LINEAR ALGEBRA -II

B V Rajarama Bhat

Indian Statistical Institute, Bangalore

References

- Basic references:

References

- Basic references:
- Linear Algebra, A. Ramachandra Rao and P. Bhimasankaram.

References

- Basic references:
- Linear Algebra, A. Ramachandra Rao and P. Bhimasankaram.
- Algebra, Michael Artin.

References

- Basic references:
- Linear Algebra, A. Ramachandra Rao and P. Bhimasankaram.
- Algebra, Michael Artin.
- Linear Algebra, Henry Helson.

Lecture 2: Permutations -2

- Recall:

Lecture 2: Permutations -2

- Recall:
- Definition 1.1: Let S be a finite set. Then a bijective function $\sigma : S \to S$ is said to be a permutation of S.

Lecture 2: Permutations -2

- Recall:
- Definition 1.1: Let S be a finite set. Then a bijective function $\sigma : S \to S$ is said to be a permutation of S.
- Example 1.4: Suppose $S = \{1, 2, \ldots 7\}$. Consider the permutation:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 5 & 7 & 4 & 2 & 6 & 1 \end{pmatrix}$$

Lecture 2: Permutations -2

- Recall:
- Definition 1.1: Let S be a finite set. Then a bijective function $\sigma : S \to S$ is said to be a permutation of S.
- Example 1.4: Suppose $S = \{1, 2, \ldots 7\}$. Consider the permutation:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 5 & 7 & 4 & 2 & 6 & 1 \end{pmatrix}$$

- We see $1 \dashrightarrow 3 \dashrightarrow 7 \dashrightarrow 1$. This we call as a cycle. It is a cycle of length 3.

- Recall:
- Definition 1.1: Let S be a finite set. Then a bijective function $\sigma : S \to S$ is said to be a permutation of S.
- Example 1.4: Suppose $S = \{1, 2, \ldots 7\}$. Consider the permutation:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 5 & 7 & 4 & 2 & 6 & 1 \end{pmatrix}$$

- We see $1 \dashrightarrow 3 \dashrightarrow 7 \dashrightarrow 1$. This we call as a cycle. It is a cycle of length 3.
- This permutation also has $2 \dashrightarrow 5 \dashrightarrow 2$, a cycle of length 2.

Lecture 2: Permutations -2

- Recall:
- Definition 1.1: Let S be a finite set. Then a bijective function $\sigma : S \to S$ is said to be a permutation of S.
- Example 1.4: Suppose $S = \{1, 2, \dots 7\}$. Consider the permutation:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 5 & 7 & 4 & 2 & 6 & 1 \end{pmatrix}$$

- We see $1 \dashrightarrow 3 \dashrightarrow 7 \dashrightarrow 1$. This we call as a cycle. It is a cycle of length 3.
- This permutation also has $2 \dashrightarrow 5 \dashrightarrow 2$, a cycle of length 2.
- It also has $4 \dashrightarrow 4$ and $6 \dashrightarrow 6$, cycles of length 1.

Lecture 2: Permutations -2

- Recall:
- Definition 1.1: Let S be a finite set. Then a bijective function $\sigma : S \to S$ is said to be a permutation of S.
- Example 1.4: Suppose $S = \{1, 2, \dots 7\}$. Consider the permutation:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 5 & 7 & 4 & 2 & 6 & 1 \end{pmatrix}$$

- We see $1 \dashrightarrow 3 \dashrightarrow 7 \dashrightarrow 1$. This we call as a cycle. It is a cycle of length 3.
- This permutation also has $2 \dashrightarrow 5 \dashrightarrow 2$, a cycle of length 2.
- It also has $4 \dashrightarrow 4$ and $6 \dashrightarrow 6$, cycles of length 1.
- For distinct $k_1, k_2, \dots, k_r$ in $\{1, 2, \dots, n\}$ (with $r \in \mathbb{N}$) we denote the cycle $k_1 \dashrightarrow k_2 \dashrightarrow \cdots \dashrightarrow k_r \dashrightarrow k_1$ simply as $(k_1, k_2, \dots, k_r)$.

Product of cycles theorem

- Theorem 1.7: Let $S = \{1, 2, \ldots, n\}$ for some $n \in \mathbb{N}$. Suppose σ is a permutation of S. Then S decomposes uniquely as a product of cycles.

Product of cycles theorem

- Theorem 1.7: Let $S = \{1, 2, \dots, n\}$ for some $n \in \mathbb{N}$. Suppose σ is a permutation of S. Then S decomposes uniquely as a product of cycles.
- We may write down a permutation by listing the cycles it has.

Product of cycles theorem

- Theorem 1.7: Let $S = \{1, 2, \ldots, n\}$ for some $n \in \mathbb{N}$. Suppose σ is a permutation of S. Then S decomposes uniquely as a product of cycles.
- We may write down a permutation by listing the cycles it has.
- For instance, the permutation of Example 1.4, is written as $(1, 3, 7)(2, 5)(4)(6)$.

Signature of a permutation

- Definition 1.8: Let $S = \{1, 2, \ldots, n\}$ for some $n \in \mathbb{N}$ and let σ be a permutation of S. Then the signature of σ is defined as the number

$$\epsilon(\sigma) = (-1)^{n-p}$$

where p is the number of cycles (including cycles of length 1) in the cycle decomposition of σ.

Signature of a permutation

- Definition 1.8: Let $S = \{1, 2, \dots, n\}$ for some $n \in \mathbb{N}$ and let σ be a permutation of S. Then the signature of σ is defined as the number
$$\epsilon(\sigma) = (-1)^{n-p}$$
where p is the number of cycles (including cycles of length 1) in the cycle decomposition of σ.
- For instance for the permutation σ of Example 1.4, $\epsilon(\sigma) = (-1)^{7-4} = (-1)^3 = -1$.

Signature of a permutation

- Definition 1.8: Let $S = \{1, 2, \ldots, n\}$ for some $n \in \mathbb{N}$ and let σ be a permutation of S. Then the signature of σ is defined as the number
$$\epsilon(\sigma) = (-1)^{n-p}$$
where p is the number of cycles (including cycles of length 1) in the cycle decomposition of σ.
- For instance for the permutation σ of Example 1.4, $\epsilon(\sigma) = (-1)^{7-4} = (-1)^3 = -1$.
- Note that the signature of identity permutation is always 1.

Signature of a permutation

- Definition 1.8: Let $S = \{1, 2, \dots, n\}$ for some $n \in \mathbb{N}$ and let σ be a permutation of S. Then the signature of σ is defined as the number
$$\epsilon(\sigma) = (-1)^{n-p}$$
where p is the number of cycles (including cycles of length 1) in the cycle decomposition of σ.
- For instance for the permutation σ of Example 1.4, $\epsilon(\sigma) = (-1)^{7-4} = (-1)^3 = -1$.
- Note that the signature of identity permutation is always 1.
- A cycle $(k_1, k_2, \dots, k_r)$ can be identified with the permutation σ defined by
$$\sigma(k_1) = k_2, \sigma(k_2) = k_3, \dots, \sigma(k_r) = k_1$$
and $\sigma(j) = j$ for $j \notin \{k_1, k_2, \dots, k_r\}$.

Signature of a permutation

- Definition 1.8: Let $S = \{1, 2, \ldots, n\}$ for some $n \in \mathbb{N}$ and let σ be a permutation of S. Then the signature of σ is defined as the number
$$\epsilon(\sigma) = (-1)^{n-p}$$
where p is the number of cycles (including cycles of length 1) in the cycle decomposition of σ.
- For instance for the permutation σ of Example 1.4, $\epsilon(\sigma) = (-1)^{7-4} = (-1)^3 = -1$.
- Note that the signature of identity permutation is always 1.
- A cycle $(k_1, k_2, \ldots, k_r)$ can be identified with the permutation σ defined by
$$\sigma(k_1) = k_2, \sigma(k_2) = k_3, \ldots, \sigma(k_r) = k_1$$
and $\sigma(j) = j$ for $j \notin \{k_1, k_2, \ldots, k_r\}$.
- Therefore the signature of a cycle is defined as $(k_1, k_2, \ldots, k_r) = (-1)^{n-(1+(n-r))} = (-1)^{r-1}$.

Continuation

- Cycles of length two are known as transpositions. We see that transpositions have signature (-1).

Continuation

- Cycles of length two are known as transpositions. We see that transpositions have signature (-1).
- Permutations with signature $(+1)$ are known as even permutations and those with signature (-1) are known as odd permutations.

Permutations as products of transpositions

- Let $\sigma = (k_1, k_2, \ldots, k_r)$ be a cycle on $S = \{1, 2, \ldots, n\}$. This means that $k_1, k_2, \ldots, k_r$ are distinct elements in S and we are looking at the permutation:

$$k_1 \dashrightarrow k_2 \dashrightarrow \cdots k_{r-1} \dashrightarrow k_r \dashrightarrow k_1.$$

Permutations as products of transpositions

- Let $\sigma = (k_1, k_2, \ldots, k_r)$ be a cycle on $S = \{1, 2, \ldots, n\}$. This means that $k_1, k_2, \ldots, k_r$ are distinct elements in S and we are looking at the permutation:

$$k_1 \dashrightarrow k_2 \dashrightarrow \cdots k_{r-1} \dashrightarrow k_r \dashrightarrow k_1.$$

- To arrive at this permutation, we may first interchange k_1, k_2 and then interchange k_1, k_3, and then k_1, k_4 and so on and finally k_1 and k_r.

Permutations as products of transpositions

- Let $\sigma = (k_1, k_2, \ldots, k_r)$ be a cycle on $S = \{1, 2, \ldots, n\}$. This means that $k_1, k_2, \ldots, k_r$ are distinct elements in S and we are looking at the permutation:

$$k_1 \dashrightarrow k_2 \dashrightarrow \cdots k_{r-1} \dashrightarrow k_r \dashrightarrow k_1.$$

- To arrive at this permutation, we may first interchange k_1, k_2 and then interchange k_1, k_3, and then k_1, k_4 and so on and finally k_1 and k_r.
- In other words, if $\tau_{i,j}$ is the transposition between i and j, then

$$\sigma = \tau_{k_1,k_r} \circ \cdots \tau_{k_1,k_3} \circ \tau_{k_1,k_2}.$$

Permutations as products of transpositions

- Let $\sigma = (k_1, k_2, \ldots, k_r)$ be a cycle on $S = \{1, 2, \ldots, n\}$. This means that $k_1, k_2, \ldots, k_r$ are distinct elements in S and we are looking at the permutation:

$$k_1 \dashrightarrow k_2 \dashrightarrow \cdots k_{r-1} \dashrightarrow k_r \dashrightarrow k_1.$$

- To arrive at this permutation, we may first interchange k_1, k_2 and then interchange k_1, k_3, and then k_1, k_4 and so on and finally k_1 and k_r.
- In other words, if $\tau_{i,j}$ is the transposition between i and j, then

$$\sigma = \tau_{k_1,k_r} \circ \cdots \tau_{k_1,k_3} \circ \tau_{k_1,k_2}.$$

- Since every permutation is a product of disjoint cycles it follows that every permutation is a product of transpositions. In other words, given any permutation σ there exist transpositions $\tau_1, \tau_2, \ldots, \tau_k$ (for some $k \in \{0, 1, \ldots\}$) such that

$$\sigma = \tau_k \circ \tau_{k-1} \circ \cdots \circ \tau_2 \circ \tau_1.$$

Permutations as products of transpositions

- Let $\sigma = (k_1, k_2, \ldots, k_r)$ be a cycle on $S = \{1, 2, \ldots, n\}$. This means that $k_1, k_2, \ldots, k_r$ are distinct elements in S and we are looking at the permutation:

$$k_1 \dashrightarrow k_2 \dashrightarrow \cdots k_{r-1} \dashrightarrow k_r \dashrightarrow k_1.$$

- To arrive at this permutation, we may first interchange k_1, k_2 and then interchange k_1, k_3, and then k_1, k_4 and so on and finally k_1 and k_r.
- In other words, if $\tau_{i,j}$ is the transposition between i and j, then

$$\sigma = \tau_{k_1,k_r} \circ \cdots \tau_{k_1,k_3} \circ \tau_{k_1,k_2}.$$

- Since every permutation is a product of disjoint cycles it follows that every permutation is a product of transpositions. In other words, given any permutation σ there exist transpositions $\tau_1, \tau_2, \ldots, \tau_k$ (for some $k \in \{0, 1, \ldots\}$) such that

$$\sigma = \tau_k \circ \tau_{k-1} \circ \cdots \circ \tau_2 \circ \tau_1.$$

- Note that this factorization is not unique.

Homomorphism property of the signature

- Theorem 2.1: Let $S = \{1, 2, \ldots, n\}$ for some $n \in \mathbb{N}$. Suppose σ, τ are two permutations of S. Then

$$\epsilon(\tau \circ \sigma) = \epsilon(\tau).\epsilon(\sigma).$$

Homomorphism property of the signature

- Theorem 2.1: Let $S = \{1, 2, \ldots, n\}$ for some $n \in \mathbb{N}$. Suppose σ, τ are two permutations of S. Then

$$\epsilon(\tau \circ \sigma) = \epsilon(\tau).\epsilon(\sigma).$$

- Proof. We first prove the theorem when τ is a transposition. So let τ be transposition of two distinct indices in S.

Homomorphism property of the signature

- Theorem 2.1: Let $S = \{1, 2, \ldots, n\}$ for some $n \in \mathbb{N}$. Suppose σ, τ are two permutations of S. Then

$$\epsilon(\tau \circ \sigma) = \epsilon(\tau).\epsilon(\sigma).$$

- Proof. We first prove the theorem when τ is a transposition. So let τ be transposition of two distinct indices in S.
- Let

$$\sigma = (k_{11}, k_{12}, \ldots, k_{1r_1})(k_{21}, k_{22}, \ldots, k_{2r_2}) \cdots (k_{p1}, k_{p2}, \ldots, k_{pr_p})$$

be the cycle decomposition of σ.

Homomorphism property of the signature

- Theorem 2.1: Let $S = \{1, 2, \ldots, n\}$ for some $n \in \mathbb{N}$. Suppose σ, τ are two permutations of S. Then

$$\epsilon(\tau \circ \sigma) = \epsilon(\tau).\epsilon(\sigma).$$

- Proof. We first prove the theorem when τ is a transposition. So let τ be transposition of two distinct indices in S.
- Let

$$\sigma = (k_{11}, k_{12}, \ldots, k_{1r_1})(k_{21}, k_{22}, \ldots, k_{2r_2}) \cdots (k_{p1}, k_{p2}, \ldots, k_{pr_p})$$

be the cycle decomposition of σ.
- By the definition of ϵ, $\epsilon(\sigma) = (-1)^{n-p}$ and we have also seen that $\epsilon(\tau) = -1$.

Homomorphism property of the signature

- Theorem 2.1: Let $S = \{1, 2, \ldots, n\}$ for some $n \in \mathbb{N}$. Suppose σ, τ are two permutations of S. Then

$$\epsilon(\tau \circ \sigma) = \epsilon(\tau).\epsilon(\sigma).$$

- Proof. We first prove the theorem when τ is a transposition. So let τ be transposition of two distinct indices in S.
- Let

$$\sigma = (k_{11}, k_{12}, \ldots, k_{1r_1})(k_{21}, k_{22}, \ldots, k_{2r_2}) \cdots (k_{p1}, k_{p2}, \ldots, k_{pr_p})$$

  be the cycle decomposition of σ.
- By the definition of ϵ, $\epsilon(\sigma) = (-1)^{n-p}$ and we have also seen that $\epsilon(\tau) = -1$.
- Therefore our aim is to show that $\epsilon(\tau \circ \sigma) = (-1)^{n-p+1}$.

Continuation

- Now there are two possibilities, the transposition τ could be of two indices of same cycle of σ or could be of different cycles of σ.

Continuation

- Now there are two possibilities, the transposition τ could be of two indices of same cycle of σ or could be of different cycles of σ.
- Case (i): Suppose τ is transposition of k_{1i} and k_{1j}.

Continuation

- Now there are two possibilities, the transposition τ could be of two indices of same cycle of σ or could be of different cycles of σ.
- Case (i): Suppose τ is transposition of k_{1i} and k_{1j}.
- With out loss of generality we may take $i < j$.

Continuation

- Now there are two possibilities, the transposition τ could be of two indices of same cycle of σ or could be of different cycles of σ.
- Case (i): Suppose τ is transposition of k_{1i} and k_{1j}.
- With out loss of generality we may take $i < j$.
- Then $\tau \circ \sigma$ acting on $\{k_{11}, k_{12}, \ldots, k_{1r_1}\}$ has two cycles, namely

$$(k_{11}, \ldots, k_{1(i-1)}, k_{1j}, k_{1(j+1)}, \ldots, k_{1r_1})$$

and

$$(k_{1i}, k_{1(i+1)}, \ldots, k_{1(j-1)}).$$

Continuation

- Case (ii): Suppose τ is transposition of k_{1i} and k_{2j}.

Continuation

- Case (ii): Suppose τ is transposition of k_{1i} and k_{2j}.
- Then $\tau \circ \sigma$ acting on these two cycles gives a single cycle:

$$(k_{11}, \ldots, k_{1(i-1)}, k_{2j}, k_{2(j+1)}, k_{2r_2}, k_{21},$$

$$\ldots, k_{2(j-1)}, k_{1i}, k_{1(i+1)}, \ldots, k_{1r_1}).$$

Continuation

- Case (ii): Suppose τ is transposition of k_{1i} and k_{2j}.
- Then $\tau \circ \sigma$ acting on these two cycles gives a single cycle:

$$(k_{11}, \ldots, k_{1(i-1)}, k_{2j}, k_{2(j+1)}, k_{2r_2}, k_{21},$$

$$\ldots, k_{2(j-1)}, k_{1i}, k_{1(i+1)}, \ldots, k_{1r_1}).$$

- In other words in both cases the number of cycles changes by 1 (either $+1$ or -1).

Continuation

- Case (ii): Suppose τ is transposition of k_{1i} and k_{2j}.
- Then $\tau \circ \sigma$ acting on these two cycles gives a single cycle:

$$(k_{11}, \ldots, k_{1(i-1)}, k_{2j}, k_{2(j+1)}, k_{2r_2}, k_{21},$$

$$\ldots, k_{2(j-1)}, k_{1i}, k_{1(i+1)}, \ldots, k_{1r_1}).$$

- In other words in both cases the number of cycles changes by 1 (either $+1$ or -1).
- Caution: We have written the proof above essentially assuming that $r_1, r_2 \geq 2$, as we have indices such as $(i-1), (j-1)$ etc. You should verify that the result is true for all r_1, r_2.

Continuation

- Case (ii): Suppose τ is transposition of k_{1i} and k_{2j}.
- Then $\tau \circ \sigma$ acting on these two cycles gives a single cycle:

$$(k_{11}, \ldots, k_{1(i-1)}, k_{2j}, k_{2(j+1)}, k_{2r_2}, k_{21},$$

$$\ldots, k_{2(j-1)}, k_{1i}, k_{1(i+1)}, \ldots, k_{1r_1}).$$

- In other words in both cases the number of cycles changes by 1 (either $+1$ or -1).
- Caution: We have written the proof above essentially assuming that $r_1, r_2 \geq 2$, as we have indices such as $(i-1), (j-1)$ etc. You should verify that the result is true for all r_1, r_2.
- This proves that $\epsilon(\tau \circ \sigma) = \epsilon(\tau).\epsilon(\sigma)$ whenever τ is a transposition.

Continuation

- Case (ii): Suppose τ is transposition of k_{1i} and k_{2j}.
- Then $\tau \circ \sigma$ acting on these two cycles gives a single cycle:

$$(k_{11}, \dots, k_{1(i-1)}, k_{2j}, k_{2(j+1)}, k_{2r_2}, k_{21},$$

$$\dots, k_{2(j-1)}, k_{1i}, k_{1(i+1)}, \dots, k_{1r_1}).$$

- In other words in both cases the number of cycles changes by 1 (either $+1$ or -1).
- Caution: We have written the proof above essentially assuming that $r_1, r_2 \geq 2$, as we have indices such as $(i-1), (j-1)$ etc. You should verify that the result is true for all r_1, r_2.
- This proves that $\epsilon(\tau \circ \sigma) = \epsilon(\tau).\epsilon(\sigma)$ whenever τ is a transposition.
- Since every permutation is a product of transpositions, by mathematical induction we get $\epsilon(\tau \circ \sigma)$ for every τ, σ.

Consequences

- Corollary 2.2: If a permutation $\tau = \tau_1 \circ \tau_2 \circ \cdots \circ \tau_k$, where $\tau_1, \ldots, \tau_k$ are transpositions then

$$\epsilon(\tau) = (-1)^k.$$

Consequences

- Corollary 2.2: If a permutation $\tau = \tau_1 \circ \tau_2 \circ \cdots \circ \tau_k$, where $\tau_1, \ldots, \tau_k$ are transpositions then
$$\epsilon(\tau) = (-1)^k.$$
- Proof. This is clear from the previous theorem and mathematical induction.

Consequences

- Corollary 2.2: If a permutation $\tau = \tau_1 \circ \tau_2 \circ \cdots \circ \tau_k$, where $\tau_1, \ldots, \tau_k$ are transpositions then
$$\epsilon(\tau) = (-1)^k.$$
- Proof. This is clear from the previous theorem and mathematical induction.
- Corollary 2.3: If a permutation $\tau = \tau_1 \circ \tau_2 \circ \cdots \circ \tau_k = \sigma_1 \circ \sigma_2 \circ \cdots \circ \sigma_l$, where $\tau_1, \tau_2, \ldots, \tau_k, \sigma_1, \sigma_2, \ldots, \sigma_l$ are transpositions, then $k - l$ is even. In particular, k is odd/even if and only if l is odd/even.

Consequences

- Corollary 2.2: If a permutation $\tau = \tau_1 \circ \tau_2 \circ \cdots \circ \tau_k$, where $\tau_1, \ldots, \tau_k$ are transpositions then
$$\epsilon(\tau) = (-1)^k.$$
- Proof. This is clear from the previous theorem and mathematical induction.
- Corollary 2.3: If a permutation $\tau = \tau_1 \circ \tau_2 \circ \cdots \circ \tau_k = \sigma_1 \circ \sigma_2 \circ \cdots \circ \sigma_l$, where $\tau_1, \tau_2, \ldots, \tau_k, \sigma_1, \sigma_2, \ldots, \sigma_l$ are transpositions, then $k - l$ is even. In particular, k is odd/even if and only if l is odd/even.
- Recall that a permutation σ is called even if $\epsilon(\sigma) = 1$ and is called odd if $\epsilon(\sigma) = -1$. These results tell us that composition of two permutations is even if and only if either both of them are even or both of them are odd.

Consequences

- Corollary 2.2: If a permutation $\tau = \tau_1 \circ \tau_2 \circ \cdots \circ \tau_k$, where $\tau_1, \ldots, \tau_k$ are transpositions then
$$\epsilon(\tau) = (-1)^k.$$
- Proof. This is clear from the previous theorem and mathematical induction.
- Corollary 2.3: If a permutation $\tau = \tau_1 \circ \tau_2 \circ \cdots \circ \tau_k = \sigma_1 \circ \sigma_2 \circ \cdots \circ \sigma_l$, where $\tau_1, \tau_2, \ldots, \tau_k, \sigma_1, \sigma_2, \ldots, \sigma_l$ are transpositions, then $k - l$ is even. In particular, k is odd/even if and only if l is odd/even.
- Recall that a permutation σ is called even if $\epsilon(\sigma) = 1$ and is called odd if $\epsilon(\sigma) = -1$. These results tell us that composition of two permutations is even if and only if either both of them are even or both of them are odd.
- In the Fifteen Puzzle game, one can see that moves of the game do not change the signature (It is an invariant.). This proves why it is not possible to reach the natural permutation on starting from (15, 14).

Permutation matrices

- Definition 2.4: Fix $n \in \mathbb{N}$ and let σ be a permutation of $\{1, 2, \ldots, n\}$. Then the $n \times n$ matrix P^σ defined by

$$p_{ij}^{\sigma} = \begin{cases} 1 & \text{if } \; i = \sigma(j) \\ 0 & \text{otherwise.} \end{cases}$$

is called the permutation matrix associated with the permutation σ. Note that every row or column of P^σ has exactly one non-zero entry which is 1.

Permutation matrices

- Definition 2.4: Fix $n \in \mathbb{N}$ and let σ be a permutation of $\{1, 2, \ldots, n\}$. Then the $n \times n$ matrix P^σ defined by

$$p_{ij}^\sigma = \begin{cases} 1 & \text{if } \ i = \sigma(j) \\ 0 & \text{otherwise.} \end{cases}$$

  is called the permutation matrix associated with the permutation σ. Note that every row or column of P^σ has exactly one non-zero entry which is 1.

Permutation matrices

- Definition 2.4: Fix $n \in \mathbb{N}$ and let σ be a permutation of $\{1, 2, \ldots, n\}$. Then the $n \times n$ matrix P^σ defined by

$$p^\sigma_{ij} = \begin{cases} 1 & \text{if } \ i = \sigma(j) \\ 0 & \text{otherwise.} \end{cases}$$

  is called the permutation matrix associated with the permutation σ. Note that every row or column of P^σ has exactly one non-zero entry which is 1.
- We also consider the matrix P^σ as the linear transformation $x \mapsto P^\sigma x$ on $\mathbb{R}^n$. More explicitly, if $x \in \mathbb{R}^n$ has the expansion $x = \sum_{j=1}^n x_j e_j$ in the standard basis $\{e_1, e_2, \ldots, e_n\}$,

Permutation matrices

- Definition 2.4: Fix $n \in \mathbb{N}$ and let σ be a permutation of $\{1, 2, \ldots, n\}$. Then the $n \times n$ matrix P^σ defined by

$$p_{ij}^{\sigma} = \begin{cases} 1 & \text{if } \; i = \sigma(j) \\ 0 & \text{otherwise.} \end{cases}$$

  is called the permutation matrix associated with the permutation σ. Note that every row or column of P^σ has exactly one non-zero entry which is 1.
- We also consider the matrix P^σ as the linear transformation $x \mapsto P^\sigma x$ on $\mathbb{R}^n$. More explicitly, if $x \in \mathbb{R}^n$ has the expansion $x = \sum_{j=1}^{n} x_j e_j$ in the standard basis $\{e_1, e_2, \ldots, e_n\}$,
-

$$(P^\sigma x)_i = \sum_j p_{ij}^{\sigma} x_j = x_{\sigma^{-1}(i)}.$$

Permutation matrices

- Definition 2.4: Fix $n \in \mathbb{N}$ and let σ be a permutation of $\{1, 2, \ldots, n\}$. Then the $n \times n$ matrix P^σ defined by

$$p_{ij}^{\sigma} = \begin{cases} 1 & \text{if } \; i = \sigma(j) \\ 0 & \text{otherwise.} \end{cases}$$

  is called the permutation matrix associated with the permutation σ. Note that every row or column of P^σ has exactly one non-zero entry which is 1.
- We also consider the matrix P^σ as the linear transformation $x \mapsto P^\sigma x$ on $\mathbb{R}^n$. More explicitly, if $x \in \mathbb{R}^n$ has the expansion $x = \sum_{j=1}^{n} x_j e_j$ in the standard basis $\{e_1, e_2, \ldots, e_n\}$,
-
$$(P^\sigma x)_i = \sum_j p_{ij}^{\sigma} x_j = x_{\sigma^{-1}(i)}.$$
- Note that $P^\sigma e_j = e_{\sigma(j)}$. Therefore P^σ just permutes the basis elements $e_1, e_2, \ldots, e_n$, sending e_j to $e_{\sigma(j)}$. Hence for any two permutations σ, τ, $P^{\tau \circ \sigma} = P^\tau . P^\sigma$.

Stochastic and doubly stochastic matrices

- Definition 2.5: A matrix $D = [d_{ij}]_{1 \leq i,j \leq n}$ is said to be stochastic if

Stochastic and doubly stochastic matrices

- Definition 2.5: A matrix $D = [d_{ij}]_{1 \leq i,j \leq n}$ is said to be stochastic if
- (i) $d_{ij} \geq 0$ for all $1 \leq i, j \leq n$ (entries are non-negative);

Stochastic and doubly stochastic matrices

- Definition 2.5: A matrix $D = [d_{ij}]_{1\le i,j\le n}$ is said to be stochastic if
- (i) $d_{ij} \geq 0$ for all $1 \leq i, j \leq n$ (entries are non-negative);
- (ii) $\sum_{j=1}^{n} d_{ij} = 1$ for every i (row sums of the matrix are equal to 1.)

Stochastic and doubly stochastic matrices

- Definition 2.5: A matrix $D = [d_{ij}]_{1 \leq i,j \leq n}$ is said to be stochastic if
- (i) $d_{ij} \geq 0$ for all $1 \leq i, j \leq n$ (entries are non-negative);
- (ii) $\sum_{j=1}^{n} d_{ij} = 1$ for every i (row sums of the matrix are equal to 1.)
- A stochastic matrix is said to be doubly stochastic if it also satisfies

Stochastic and doubly stochastic matrices

- Definition 2.5: A matrix $D = [d_{ij}]_{1 \leq i,j \leq n}$ is said to be stochastic if
- (i) $d_{ij} \geq 0$ for all $1 \leq i, j \leq n$ (entries are non-negative);
- (ii) $\sum_{j=1}^{n} d_{ij} = 1$ for every i (row sums of the matrix are equal to 1.)
- A stochastic matrix is said to be doubly stochastic if it also satisfies
- (iii) $\sum_{i=1}^{n} d_{ij} = 1$ (column sums are also equal to 1.)

Stochastic and doubly stochastic matrices

- Definition 2.5: A matrix $D = [d_{ij}]_{1 \leq i,j \leq n}$ is said to be stochastic if
- (i) $d_{ij} \geq 0$ for all $1 \leq i, j \leq n$ (entries are non-negative);
- (ii) $\sum_{j=1}^{n} d_{ij} = 1$ for every i (row sums of the matrix are equal to 1.)
- A stochastic matrix is said to be doubly stochastic if it also satisfies
- (iii) $\sum_{i=1}^{n} d_{ij} = 1$ (column sums are also equal to 1.)
- Clearly all permutation matrices are doubly stochastic matrices.

Extreme points of convex sets

- Let $\mathcal{D}$ be the set of all $n \times n$ doubly stochastic matrices.

Extreme points of convex sets

- Let $\mathcal{D}$ be the set of all $n \times n$ doubly stochastic matrices.
- Then $\mathcal{D}$ is a convex set, that is for $D, E \in \mathcal{D}$,

$$pD + (1-p)E \in \mathcal{D}, \quad \forall\, 0 \leq p \leq 1,$$

Extreme points of convex sets

- Let $\mathcal{D}$ be the set of all $n \times n$ doubly stochastic matrices.
- Then $\mathcal{D}$ is a convex set, that is for $D, E \in \mathcal{D}$,

$$pD + (1-p)E \in \mathcal{D}, \quad \forall\ 0 \leq p \leq 1,$$

- that is, the line segment joining D, E is contained in $\mathcal{D}$.

Extreme points of convex sets

- Let $\mathcal{D}$ be the set of all $n \times n$ doubly stochastic matrices.
- Then $\mathcal{D}$ is a convex set, that is for $D, E \in \mathcal{D}$,

$$pD + (1-p)E \in \mathcal{D}, \quad \forall\, 0 \leq p \leq 1,$$

- that is, the line segment joining D, E is contained in $\mathcal{D}$.
- A matrix $F \in \mathcal{D}$ is said to be an extreme point if

$$F = pD + (1-p)E$$

with $0 < p < 1$, implies $D = E = F$.

Extreme points of convex sets

- Let $\mathcal{D}$ be the set of all $n \times n$ doubly stochastic matrices.
- Then $\mathcal{D}$ is a convex set, that is for $D, E \in \mathcal{D}$,

$$pD + (1-p)E \in \mathcal{D}, \quad \forall\, 0 \leq p \leq 1,$$

- that is, the line segment joining D, E is contained in $\mathcal{D}$.
- A matrix $F \in \mathcal{D}$ is said to be an extreme point if

$$F = pD + (1-p)E$$

  with $0 < p < 1$, implies $D = E = F$.
- Simple Exercise: Permutation matrices are extreme points of $\mathcal{D}$.

Birkhoff-von Neumann theorem

- Theorem 2.6: A doubly stochastic matrix is an extreme point of the convex set of doubly stochastic matrices if and only if it is a permutation matrix.

Birkhoff-von Neumann theorem

- Theorem 2.6: A doubly stochastic matrix is an extreme point of the convex set of doubly stochastic matrices if and only if it is a permutation matrix.
- Proof. Difficult exercise. (Omitted).

Birkhoff-von Neumann theorem

- Theorem 2.6: A doubly stochastic matrix is an extreme point of the convex set of doubly stochastic matrices if and only if it is a permutation matrix.
- Proof. Difficult exercise. (Omitted).
- END OF LECTURE 2.