand, multiplying a matrix by a permutation matrix on the right permutes the columns of the original matrix.

Theorem 2.3.15

Let $A = \begin{bmatrix} \mathbf{r}_1 \\ \vdots \\ \mathbf{r}_n \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{bmatrix}$ be an $n \times m$ matrix with entries in some field $\mathbb{F}$. Then all the following hold:

(a) for all $\pi \in S_n$, we have that

$$P_\pi A = \begin{bmatrix} \mathbf{r}_{\pi(1)} \\ \vdots \\ \mathbf{r}_{\pi(n)} \end{bmatrix};$$

(b) for all $\pi \in S_m$, we have that

$$AP_\pi = \begin{bmatrix} \mathbf{a}_{\pi^{-1}(1)} & \dots & \mathbf{a}_{\pi^{-1}(m)} \end{bmatrix};$$

(c) for all $\pi \in S_m$, we have that

$$AP_\pi^T = \begin{bmatrix} \mathbf{a}_{\pi(1)} & \dots & \mathbf{a}_{\pi(m)} \end{bmatrix}.$$

Proof.

Theorem 2.3.15

Let $A = \begin{bmatrix} \mathbf{r}_1 \\ \vdots \\ \mathbf{r}_n \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{bmatrix}$ be an $n \times m$ matrix with entries in some field $\mathbb{F}$. Then all the following hold:

(a) for all $\pi \in S_n$, we have that

$$P_\pi A = \begin{bmatrix} \mathbf{r}_{\pi(1)} \\ \vdots \\ \mathbf{r}_{\pi(n)} \end{bmatrix};$$

(b) for all $\pi \in S_m$, we have that

$$AP_\pi = \begin{bmatrix} \mathbf{a}_{\pi^{-1}(1)} & \dots & \mathbf{a}_{\pi^{-1}(m)} \end{bmatrix};$$

(c) for all $\pi \in S_m$, we have that

$$AP_\pi^T = \begin{bmatrix} \mathbf{a}_{\pi(1)} & \dots & \mathbf{a}_{\pi(m)} \end{bmatrix}.$$

Proof. We prove (b). Parts (a) and (c) are in the Lecture Notes.

Theorem 2.3.15

Let $A = \begin{bmatrix} \mathbf{r}_1 \\ \vdots \\ \mathbf{r}_n \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{bmatrix}$ be an $n \times m$ matrix with entries in some field $\mathbb{F}$. Then all the following hold:

- (b) for all $\pi \in S_m$, we have that

$$AP_\pi \quad = \quad \begin{bmatrix} \mathbf{a}_{\pi^{-1}(1)} & \dots & \mathbf{a}_{\pi^{-1}(m)} \end{bmatrix};$$

Proof of (b).

Theorem 2.3.15

Let $A = \begin{bmatrix} \mathbf{r}_1 \\ \vdots \\ \mathbf{r}_n \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{bmatrix}$ be an $n \times m$ matrix with entries in some field $\mathbb{F}$. Then all the following hold:

(b) for all $\pi \in S_m$, we have that

$$AP_\pi = \begin{bmatrix} \mathbf{a}_{\pi^{-1}(1)} & \dots & \mathbf{a}_{\pi^{-1}(m)} \end{bmatrix};$$

Proof of (b). Fix any permutation $\pi \in S_m$. In what follows, $\mathbf{e}_1, \dots, \mathbf{e}_m$ are the standard basis vectors of $\mathbb{F}^m$. We compute:

$$\begin{aligned} AP_\pi &= A\begin{bmatrix} \mathbf{e}_{\pi^{-1}(1)} & \dots & \mathbf{e}_{\pi^{-1}(m)} \end{bmatrix} && \text{by Proposition 2.3.11} \\ &= \begin{bmatrix} A\mathbf{e}_{\pi^{-1}(1)} & \dots & A\mathbf{e}_{\pi^{-1}(m)} \end{bmatrix} && \text{by the definition of matrix multiplication} \\ &= \begin{bmatrix} \mathbf{a}_{\pi^{-1}(1)} & \dots & \mathbf{a}_{\pi^{-1}(m)} \end{bmatrix} && \text{by Proposition 1.4.4.} \end{aligned}$$

This proves (b). □

Theorem 2.3.15

Let $A = \begin{bmatrix} \mathbf{r}_1 \\ \vdots \\ \mathbf{r}_n \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{bmatrix}$ be an $n \times m$ matrix with entries in some field $\mathbb{F}$. Then all the following hold:

(a) for all $\pi \in S_n$, we have that

$$P_\pi A = \begin{bmatrix} \mathbf{r}_{\pi(1)} \\ \vdots \\ \mathbf{r}_{\pi(n)} \end{bmatrix};$$

(b) for all $\pi \in S_m$, we have that

$$AP_\pi = \begin{bmatrix} \mathbf{a}_{\pi^{-1}(1)} & \dots & \mathbf{a}_{\pi^{-1}(m)} \end{bmatrix};$$

(c) for all $\pi \in S_m$, we have that

$$AP_\pi^T = \begin{bmatrix} \mathbf{a}_{\pi(1)} & \dots & \mathbf{a}_{\pi(m)} \end{bmatrix}.$$

2. Orthogonal matrices

2. Orthogonal matrices

- In our study of orthogonal matrices, we assume that $\mathbb{R}^n$ is equipped with the standard scalar product $\cdot$ and the induced norm $||\cdot||$.

2. Orthogonal matrices

- In our study of orthogonal matrices, we assume that $\mathbb{R}^n$ is equipped with the standard scalar product $\cdot$ and the induced norm $||\cdot||$.

Definition

A matrix $Q \in \mathbb{R}^{n\times n}$ is *orthogonal* if it satisfies $Q^T Q = I_n$.

Definition

A matrix $Q \in \mathbb{R}^{n \times n}$ is *orthogonal* if it satisfies $Q^T Q = I_n$.

Definition

A matrix $Q \in \mathbb{R}^{n \times n}$ is *orthogonal* if it satisfies $Q^T Q = I_n$.

- Obviously, matrices I_n and $-I_n$ are orthogonal.

Definition

A matrix $Q \in \mathbb{R}^{n \times n}$ is *orthogonal* if it satisfies $Q^T Q = I_n$.

- Obviously, matrices I_n and $-I_n$ are orthogonal.
- By Theorem 2.3.14, permutation matrices are orthogonal (as long as we consider the 0's and 1's in those matrices as being real numbers).

Theorem 2.3.14

Let n be a positive integer, and let $\pi \in S_n$. Then P_π is invertible, and moreover,

$$P_\pi^{-1} \;=\; P_{\pi^{-1}} \;=\; P_\pi^T.$$

Definition

A matrix $Q \in \mathbb{R}^{n \times n}$ is *orthogonal* if it satisfies $Q^T Q = I_n$.

- Obviously, matrices I_n and $-I_n$ are orthogonal.
- By Theorem 2.3.14, permutation matrices are orthogonal (as long as we consider the 0's and 1's in those matrices as being real numbers).

Theorem 2.3.14

Let n be a positive integer, and let $\pi \in S_n$. Then P_π is invertible, and moreover,

$$P_\pi^{-1} \ = \ P_{\pi^{-1}} \ = \ P_\pi^T.$$

- The matrices mentioned so far all have entries only $-1, 0, 1$. However, there are many other orthogonal matrices, and we will see a couple of examples later.

- Reminder:

Corollary 3.3.18

Let $\mathbb{F}$ be field, and let $A, B \in \mathbb{F}^{n \times n}$ be such that $AB = I_n$ or $BA = I_n$. Then $AB = BA = I_n$, i.e. A and B are both invertible and are each other's inverses.

Theorem 6.8.1

Let $Q \in \mathbb{R}^{n \times n}$. Then the following are equivalent:

(a) Q is orthogonal (i.e. satisfies $Q^T Q = I_n$);

(b) Q is invertible and satisfies $Q^{-1} = Q^T$;

(c) $QQ^T = I_n$;

(d) Q^T is orthogonal;

(e) Q is invertible and Q^{-1} is orthogonal;

(f) the columns of Q form an orthonormal basis of $\mathbb{R}^n$;

(g) the columns of Q^T form an orthonormal basis of $\mathbb{R}^n$.

Proof.

Theorem 6.8.1

Let $Q \in \mathbb{R}^{n \times n}$. Then the following are equivalent:

(a) Q is orthogonal (i.e. satisfies $Q^T Q = I_n$);
(b) Q is invertible and satisfies $Q^{-1} = Q^T$;
(c) $QQ^T = I_n$;
(d) Q^T is orthogonal;
(e) Q is invertible and Q^{-1} is orthogonal;
(f) the columns of Q form an orthonormal basis of $\mathbb{R}^n$;
(g) the columns of Q^T form an orthonormal basis of $\mathbb{R}^n$.

Proof. By Corollary 3.3.18, we have that (a), (b), and (c) are equivalent.

Theorem 6.8.1

Let $Q \in \mathbb{R}^{n \times n}$. Then the following are equivalent:

(a) Q is orthogonal (i.e. satisfies $Q^T Q = I_n$);
(b) Q is invertible and satisfies $Q^{-1} = Q^T$;
(c) $QQ^T = I_n$;
(d) Q^T is orthogonal;
(e) Q is invertible and Q^{-1} is orthogonal;
(f) the columns of Q form an orthonormal basis of $\mathbb{R}^n$;
(g) the columns of Q^T form an orthonormal basis of $\mathbb{R}^n$.

Proof. By Corollary 3.3.18, we have that (a), (b), and (c) are equivalent. Moreover, since $(Q^T)^T = Q$, we have that (c) and (d) are equivalent.

Theorem 6.8.1

Let $Q \in \mathbb{R}^{n\times n}$. Then the following are equivalent:

(a) Q is orthogonal (i.e. satisfies $Q^TQ = I_n$);

(b) Q is invertible and satisfies $Q^{-1} = Q^T$;

(c) $QQ^T = I_n$;

(d) Q^T is orthogonal;

(e) Q is invertible and Q^{-1} is orthogonal;

(f) the columns of Q form an orthonormal basis of $\mathbb{R}^n$;

(g) the columns of Q^T form an orthonormal basis of $\mathbb{R}^n$.

Proof. By Corollary 3.3.18, we have that (a), (b), and (c) are equivalent. Moreover, since $(Q^T)^T = Q$, we have that (c) and (d) are equivalent. This proves that (a), (b), (c), and (d) are equivalent.

Theorem 6.8.1

Let $Q \in \mathbb{R}^{n \times n}$. Then the following are equivalent:

1. (a) Q is orthogonal (i.e. satisfies $Q^T Q = I_n$);
2. (b) Q is invertible and satisfies $Q^{-1} = Q^T$;
3. (c) $QQ^T = I_n$;
4. (d) Q^T is orthogonal;
5. (e) Q is invertible and Q^{-1} is orthogonal;
6. (f) the columns of Q form an orthonormal basis of $\mathbb{R}^n$;
7. (g) the columns of Q^T form an orthonormal basis of $\mathbb{R}^n$.

Proof (continued). Next, (b) and (d) together imply (e).

Theorem 6.8.1

Let $Q \in \mathbb{R}^{n \times n}$. Then the following are equivalent:

(a) Q is orthogonal (i.e. satisfies $Q^T Q = I_n$);

(b) Q is invertible and satisfies $Q^{-1} = Q^T$;

(c) $QQ^T = I_n$;

(d) Q^T is orthogonal;

(e) Q is invertible and Q^{-1} is orthogonal;

(f) the columns of Q form an orthonormal basis of $\mathbb{R}^n$;

(g) the columns of Q^T form an orthonormal basis of $\mathbb{R}^n$.

Proof (continued). Next, (b) and (d) together imply (e).

Suppose now that (e) holds. Then by applying "(a) $\Longrightarrow$ (b)" to the matrix Q^{-1}, we see that Q^{-1} is invertible and satisfies $(Q^{-1})^{-1} = (Q^{-1})^T$. Consequently, $Q^{-1} = Q^T$, and it follows that (b) holds.

Theorem 6.8.1

Let $Q \in \mathbb{R}^{n \times n}$. Then the following are equivalent:

(a) Q is orthogonal (i.e. satisfies $Q^T Q = I_n$);

(b) Q is invertible and satisfies $Q^{-1} = Q^T$;

(c) $QQ^T = I_n$;

(d) Q^T is orthogonal;

(e) Q is invertible and Q^{-1} is orthogonal;

(f) the columns of Q form an orthonormal basis of $\mathbb{R}^n$;

(g) the columns of Q^T form an orthonormal basis of $\mathbb{R}^n$.

Proof (continued). So far, we have established that (a), (b), (c), (d), and (e) are equivalent.

Theorem 6.8.1

Let $Q \in \mathbb{R}^{n \times n}$. Then the following are equivalent:

(a) Q is orthogonal (i.e. satisfies $Q^T Q = I_n$);
(b) Q is invertible and satisfies $Q^{-1} = Q^T$;
(c) $QQ^T = I_n$;
(d) Q^T is orthogonal;
(e) Q is invertible and Q^{-1} is orthogonal;
(f) the columns of Q form an orthonormal basis of $\mathbb{R}^n$;
(g) the columns of Q^T form an orthonormal basis of $\mathbb{R}^n$.

Proof (continued). So far, we have established that (a), (b), (c), (d), and (e) are equivalent.

Let us now show that (a) and (f) are equivalent.

Theorem 6.8.1

- (a) Q is orthogonal (i.e. satisfies $Q^T Q = I_n$);
- (f) the columns of Q form an orthonormal basis of $\mathbb{R}^n$;

Proof (continued).

Theorem 6.8.1

(a) Q is orthogonal (i.e. satisfies $Q^TQ = I_n$);

(f) the columns of Q form an orthonormal basis of $\mathbb{R}^n$;

Proof (continued). Set $Q = \begin{bmatrix} \mathbf{q}_1 & \dots & \mathbf{q}_n \end{bmatrix}$. Then

$$
\begin{aligned}
Q^TQ &= \begin{bmatrix} \mathbf{q}_1^T \\ \mathbf{q}_2^T \\ \vdots \\ \mathbf{q}_n^T \end{bmatrix} \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \dots & \mathbf{q}_n \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{q}_1 \cdot \mathbf{q}_1 & \mathbf{q}_1 \cdot \mathbf{q}_2 & \dots & \mathbf{q}_1 \cdot \mathbf{q}_n \\ \mathbf{q}_2 \cdot \mathbf{q}_1 & \mathbf{q}_2 \cdot \mathbf{q}_2 & \dots & \mathbf{q}_2 \cdot \mathbf{q}_n \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{q}_n \cdot \mathbf{q}_1 & \mathbf{q}_n \cdot \mathbf{q}_2 & \dots & \mathbf{q}_n \cdot \mathbf{q}_n \end{bmatrix}.
\end{aligned}
$$

Theorem 6.8.1

- (a) Q is orthogonal (i.e. satisfies $Q^TQ = I_n$);
- (f) the columns of Q form an orthonormal basis of $\mathbb{R}^n$;

Proof (continued). Set $Q = \begin{bmatrix} \mathbf{q}_1 & \dots & \mathbf{q}_n \end{bmatrix}$. Then

$$
\begin{aligned}
Q^TQ &= \begin{bmatrix} \mathbf{q}_1^T \\ \mathbf{q}_2^T \\ \vdots \\ \mathbf{q}_n^T \end{bmatrix} \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \dots & \mathbf{q}_n \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{q}_1 \cdot \mathbf{q}_1 & \mathbf{q}_1 \cdot \mathbf{q}_2 & \dots & \mathbf{q}_1 \cdot \mathbf{q}_n \\ \mathbf{q}_2 \cdot \mathbf{q}_1 & \mathbf{q}_2 \cdot \mathbf{q}_2 & \dots & \mathbf{q}_2 \cdot \mathbf{q}_n \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{q}_n \cdot \mathbf{q}_1 & \mathbf{q}_n \cdot \mathbf{q}_2 & \dots & \mathbf{q}_n \cdot \mathbf{q}_n \end{bmatrix}.
\end{aligned}
$$

So, $Q^TQ = I_n$ iff $\{\mathbf{q}_1, \dots, \mathbf{q}_n\}$ is an orthonormal set.

Theorem 6.8.1

(a) Q is orthogonal (i.e. satisfies $Q^T Q = I_n$);

(f) the columns of Q form an orthonormal basis of $\mathbb{R}^n$;

Proof (continued). Set $Q = \begin{bmatrix} \mathbf{q}_1 & \dots & \mathbf{q}_n \end{bmatrix}$. Then

$$
\begin{aligned}
Q^T Q &= \begin{bmatrix} \mathbf{q}_1^T \\ \mathbf{q}_2^T \\ \vdots \\ \mathbf{q}_n^T \end{bmatrix} \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \dots & \mathbf{q}_n \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{q}_1 \cdot \mathbf{q}_1 & \mathbf{q}_1 \cdot \mathbf{q}_2 & \dots & \mathbf{q}_1 \cdot \mathbf{q}_n \\ \mathbf{q}_2 \cdot \mathbf{q}_1 & \mathbf{q}_2 \cdot \mathbf{q}_2 & \dots & \mathbf{q}_2 \cdot \mathbf{q}_n \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{q}_n \cdot \mathbf{q}_1 & \mathbf{q}_n \cdot \mathbf{q}_2 & \dots & \mathbf{q}_n \cdot \mathbf{q}_n \end{bmatrix}.
\end{aligned}
$$

So, $Q^T Q = I_n$ iff $\{\mathbf{q}_1, \dots, \mathbf{q}_n\}$ is an orthonormal set. But by Proposition 6.3.4(b), any orthonormal set of n vectors in $\mathbb{R}^n$ is in fact an orthonormal basis of $\mathbb{R}^n$. So, (a) and (f) are equivalent.

Theorem 6.8.1

Let $Q \in \mathbb{R}^{n \times n}$. Then the following are equivalent:

(a) Q is orthogonal (i.e. satisfies $Q^T Q = I_n$);

(b) Q is invertible and satisfies $Q^{-1} = Q^T$;

(c) $QQ^T = I_n$;

(d) Q^T is orthogonal;

(e) Q is invertible and Q^{-1} is orthogonal;

(f) the columns of Q form an orthonormal basis of $\mathbb{R}^n$;

(g) the columns of Q^T form an orthonormal basis of $\mathbb{R}^n$.

Proof (continued). Analogously to "(a) $\iff$ (f)," we have that (d) and (g) are equivalent. $\square$

Theorem 6.8.1

Let $Q \in \mathbb{R}^{n\times n}$. Then the following are equivalent:

(a) Q is orthogonal (i.e. satisfies $Q^TQ = I_n$);
(b) Q is invertible and satisfies $Q^{-1} = Q^T$;
(c) $QQ^T = I_n$;
(d) Q^T is orthogonal;
(e) Q is invertible and Q^{-1} is orthogonal;
(f) the columns of Q form an orthonormal basis of $\mathbb{R}^n$;
(g) the columns of Q^T form an orthonormal basis of $\mathbb{R}^n$.

- We can make new orthogonal matrices out of old ones, as Propositions 6.8.2, 6.8.3, and 6.8.4 (below and next slide) show.
- The proofs of these propositions are easy and are in the Lecture Notes (we omit them here).

Proposition 6.8.2

Let

$$Q = \begin{bmatrix} \mathbf{q}_1 & \dots & \mathbf{q}_n \end{bmatrix} = \begin{bmatrix} \mathbf{r}_1^T \\ \vdots \\ \mathbf{r}_n^T \end{bmatrix}$$

be an orthogonal matrix in $\mathbb{R}^n$. Then all the following hold:

(a) $\forall \alpha_1, \dots, \alpha_n \in \{-1, 1\}$: $\begin{bmatrix} \alpha_1\mathbf{q}_1 & \dots & \alpha_n\mathbf{q}_n \end{bmatrix}$ is orthogonal;

(b) $\forall \alpha_1, \dots, \alpha_n \in \{-1, 1\}$: $\begin{bmatrix} \alpha_1\mathbf{r}_1^T \\ \vdots \\ \alpha_n\mathbf{r}_n^T \end{bmatrix}$ is orthogonal;

(c) the matrix $-Q$ is orthogonal.

Proposition 6.8.3

If $Q_1, Q_2 \in \mathbb{R}^{n \times n}$ are orthogonal, then so is their product $Q_1 Q_2$.

Proposition 6.8.4

Let $Q_1 \in \mathbb{R}^{m \times m}$ and $Q_2 \in \mathbb{R}^{n \times n}$ be orthogonal matrices. Then the $(m+n) \times (m+n)$ matrix

$$Q = \left[\begin{array}{c|c} Q_1 & O_{m \times n} \\ \hline O_{n \times m} & Q_2 \end{array}\right]$$

is an orthogonal matrix in $\mathbb{R}^{(m+n) \times (m+n)}$.

- Next, we discuss two particularly significant orthogonal matrices: the Householder matrix and the Givens matrix.

- Next, we discuss two particularly significant orthogonal matrices: the Householder matrix and the Givens matrix.
- In our discussion of the Householder matrix, we will need the following result.

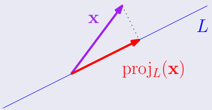
Corollary 6.6.4

Let $\mathbf{a}$ be a non-zero vector in $\mathbb{R}^n$. Then the standard matrix of orthogonal projection onto the line $L := \text{Span}(\mathbf{a})$ is the matrix

$$\mathbf{a}(\mathbf{a}^T\mathbf{a})^{-1}\mathbf{a}^T \quad = \quad \mathbf{a}(\mathbf{a}\cdot\mathbf{a})^{-1}\mathbf{a}^T \quad = \quad \frac{1}{\mathbf{a}\cdot\mathbf{a}}\mathbf{a}\mathbf{a}^T.$$

Consequently, for every vector $\mathbf{x} \in \mathbb{R}^n$, we have that

$$\mathbf{x}_L \quad = \quad \text{proj}_L(\mathbf{x}) \quad = \quad \frac{1}{\mathbf{a}\cdot\mathbf{a}}\mathbf{a}\mathbf{a}^T\mathbf{x}.$$



- Proof: Lecture Notes.

Definition

For a non-zero vector $\mathbf{a}$ in $\mathbb{R}^n$, the *Householder matrix* is the $n \times n$ matrix

$$H(\mathbf{a}) \;:=\; I_n - \frac{2}{\mathbf{a}^T\mathbf{a}}\mathbf{a}\mathbf{a}^T \;=\; I_n - \frac{2}{\mathbf{a}\cdot\mathbf{a}}\mathbf{a}\mathbf{a}^T.$$

Definition

For a non-zero vector $\mathbf{a}$ in $\mathbb{R}^n$, the *Householder matrix* is the $n \times n$ matrix

$$H(\mathbf{a}) \;:=\; I_n - \frac{2}{\mathbf{a}^T\mathbf{a}}\mathbf{a}\mathbf{a}^T \;=\; I_n - \frac{2}{\mathbf{a}\cdot\mathbf{a}}\mathbf{a}\mathbf{a}^T.$$

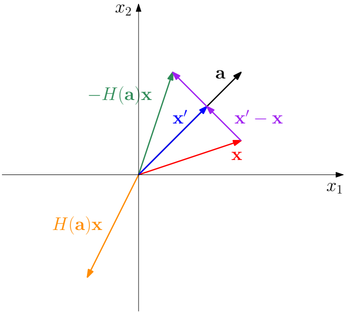
- To see that $H(\mathbf{a})$ is an orthogonal matrix, we compute:

$$\begin{aligned}
H(\mathbf{a})^T\, H(\mathbf{a}) &= \left(I_n - \tfrac{2}{\mathbf{a}\cdot\mathbf{a}}\mathbf{a}\mathbf{a}^T\right)^T \left(I_n - \tfrac{2}{\mathbf{a}\cdot\mathbf{a}}\mathbf{a}\mathbf{a}^T\right) \\
&= \left(I_n^T - \tfrac{2}{\mathbf{a}\cdot\mathbf{a}}(\mathbf{a}\mathbf{a}^T)^T\right) \left(I_n - \tfrac{2}{\mathbf{a}\cdot\mathbf{a}}\mathbf{a}\mathbf{a}^T\right) \\
&= \left(I_n - \tfrac{2}{\mathbf{a}\cdot\mathbf{a}}\mathbf{a}\mathbf{a}^T\right) \left(I_n - \tfrac{2}{\mathbf{a}\cdot\mathbf{a}}\mathbf{a}\mathbf{a}^T\right) \\
&= I_n - \tfrac{4}{\mathbf{a}\cdot\mathbf{a}}\mathbf{a}\mathbf{a}^T + \tfrac{4}{(\mathbf{a}\cdot\mathbf{a})^2}\mathbf{a}\,\underbrace{\mathbf{a}^T\mathbf{a}}_{=\mathbf{a}\cdot\mathbf{a}}\,\mathbf{a}^T \\
&= I_n - \tfrac{4}{\mathbf{a}\cdot\mathbf{a}}\mathbf{a}\mathbf{a}^T + \tfrac{4}{\mathbf{a}\cdot\mathbf{a}}\mathbf{a}\mathbf{a}^T \\
&= I_n.
\end{aligned}$$

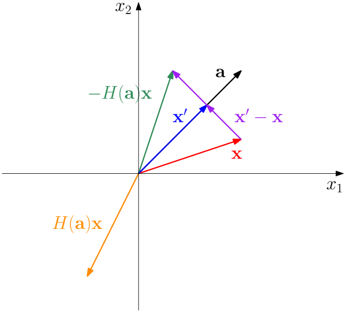
- Reminder: $H(\mathbf{a}) := I_n - \frac{2}{\mathbf{a}^T\mathbf{a}}\mathbf{a}\mathbf{a}^T = I_n - \frac{2}{\mathbf{a}\cdot\mathbf{a}}\mathbf{a}\mathbf{a}^T$.

- Reminder: $H(\mathbf{a}) := I_n - \frac{2}{\mathbf{a}^T\mathbf{a}}\mathbf{a}\mathbf{a}^T = I_n - \frac{2}{\mathbf{a}\cdot\mathbf{a}}\mathbf{a}\mathbf{a}^T$.
- if $\mathbf{x}$ is any vector in $\mathbb{R}^n$, and $\mathbf{x}'$ represents the orthogonal projection of $\mathbf{x}$ onto Span($\mathbf{a}$), then the reflection of $\mathbf{x}$ about the line Span($\mathbf{a}$) is given by

$$\begin{aligned}\mathbf{x} + 2(\mathbf{x}' - \mathbf{x}) = 2\mathbf{x}' - \mathbf{x} &= \frac{2}{\mathbf{a}\cdot\mathbf{a}}\mathbf{a}\mathbf{a}^T\mathbf{x} - I_n\mathbf{x} \\ &= (\frac{2}{\mathbf{a}\cdot\mathbf{a}}\mathbf{a}\mathbf{a}^T - I_n)\mathbf{x} \\ &= -H(\mathbf{a})\mathbf{x}.\end{aligned}$$

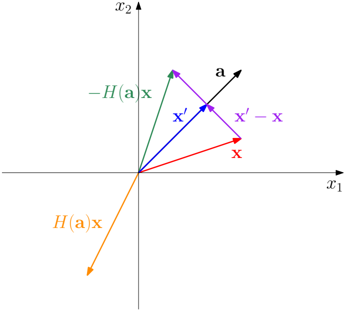


- Reminder: $H(\mathbf{a}) := I_n - \frac{2}{\mathbf{a}^T\mathbf{a}}\mathbf{a}\mathbf{a}^T = I_n - \frac{2}{\mathbf{a}\cdot\mathbf{a}}\mathbf{a}\mathbf{a}^T$.



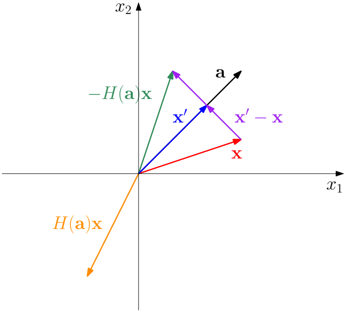
- Thus, $-H(\mathbf{a})$ is the standard matrix of reflection about the Span($\mathbf{a}$) line.

- Reminder: $H(\mathbf{a}) := I_n - \frac{2}{\mathbf{a}^T\mathbf{a}}\mathbf{a}\mathbf{a}^T = I_n - \frac{2}{\mathbf{a}\cdot\mathbf{a}}\mathbf{a}\mathbf{a}^T$.



- Thus, $-H(\mathbf{a})$ is the standard matrix of reflection about the Span($\mathbf{a}$) line.
- The Householder matrix $H(\mathbf{a})$ itself is the standard matrix of the linear operation that first reflects about the Span($\mathbf{a}$) line and then reflects about the origin.

- Reminder: $H(\mathbf{a}) := I_n - \frac{2}{\mathbf{a}^T\mathbf{a}}\mathbf{a}\mathbf{a}^T = I_n - \frac{2}{\mathbf{a}\cdot\mathbf{a}}\mathbf{a}\mathbf{a}^T$.



- **Remark:** Suppose that $\mathbf{a}$ is a non-zero vector in $\mathbb{R}^n$.
  - Then the standard matrix of reflection about the line $L := \text{Span}(\mathbf{a})$ in $\mathbb{R}^n$ is an orthogonal matrix.
  - Indeed, as we saw, the Householder matrix $H(\mathbf{a})$ is an orthogonal matrix.
  - By Proposition 6.8.2(c), it follows that $-H(\mathbf{a})$ is also an orthogonal matrix, and as we saw above, $-H(\mathbf{a})$ is the standard matrix of reflection about the line $L = \text{Span}(\mathbf{a})$ in $\mathbb{R}^n$.

- Given an integer $n \geq 2$, indices $i, j \in \{1, \dots, n\}$ such that $i < j$, and real numbers c and s such that $c^2 + s^2 = 1$, we define the *Givens matrix* $G_{i,j}(c, s)$ as follows:

$$
G_{i,j}(c,s) = \begin{array}{c} \\ \\ \\ \\ i \\ \\ \\ \\ j \\ \\ \\ \\ \end{array}
\overset{\begin{array}{ccccccccc} & & & i & & & j & & \end{array}}{\left[\begin{array}{ccccccccc}
1 & 0 & \dots\ 0 & 0 & 0 & \dots\ 0 & 0 & 0 & \dots\ 0 \\
0 & 1 & \dots\ 0 & 0 & 0 & \dots\ 0 & 0 & 0 & \dots\ 0 \\
\vdots & \vdots & \ddots\ \vdots & \vdots & \vdots & \ddots\ \vdots & \vdots & \vdots & \vdots \\
0 & 0 & \dots\ 1 & 0 & 0 & \dots\ 0 & 0 & 0 & \dots\ 0 \\
0 & 0 & \dots\ 0 & c & 0 & \dots\ 0 & -s & 0 & \dots\ 0 \\
0 & 0 & \dots\ 0 & 0 & 1 & \dots\ 0 & 0 & 0 & \dots\ 0 \\
\vdots & \vdots & \dots\ \vdots & \vdots & \vdots & \ddots\ \vdots & \vdots & \vdots & 0 \\
0 & 0 & \dots\ 0 & 0 & 0 & \dots\ 1 & 0 & 0 & \dots\ 0 \\
0 & 0 & \dots\ 0 & s & 0 & \dots\ 0 & c & 0 & \dots\ 0 \\
0 & 0 & \dots\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots\ \vdots \\
0 & 0 & \dots\ 0 & 0 & 0 & 0 & 0 & 0 & \dots\ 1
\end{array}\right]}
$$

- Given an integer $n \geq 2$, indices $i, j \in \{1, \ldots, n\}$ such that $i < j$, and real numbers c and s such that $c^2 + s^2 = 1$, we define the *Givens matrix* $G_{i,j}(c, s)$ as follows:

$$
\begin{array}{c} \\ \\ \\ \\ \\ G_{i,j}(c,s) \; = \end{array}
\begin{array}{c} \\ \\ \\ \\ i \\ \\ \\ \\ j \\ \\ \\ \\ \end{array}
\begin{array}{c}
\begin{array}{cccccccccc} & & & i & & & j & & & \end{array} \\
\begin{bmatrix}
1 & 0 & \ldots 0 & 0 & 0 & \ldots 0 & 0 & 0 & \ldots 0 \\
0 & 1 & \ldots 0 & 0 & 0 & \ldots 0 & 0 & 0 & \ldots 0 \\
\vdots & \vdots & \ddots \vdots & \vdots & \vdots & \ddots \vdots & \vdots & \vdots & \vdots \\
0 & 0 & \ldots 1 & 0 & 0 & \ldots 0 & 0 & 0 & \ldots 0 \\
0 & 0 & \ldots 0 & c & 0 & \ldots 0 & -s & 0 & \ldots 0 \\
0 & 0 & \ldots 0 & 0 & 1 & \ldots 0 & 0 & 0 & \ldots 0 \\
\vdots & \vdots & \ldots \vdots & \vdots & \vdots & \ddots \vdots & \vdots & \vdots & 0 \\
0 & 0 & \ldots 0 & 0 & 0 & \ldots 1 & 0 & 0 & \ldots 0 \\
0 & 0 & \ldots 0 & s & 0 & \ldots 0 & c & 0 & \ldots 0 \\
0 & 0 & \ldots 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \vdots \\
0 & 0 & \ldots 0 & 0 & 0 & 0 & 0 & 0 & \ldots 1
\end{bmatrix}
\end{array}
$$

- It is not hard to check that the columns of $G_{i,j}(c, s)$ form an orthonormal set of vectors in $\mathbb{R}^n$, and therefore (by Proposition 6.3.4) an orthonormal basis of $\mathbb{R}^n$.
- So, by Theorem 6.8.1, $G_{i,j}(c, s)$ is orthogonal.

- Let us now give a geometric interpretation of this matrix.

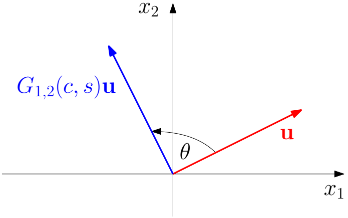
- Let us now give a geometric interpretation of this matrix.
- Since $c^2 + s^2 = 1$, we see that there exists a real number (angle in radians) θ such that $c = \cos\theta$ and $s = \sin\theta$.

- Let us now give a geometric interpretation of this matrix.
- Since $c^2 + s^2 = 1$, we see that there exists a real number (angle in radians) θ such that $c = \cos\theta$ and $s = \sin\theta$.
- With this set-up, we see that $G_{i,j}(c, s)$ represents rotation about the origin by angle θ in the $x_i x_j$-plane.

- Let us now give a geometric interpretation of this matrix.
- Since $c^2 + s^2 = 1$, we see that there exists a real number (angle in radians) θ such that $c = \cos\theta$ and $s = \sin\theta$.
- With this set-up, we see that $G_{i,j}(c, s)$ represents rotation about the origin by angle θ in the $x_i x_j$-plane.
- This is particularly easy to see in the case when $n = 2$. In that case, we have that

$$G_{1,2}(c,s) = \begin{bmatrix} c & -s \\ s & c \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix},$$

which is precisely the standard matrix of counterclockwise rotation about the origin by angle θ.



$c = \cos\theta$
$s = \sin\theta$

Theorem 6.8.5

Let $Q = \begin{bmatrix} q_{i,j} \end{bmatrix}_{n \times n}$ be an orthogonal matrix in $\mathbb{R}^{n \times n}$. Then:

- (a) for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $(Q\mathbf{x}) \cdot (Q\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$;
- (b) for all $\mathbf{x} \in \mathbb{R}^n$, $||Q\mathbf{x}|| = ||\mathbf{x}||$;
- (c) for all $i, j \in \{1, \dots, n\}$, $|q_{i,j}| \leq 1$.

- Proof: next slide.

Theorem 6.8.5

Let $Q = \left[\begin{array}{c} q_{i,j} \end{array} \right]_{n \times n}$ be an orthogonal matrix in $\mathbb{R}^{n \times n}$. Then:

(a) for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $(Q\mathbf{x}) \cdot (Q\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$;

(b) for all $\mathbf{x} \in \mathbb{R}^n$, $||Q\mathbf{x}|| = ||\mathbf{x}||$;

(c) for all $i, j \in \{1, \dots, n\}$, $|q_{i,j}| \leq 1$.

- Proof: next slide.
- **Remark:** By Theorem 6.8.5(b), multiplication by an orthogonal matrix (on the left) preserves vector length.

Theorem 6.8.5

Let $Q = \left[\begin{array}{c} q_{i,j} \end{array}\right]_{n\times n}$ be an orthogonal matrix in $\mathbb{R}^{n\times n}$. Then:

(a) for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $(Q\mathbf{x}) \cdot (Q\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$;

(b) for all $\mathbf{x} \in \mathbb{R}^n$, $||Q\mathbf{x}|| = ||\mathbf{x}||$;

(c) for all $i, j \in \{1, \dots, n\}$, $|q_{i,j}| \leq 1$.

- Proof: next slide.
- **Remark:** By Theorem 6.8.5(b), multiplication by an orthogonal matrix (on the left) preserves vector length.
  - On the other hand, recall that for non-zero vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, we have that $\mathbf{x} \cdot \mathbf{y} = ||\mathbf{x}|| \; ||\mathbf{y}|| \cos\theta$, where θ is the angle between $\mathbf{x}$ and $\mathbf{y}$.

Theorem 6.8.5

Let $Q = \begin{bmatrix} q_{i,j} \end{bmatrix}_{n \times n}$ be an orthogonal matrix in $\mathbb{R}^{n \times n}$. Then:

- (a) for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $(Q\mathbf{x}) \cdot (Q\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$;
- (b) for all $\mathbf{x} \in \mathbb{R}^n$, $||Q\mathbf{x}|| = ||\mathbf{x}||$;
- (c) for all $i, j \in \{1, \dots, n\}$, $|q_{i,j}| \leq 1$.

- Proof: next slide.
- **Remark:** By Theorem 6.8.5(b), multiplication by an orthogonal matrix (on the left) preserves vector length.
  - On the other hand, recall that for non-zero vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, we have that $\mathbf{x} \cdot \mathbf{y} = ||\mathbf{x}|| \, ||\mathbf{y}|| \cos \theta$, where θ is the angle between $\mathbf{x}$ and $\mathbf{y}$.
  - So, Theorem 6.8.5(a-b) implies that multiplication (on the left) by an orthogonal matrix preserves angles between non-zero vectors.

Theorem 6.8.5

Let $Q = \begin{bmatrix} q_{i,j} \end{bmatrix}_{n \times n}$ be an orthogonal matrix in $\mathbb{R}^{n \times n}$. Then:

a. for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $(Q\mathbf{x}) \cdot (Q\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$;
b. for all $\mathbf{x} \in \mathbb{R}^n$, $||Q\mathbf{x}|| = ||\mathbf{x}||$;
c. for all $i, j \in \{1, \dots, n\}$, $|q_{i,j}| \leq 1$.

Proof.

Theorem 6.8.5

Let $Q = \left[\ q_{i,j}\ \right]_{n\times n}$ be an orthogonal matrix in $\mathbb{R}^{n\times n}$. Then:

(a) for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $(Q\mathbf{x}) \cdot (Q\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$;

(b) for all $\mathbf{x} \in \mathbb{R}^n$, $||Q\mathbf{x}|| = ||\mathbf{x}||$;

(c) for all $i, j \in \{1, \dots, n\}$, $|q_{i,j}| \leq 1$.

Proof. (a) For $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, we have the following:

$$(Q\mathbf{x}) \cdot (Q\mathbf{y}) \;=\; (Q\mathbf{x})^T (Q\mathbf{x}) \;=\; \mathbf{x}^T \underbrace{Q^T Q}_{=I_n} \mathbf{y} \;=\; \mathbf{x}^T \mathbf{y} \;=\; \mathbf{x} \cdot \mathbf{y}.$$

Theorem 6.8.5

Let $Q = \begin{bmatrix} q_{i,j} \end{bmatrix}_{n \times n}$ be an orthogonal matrix in $\mathbb{R}^{n \times n}$. Then:

- (a) for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $(Q\mathbf{x}) \cdot (Q\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$;
- (b) for all $\mathbf{x} \in \mathbb{R}^n$, $||Q\mathbf{x}|| = ||\mathbf{x}||$;
- (c) for all $i, j \in \{1, \dots, n\}$, $|q_{i,j}| \leq 1$.

Proof. (a) For $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, we have the following:

$$(Q\mathbf{x}) \cdot (Q\mathbf{y}) \quad = \quad (Q\mathbf{x})^T (Q\mathbf{x}) \quad = \quad \mathbf{x}^T \underbrace{Q^T Q}_{= I_n} \mathbf{y} \quad = \quad \mathbf{x}^T \mathbf{y} \quad = \quad \mathbf{x} \cdot \mathbf{y}.$$

(b) For $\mathbf{x} \in \mathbb{R}^n$, we have the following:

$$||Q\mathbf{x}|| \quad = \quad \sqrt{(Q\mathbf{x}) \cdot (Q\mathbf{x})} \quad \overset{(a)}{=} \quad \sqrt{\mathbf{x} \cdot \mathbf{x}} \quad = \quad ||\mathbf{x}||.$$

Theorem 6.8.5

Let $Q = \begin{bmatrix} q_{i,j} \end{bmatrix}_{n \times n}$ be an orthogonal matrix in $\mathbb{R}^{n \times n}$. Then:

- (a) for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $(Q\mathbf{x}) \cdot (Q\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$;
- (b) for all $\mathbf{x} \in \mathbb{R}^n$, $||Q\mathbf{x}|| = ||\mathbf{x}||$;
- (c) for all $i, j \in \{1, \dots, n\}$, $|q_{i,j}| \leq 1$.

Proof. (a) For $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, we have the following:

$$(Q\mathbf{x}) \cdot (Q\mathbf{y}) = (Q\mathbf{x})^T (Q\mathbf{x}) = \mathbf{x}^T \underbrace{Q^T Q}_{= I_n} \mathbf{y} = \mathbf{x}^T \mathbf{y} = \mathbf{x} \cdot \mathbf{y}.$$

(b) For $\mathbf{x} \in \mathbb{R}^n$, we have the following:

$$||Q\mathbf{x}|| = \sqrt{(Q\mathbf{x}) \cdot (Q\mathbf{x})} \overset{(a)}{=} \sqrt{\mathbf{x} \cdot \mathbf{x}} = ||\mathbf{x}||.$$

(c) By Theorem 6.8.1, the columns of Q form an orthonormal basis. In particular, all columns of Q are unit vectors, and it follows that all entries of Q have absolute value at most 1. □

3. Scalar product, coordinate vectors, and matrices of linear functions

3. Scalar product, coordinate vectors, and matrices of linear functions

Proposition 6.9.1

Let V be a real or complex vector space, equipped with the scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $|| \cdot ||$, and let $\mathcal{B} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ be an **orthonormal** basis of V. Let $\cdot$ be the standard scalar product in $\mathbb{R}^n$ or $\mathbb{C}^n$ (depending on whether the vector space V is real or complex). Then for all $\mathbf{x}, \mathbf{y} \in V$, we have that

$$\langle \mathbf{x}, \mathbf{y} \rangle \ = \ \left[\ \mathbf{x} \ \right]_{\mathcal{B}} \cdot \left[\ \mathbf{y} \ \right]_{\mathcal{B}}.$$

- Proof: Lecture Notes.

Theorem 6.9.2

Let U and V be non-trivial, finite-dimensional **real** vector spaces. Assume that U is equipped with a scalar product $\langle \cdot, \cdot \rangle_U$ and the induced norm $|| \cdot ||_U$, and that V is equipped with a scalar product $\langle \cdot, \cdot \rangle_V$ and the induced norm $|| \cdot ||_V$. Let $\mathcal{B}_U = \{\mathbf{u}_1, \dots, \mathbf{u}_m\}$ and $\mathcal{B}_V = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ be **orthonormal** bases of U and V, respectively, and let $f : U \to V$ be a linear function. Then the following two statements are equivalent:

1. (i) the columns of the $n \times m$ matrix ${}_{\mathcal{B}_V}\left[\ f\ \right]_{\mathcal{B}_U}$ form an orthonormal set of vectors in $\mathbb{R}^n$ (with respect to the standard scalar product $\cdot$ and the induced norm $|| \cdot ||$);[a]
2. (ii) f preserves the scalar product, that is, for all vectors $\mathbf{x}, \mathbf{y} \in U$, we have that $\langle f(\mathbf{x}), f(\mathbf{y}) \rangle_V = \langle \mathbf{x}, \mathbf{y} \rangle_U$.

---

[a] However, despite Theorem 6.8.1, this does not necessarily mean that the matrix ${}_{\mathcal{B}_V}\left[\ f\ \right]_{\mathcal{B}_U}$ is orthogonal. This is because ${}_{\mathcal{B}_V}\left[\ f\ \right]_{\mathcal{B}_U}$ is an $n \times m$ matrix, and it is possible that $m \neq n$, in which case ${}_{\mathcal{B}_V}\left[\ f\ \right]_{\mathcal{B}_U}$ is not a square matrix. Only square matrices can be orthogonal!

Proof. Set ${}_{\mathcal{B}_V}\left[\ f\ \right]_{\mathcal{B}_U} = \left[\ \mathbf{c}_1 \quad \dots \quad \mathbf{c}_m\ \right]$.

Proof. Set ${}_{\mathcal{B}_V}\big[\ f\ \big]_{\mathcal{B}_U} = \big[\ \mathbf{c}_1 \quad \dots \quad \mathbf{c}_m\ \big]$. We observe that

$$
\begin{aligned}
\left({}_{\mathcal{B}_V}\big[\ f\ \big]_{\mathcal{B}_U}\right)^T\ {}_{\mathcal{B}_V}\big[\ f\ \big]_{\mathcal{B}_U} &= \begin{bmatrix} \mathbf{c}_1^T \\ \mathbf{c}_2^T \\ \vdots \\ \mathbf{c}_m^T \end{bmatrix} \begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 & \dots & \mathbf{c}_m \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{c}_1 \cdot \mathbf{c}_1 & \mathbf{c}_1 \cdot \mathbf{c}_2 & \dots & \mathbf{c}_1 \cdot \mathbf{c}_m \\ \mathbf{c}_2 \cdot \mathbf{c}_1 & \mathbf{c}_2 \cdot \mathbf{c}_2 & \dots & \mathbf{c}_2 \cdot \mathbf{c}_m \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{c}_m \cdot \mathbf{c}_1 & \mathbf{c}_m \cdot \mathbf{c}_2 & \dots & \mathbf{c}_m \cdot \mathbf{c}_m \end{bmatrix}.
\end{aligned}
$$

So, we see that (i) holds iff $\left({}_{\mathcal{B}_V}\big[\ f\ \big]_{\mathcal{B}_U}\right)^T\ {}_{\mathcal{B}_V}\big[\ f\ \big]_{\mathcal{B}_U} = I_m$.

Proof (cont.). Reminder: (i) holds iff $({}_{\mathcal{B}_V}[\ f\ ]_{\mathcal{B}_U})^T\ {}_{\mathcal{B}_V}[\ f\ ]_{\mathcal{B}_U} = I_m$.

Proof (cont.). Reminder: (i) holds iff $({}_{\mathcal{B}_V}[\ f\ ]_{\mathcal{B}_U})^T\ {}_{\mathcal{B}_V}[\ f\ ]_{\mathcal{B}_U} = I_m$.

Next, by Proposition 6.9.1, the following hold for all $\mathbf{x}, \mathbf{y} \in U$:

(1) $\langle \mathbf{x}, \mathbf{y} \rangle_U = [\ \mathbf{x}\ ]_{\mathcal{B}_U} \cdot [\ \mathbf{y}\ ]_{\mathcal{B}_U}$;

(2) $\big\langle f(\mathbf{x}), f(\mathbf{y}) \big\rangle_V = [\ f(\mathbf{x})\ ]_{\mathcal{B}_V} \cdot [\ f(\mathbf{y})\ ]_{\mathcal{B}_V}$.

Proof (cont.). Reminder: (i) holds iff $({}_{\mathcal{B}_V}[\ f\ ]_{\mathcal{B}_U})^T\ {}_{\mathcal{B}_V}[\ f\ ]_{\mathcal{B}_U} = I_m$.

Next, by Proposition 6.9.1, the following hold for all $\mathbf{x}, \mathbf{y} \in U$:

(1) $\langle \mathbf{x}, \mathbf{y} \rangle_U = [\ \mathbf{x}\ ]_{\mathcal{B}_U} \cdot [\ \mathbf{y}\ ]_{\mathcal{B}_U}$;

(2) $\big\langle f(\mathbf{x}), f(\mathbf{y}) \big\rangle_V = [\ f(\mathbf{x})\ ]_{\mathcal{B}_V} \cdot [\ f(\mathbf{y})\ ]_{\mathcal{B}_V}$.

Now, for all $\mathbf{x}, \mathbf{y} \in U$, we have that

$$\begin{aligned}
\big\langle f(\mathbf{x}), f(\mathbf{y}) \big\rangle_V &\overset{(2)}{=} [\ f(\mathbf{x})\ ]_{\mathcal{B}_V} \cdot [\ f(\mathbf{y})\ ]_{\mathcal{B}_V} \\
&= \big([\ f(\mathbf{x})\ ]_{\mathcal{B}_V}\big)^T\ [\ f(\mathbf{y})\ ]_{\mathcal{B}_V} \\
&= \Big({}_{\mathcal{B}_V}[\ f\ ]_{\mathcal{B}_U}\ [\ \mathbf{x}\ ]_{\mathcal{B}_U}\Big)^T \Big({}_{\mathcal{B}_V}[\ f\ ]_{\mathcal{B}_U}\ [\ \mathbf{y}\ ]_{\mathcal{B}_U}\Big) \\
&= \big([\ \mathbf{x}\ ]_{\mathcal{B}_U}\big)^T \big({}_{\mathcal{B}_V}[\ f\ ]_{\mathcal{B}_U}\big)^T\ {}_{\mathcal{B}_V}[\ f\ ]_{\mathcal{B}_U}\ [\ \mathbf{y}\ ]_{\mathcal{B}_U}.
\end{aligned}$$

Proof (continued). Suppose first that (i) holds. Then $({}_{\mathcal{B}_V}[\ f\ ]_{\mathcal{B}_U})^T\ {}_{\mathcal{B}_V}[\ f\ ]_{\mathcal{B}_U} = I_m$, and consequently, for all $\mathbf{x}, \mathbf{y} \in U$, we have that

$$\begin{aligned}
\big\langle f(\mathbf{x}), f(\mathbf{y}) \big\rangle_V &= \big([\ \mathbf{x}\ ]_{\mathcal{B}_U} \big)^T \underbrace{({}_{\mathcal{B}_V}[\ f\ ]_{\mathcal{B}_U})^T\ {}_{\mathcal{B}_V}[\ f\ ]_{\mathcal{B}_U}}_{=I_m}\ [\ \mathbf{y}\ ]_{\mathcal{B}_U} \\
&= \big([\ \mathbf{x}\ ]_{\mathcal{B}_U} \big)^T\ [\ \mathbf{y}\ ]_{\mathcal{B}_U} \\
&= [\ \mathbf{x}\ ]_{\mathcal{B}_U} \cdot [\ \mathbf{y}\ ]_{\mathcal{B}_U} \\
&\stackrel{(1)}{=} \langle \mathbf{x}, \mathbf{y} \rangle_U.
\end{aligned}$$

Thus, (ii) holds.

Proof (continued). Reminder: ${}_{\mathcal{B}_V}\big[\ f\ \big]_{\mathcal{B}_U} = \big[\ \mathbf{c}_1 \quad \dots \quad \mathbf{c}_m\ \big]$

Suppose now that (ii) holds. Then for all $i, j \in \{1, \dots, m\}$, we have that

$$\begin{aligned}
\mathbf{e}_i^m \cdot \mathbf{e}_j^m &= \big[\ \mathbf{u}_i\ \big]_{\mathcal{B}_U} \cdot \big[\ \mathbf{u}_j\ \big]_{\mathcal{B}_U} \\
&\overset{(1)}{=} \langle \mathbf{u}_i, \mathbf{u}_j \rangle_U \\
&\overset{(ii)}{=} \big\langle f(\mathbf{u}_i), f(\mathbf{u}_j) \big\rangle_V \\
&\overset{(2)}{=} \big[\ f(\mathbf{u}_i)\ \big]_{\mathcal{B}_V} \cdot \big[\ f(\mathbf{u}_j)\ \big]_{\mathcal{B}_V} \\
&= \big({}_{\mathcal{B}_V}\big[\ f\ \big]_{\mathcal{B}_U}\ \big[\ \mathbf{u}_i\ \big]_{\mathcal{B}_U}\big) \cdot \big({}_{\mathcal{B}_V}\big[\ f\ \big]_{\mathcal{B}_U}\ \big[\ \mathbf{u}_j\ \big]_{\mathcal{B}_U}\big) \\
&= \big({}_{\mathcal{B}_V}\big[\ f\ \big]_{\mathcal{B}_U}\ \mathbf{e}_i^m\big) \cdot \big({}_{\mathcal{B}_V}\big[\ f\ \big]_{\mathcal{B}_U}\ \mathbf{e}_j^m\big) \\
&= \mathbf{c}_i \cdot \mathbf{c}_j.
\end{aligned}$$

So, $\{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ is an orthonormal set of vectors in $\mathbb{R}^n$, that is, (i) holds. $\square$

Theorem 6.9.2

Let U and V be non-trivial, finite-dimensional **real** vector spaces. Assume that U is equipped with a scalar product $\langle \cdot, \cdot \rangle_U$ and the induced norm $|| \cdot ||_U$, and that V is equipped with a scalar product $\langle \cdot, \cdot \rangle_V$ and the induced norm $|| \cdot ||_V$. Let $\mathcal{B}_U = \{\mathbf{u}_1, \dots, \mathbf{u}_m\}$ and $\mathcal{B}_V = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ be **orthonormal** bases of U and V, respectively, and let $f : U \to V$ be a linear function. Then the following two statements are equivalent:

1. (i) the columns of the $n \times m$ matrix ${}_{\mathcal{B}_V}\left[\ f\ \right]_{\mathcal{B}_U}$ form an orthonormal set of vectors in $\mathbb{R}^n$ (with respect to the standard scalar product $\cdot$ and the induced norm $|| \cdot ||$);[a]
2. (ii) f preserves the scalar product, that is, for all vectors $\mathbf{x}, \mathbf{y} \in U$, we have that $\langle f(\mathbf{x}), f(\mathbf{y}) \rangle_V = \langle \mathbf{x}, \mathbf{y} \rangle_U$.

---

[a] However, despite Theorem 6.8.1, this does not necessarily mean that the matrix ${}_{\mathcal{B}_V}\left[\ f\ \right]_{\mathcal{B}_U}$ is orthogonal. This is because ${}_{\mathcal{B}_V}\left[\ f\ \right]_{\mathcal{B}_U}$ is an $n \times m$ matrix, and it is possible that $m \neq n$, in which case ${}_{\mathcal{B}_V}\left[\ f\ \right]_{\mathcal{B}_U}$ is not a square matrix. Only square matrices can be orthogonal!