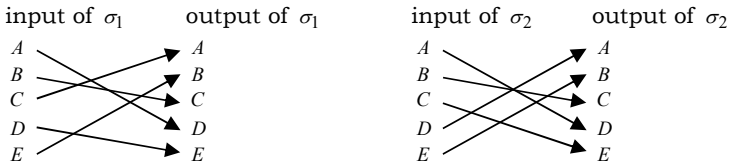
# Permutation

**Introduction**

Permutation of set X is bijective mapping from set X to itself (recall that : bijective mapping is a one–to–one mapping, surjective mapping is a many–to–one mapping, while injective mapping is a one–to–many mapping). Given a set X, all possible permutations of X form a set, denoted by Perm(X). The size of Perm(X) is thus ${}_N P_N = N!$, where N is size of X. A permutation can be denoted in three ways : (1) graphical notation, (2) Cauchy's two line notation and (3) cycle notation. Suppose set X = {A, B, C, D, E}, $\sigma_1$ and $\sigma_2$, are two possible permutations of set X, i.e. $\sigma_1, \sigma_2 \in \text{Perm}(X)$, and given that the mappings are :

$$\begin{aligned} \sigma_1(A) &= D & \qquad \sigma_2(A) &= D \\ \sigma_1(B) &= C & \qquad \sigma_2(B) &= C \\ \sigma_1(C) &= A & \qquad \sigma_2(C) &= E \\ \sigma_1(D) &= E & \qquad \sigma_2(D) &= A \\ \sigma_1(E) &= B & \qquad \sigma_2(E) &= B \end{aligned}$$

(1) Graphical notation




(2) Cauchy's two line notation (input in the 1st line, output in the 2nd line)

$$\sigma_1 = \begin{pmatrix} A & B & C & D & E \\ D & C & A & E & B \end{pmatrix} = \begin{pmatrix} A & D & E & B & C \\ D & E & B & C & A \end{pmatrix}$$

$$\sigma_2 = \begin{pmatrix} A & B & C & D & E \\ D & C & E & A & B \end{pmatrix} = \begin{pmatrix} A & D & B & C & E \\ D & A & C & E & B \end{pmatrix}$$

(3) Cycle notation (all permutations can be written as composition of disjoint cycles, it will be proved later)

$$\begin{aligned} \sigma_1 &= (A \;\; D \;\; E \;\; B \;\; C) \\ \sigma_2 &= (A \;\; D) \circ (B \;\; C \;\; E) \end{aligned}$$

Remark : In Cauchy's two line notation, the order of the columns are not important, columns can be interchanged. Just imagine a function defined as $f(n) = n^2$, where n = 1, 2, 3 only, it can be represented as the following, the order of mapping does not matter. Always remember that permutation is simply a bijective function.

$$f = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 4 & 9 \end{pmatrix} = \begin{pmatrix} 3 & 2 & 1 \\ 9 & 4 & 1 \end{pmatrix}$$

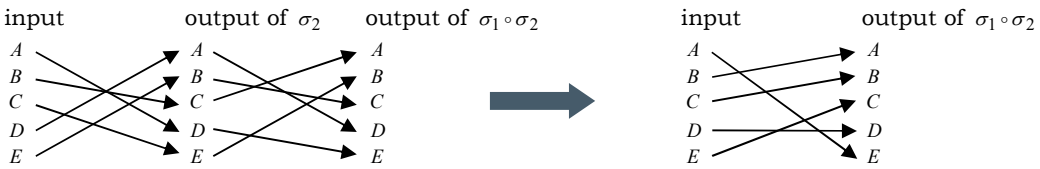
**Composition of permutations**

Similar to the composition of functions (please read the 'Chain Rule' document), composition of permutations means cascade of permutation. Since permutations are bijective mapping of a set, hence composition of permutations is also another permutation of the same set. For example, suppose $\sigma_1, \sigma_2, \sigma_3 \in \text{Perm}(X)$, composition is denoted as $\sigma_1 \circ \sigma_2$, or $\sigma_1 \circ \sigma_2 \circ \sigma_3$, which are also permutations of set X, while the mapping result is denoted as $(\sigma_1 \circ \sigma_2)(x)$ or $(\sigma_1 \circ \sigma_2 \circ \sigma_3)(x)$, which means applying mapping $\sigma_1$ on the result of $\sigma_2$, and so on (recall : mapping is right–associative).

$$\begin{aligned} (\sigma_1 \circ \sigma_2)(x) &= \sigma_1(\sigma_2(x)) && \text{where } \sigma_1 \circ \sigma_2 \in \text{Perm}(X) \text{ and } x \in X \\ (\sigma_1 \circ \sigma_2 \circ \sigma_3)(x) &= \sigma_1(\sigma_2(\sigma_3(x))) && \text{where } \sigma_1 \circ \sigma_2 \circ \sigma_3 \in \text{Perm}(X) \text{ and } x \in X \end{aligned}$$

Using $\sigma_1$ and $\sigma_2$ in the previous example, graphical notation, Cauchy's two line notation and cycle notation of the composition $\sigma_1 \circ \sigma_2$ can be written as the following.

(1) Graphical notation




(2) Cauchy's two line notation

$$\sigma_2 = \begin{pmatrix} A & B & C & D & E \\ D & C & E & A & B \end{pmatrix} \qquad \sigma_1 = \begin{pmatrix} A & B & C & D & E \\ D & C & A & E & B \end{pmatrix}$$

$$\sigma_1 \circ \sigma_2 = \begin{pmatrix} A & B & C & D & E \\ E & A & B & D & C \end{pmatrix} = \begin{pmatrix} A & E & C & B & D \\ E & C & B & A & D \end{pmatrix}$$

(3) Cycle notation

There is no direct method for simplifying composition represented in cycle notation, we need to trace the permutation step by step using first principle, i.e. finding $\sigma_1 \circ \sigma_2$ using above methods, and decompose the permutation into disjoint cycles. We will discuss composition and cycle notation later.

$$\sigma_1 \circ \sigma_2 = (A \quad E \quad C \quad B) \circ (D)$$

Remark 1 : Just like the composition of function, composition of permutation is associative.

$$\sigma_1 \circ \sigma_2 \circ \sigma_3 \circ \sigma_4 = \sigma_{12} \circ \sigma_3 \circ \sigma_4 \quad \text{where } \sigma_{12} = \sigma_1 \circ \sigma_2$$
$$= \sigma_1 \circ \sigma_{23} \circ \sigma_4 \quad \text{where } \sigma_{23} = \sigma_2 \circ \sigma_3$$
$$= \sigma_1 \circ \sigma_2 \circ \sigma_{34} \quad \text{where } \sigma_{34} = \sigma_3 \circ \sigma_4$$
$$\text{recall } ((f \circ g) \circ h)(x) = (f \circ (g \circ h))(x)$$

Remark 2 : Just like the composition of function, composition of permutation is in general not commutative.

$$\sigma_1 \circ \sigma_2 \neq \sigma_2 \circ \sigma_1 \qquad \text{recall } (f \circ g)(x) \neq (g \circ f)(x)$$

**Inverse of permutation**

Suppose $X = \{x_n, n \in [1, N]\}$ , the inverse of permutation is defined as :

$$\sigma^{-1}(x_m) = x_n \quad \text{where } \sigma(x_n) = x_m \quad \forall x_n, x_m \in X \text{ and } \sigma, \sigma^{-1} \in Perm(X)$$
$$\Rightarrow (\sigma^{-1} \circ \sigma)(x_n) = x_n$$
$$\Rightarrow (\sigma^{-1} \circ \sigma) = e \quad \text{where } e(x_n) = x_n \quad \forall x_n \in X \text{ and } e \in Perm(X) \text{, see identity permutation}$$

<u>Inverse permutation of composition</u>

Consider :

$$(\sigma_2^{-1} \circ \sigma_1^{-1}) \circ (\sigma_1 \circ \sigma_2) = \sigma_2^{-1} \circ (\sigma_1^{-1} \circ \sigma_1) \circ \sigma_2 \quad \text{since composition is associative}$$
$$= \sigma_2^{-1} \circ e \circ \sigma_2 \quad \text{since } \sigma^{-1} \circ \sigma = e$$
$$= \sigma_2^{-1} \circ \sigma_2 \quad \text{since } e \circ \sigma_2 = \sigma_2$$
$$= e$$
$$\Rightarrow (\sigma_1 \circ \sigma_2)^{-1} = (\sigma_2^{-1} \circ \sigma_1^{-1})$$

<u>Inverse permutation of cyclic permutation</u>

Given that :

(1) $\sigma(x_n) = \begin{cases} x_{n+1} & 1 \le n < N \\ x_1 & n = N \end{cases}$ where $\sigma \in Perm(X)$

(2) $\pi(x_n) = \begin{cases} x_{n-1} & 1 < n \le N \\ x_N & n = 1 \end{cases}$ where $\pi \in Perm(X)$ , i.e. cyclic permutation in a reverse direction

$$(\pi \circ \sigma)(x_n) = \begin{cases} x_n & 1 \le n < N \\ x_N & n = N \end{cases} = e$$

$$\Rightarrow \sigma^{-1} = \pi \quad \text{inverse of } \sigma \text{ is a cyclic permutation in a reverse direction}$$

**Special permutations**
Lets look at some special permutations : (1) identity permutation, (2) cyclic permutation, (3) orbit, (4) transposition, (5) adjacent transposition and (6) permutation of set $S_n$. For the following, we suppose set $X = \{x_n, n \in [1,N]\}$.

Identity permutation
Identity permutation is defined as permutation that maps an element to itself, i.e. it maintains the same element order in the set. Suppose $\sigma \in Perm(X)$, then $\sigma$ is an identity permutation if :

$$\sigma(x_n) = x_n \qquad \forall x_n \in X$$

We usually represent identity permutation as e, and here are its properties.

$$\sigma \circ e = \sigma$$
$$e \circ \sigma = \sigma$$
$$e^{-1} = e$$

Cyclic permutation
Cyclic permutation is defined as permutation that maps an element to itself after apply the same mapping for N times (i.e. cascading N identical permutations through composition), thus it forms a cycle. Suppose $\sigma \in Perm(X)$, then $\sigma$ is a cyclic permutation if :

$$\underbrace{\sigma(\sigma(\sigma(...\sigma}_{N}(x_n)))) = x_n \qquad \forall x_n \in X \qquad \text{(equation 1)}$$

or $$\underbrace{\sigma \circ \sigma \circ \sigma \circ ... \sigma}_{N} = e$$

The simplest cyclic permutation performs one–element–shift, while the general cyclic permutation performs one–element–shift after applying column rearrangment f, where f can also be regarded as another permutation of X, i.e. $f \in Perm(X)$. If there exists no column rearrangement that makes a permutation a one–element–shift permutation, then the permutation is non cyclic.

(1) one–element–shift without column rearrangement (i.e. the simplest cyclic permutation)

$$\sigma(x_n) = \begin{cases} x_{n+1} & 1 \le n < N \\ x_1 & n = N \end{cases} \qquad \text{(equation 2)}$$

$$\sigma = \begin{pmatrix} x_1 & x_2 & ... & x_{N-1} & x_N \\ x_2 & x_3 & ... & x_N & x_1 \end{pmatrix}$$

(2) one–element–shift after suitable column rearrangement (i.e. the general cyclic permutation)

$$\sigma(f(x_n)) = \begin{cases} f(x_{n+1}) & 1 \le n < N \\ f(x_1) & n = N \end{cases} \qquad \text{(equation 3)}$$

$$\sigma = \begin{pmatrix} f(x_1) & f(x_2) & ... & f(x_{N-1}) & f(x_N) \\ f(x_2) & f(x_3) & ... & f(x_N) & f(x_1) \end{pmatrix} \qquad \text{where } f \in Perm(X) \text{ is a column rearrangement}$$

It is easy to observe that equation 2 satisfies the cyclic permutation definition in equation 1. Now, lets prove that equation 3 does also satisfy the cyclic permutation definition in equation 1. From equation 3, we have :

$$(f^{-1} \circ \sigma \circ f)(x_n) = \begin{cases} x_{n+1} & 1 \le n < N \\ x_1 & n = N \end{cases}$$

$$\underbrace{(f^{-1} \circ \sigma \circ f) \circ ... \circ (f^{-1} \circ \sigma \circ f)}_{N}(x_n) = x_n \qquad \forall x_n \in X$$

$$f^{-1} \circ \sigma \circ (f \circ f^{-1}) \circ ... (f \circ f^{-1}) \circ \sigma \circ f = e \qquad \text{since composition is associative}$$

$$f^{-1} \circ \underbrace{\sigma \circ \sigma \circ ... \circ \sigma}_{N} \circ f = e \qquad \text{since } f \circ f^{-1} = f^{-1} \circ f = e$$

$$\underbrace{\sigma \circ \sigma \circ ... \circ \sigma}_{N} = f \circ e \circ f^{-1}$$

$$\underbrace{\sigma \circ \sigma \circ ... \circ \sigma}_{N} = e$$

Example 1 : one element shift

$$\sigma = \begin{pmatrix} A & B & C & D & E \\ B & C & D & E & A \end{pmatrix}$$

Example 2 : multi element shift

$$\sigma = \begin{pmatrix} A & B & C & D & E \\ D & E & A & B & C \end{pmatrix} \quad \text{where} \quad f = \begin{pmatrix} A & B & C & D & E \\ A & D & B & E & C \end{pmatrix} \quad \text{and} \quad f^{-1} = \begin{pmatrix} A & B & C & D & E \\ A & C & E & B & D \end{pmatrix}$$

$$\sigma \circ f = \begin{pmatrix} A & B & C & D & E \\ D & B & E & C & A \end{pmatrix}$$

$$f^{-1} \circ \sigma \circ f = \begin{pmatrix} A & B & C & D & E \\ B & C & D & E & A \end{pmatrix} \quad \text{besides } f^{-1} \circ f = e \text{. How to choose } f \text{ ? Hints : find } f^{-1} \text{ first.}$$

Example 3 : with some column rearrangement

$$\sigma = \begin{pmatrix} A & B & C & D & E \\ C & E & D & B & A \end{pmatrix} \quad \text{where} \quad f = \begin{pmatrix} A & B & C & D & E \\ A & C & D & B & E \end{pmatrix} \quad \text{and} \quad f^{-1} = \begin{pmatrix} A & B & C & D & E \\ A & D & B & C & E \end{pmatrix}$$

$$\sigma \circ f = \begin{pmatrix} A & B & C & D & E \\ C & D & B & E & A \end{pmatrix}$$

$$f^{-1} \circ \sigma \circ f = \begin{pmatrix} A & B & C & D & E \\ B & C & D & E & A \end{pmatrix} \quad \text{besides } f^{-1} \circ f = e \text{. How to choose } f \text{ ? Hints : find } f^{-1} \text{ first.}$$

Example 4 : exists no column rearrangement

$$\sigma = \begin{pmatrix} A & B & C & D & E \\ C & E & D & A & B \end{pmatrix} \quad \text{because } \sigma \text{ forms two disjoint orbits.}$$

Orbit

Orbit is defined as permutation of subset of set X, orbits are disjoint if they are generated from disjoint subsets of X, i.e. subsets having empty intersection. Suppose $\sigma_1$, $\sigma_2$, $\sigma_3 \in$ Perm(X), then $\sigma_1$, $\sigma_2$ and $\sigma_3$ are disjoint orbits if :

$$\sigma_1 \in Perm(X_1) \quad \text{where } X_1 \subset X$$
$$\sigma_2 \in Perm(X_2) \quad \text{where } X_2 \subset X$$
$$\sigma_3 \in Perm(X_3) \quad \text{where } X_3 \subset X \quad \text{such that } X_1 \cap X_2 = 0 \text{ , } X_2 \cap X_3 = 0 \text{ and } X_3 \cap X_1 = 0$$

The composition of permutations is associative but not commutive, yet the composition of disjoint orbits is associative and commutative.

$$\sigma_1 \circ (\sigma_2 \circ \sigma_3) = (\sigma_1 \circ \sigma_2) \circ \sigma_3$$
$$\sigma_1 \circ \sigma_2 \circ \sigma_3 = \sigma_1 \circ \sigma_3 \circ \sigma_2$$
$$= \sigma_3 \circ \sigma_1 \circ \sigma_2$$
$$= \dots$$

Transposition and adjacent transposition

Transposition is an orbit with exactly two elements. Suppose $\sigma \in$ Perm(X), then $\sigma$ is a transposition if :

$$\sigma(x_n) = x_m$$
$$\sigma(x_m) = x_n \quad \text{where } x_n, x_m \in X$$

which is a swap of the two elements. The inverse of transposition is the transposition itself. Lets prove it.

$$(\sigma \circ \sigma)(x_n) = x_n$$
$$(\sigma \circ \sigma)(x_m) = x_m \quad \text{where } x_n, x_m \in X$$

Adjacent transposition is a transposition, in which the two elements are adjacent in set X, i.e. m = n + 1.

Permutation of $S_n$

Normally when we are talking about permutation, we do not need to handle general set X, instead X is simply a set of integer starting from 1. For convenience, we define set $S_n$ as a set of consecutive integers, starting from 1 to n, i.e. $S_n$ = {1, 2, 3, …, n}. Please note that the integers in set $S_n$ **stand for all possible values of input and output** in the bijective mapping of permutation, they **do not stand for location** in permutation (please don't get confused). Here is an example of composition of permutation from $S_n$. Suppose $\sigma_1$, $\sigma_2 \in$ Perm($S_n$), we have :

$$\sigma_2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 3 & 5 & 1 & 2 \end{pmatrix}$$

$$\sigma_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 3 & 1 & 5 & 2 \end{pmatrix}$$

$$\sigma_1 \circ \sigma_2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 1 & 2 & 4 & 3 \end{pmatrix}$$

**Property 1 : Decomposition into disjoint orbits**

All permutations can always be decomposed into disjoint orbits (there is only one orbit when it is a cyclic permutation). Lets prove it by mathematical induction. Suppose set $X_N = \{x_n, n \in [1,N]\}$, we will then go through two steps : (1) prove that σ can be decomposed into disjoint orbits $\forall\sigma \in Perm(X_2)$ (and also $\forall\sigma \in Perm(X_3)$ for demonstration purpose), (2) given that σ can be decomposed into disjoint orbits $\forall\sigma \in Perm(X_{N-1})$, prove that σ can be decomposed into disjoint orbits $\forall\sigma \in Perm(X_N)$. Note : there are (N-1)! and N! permutations inside set $Perm(X_{N-1})$ and set $Perm(X_N)$ respectively.

Step 1

For case N = 2

$$\begin{pmatrix} x_1 & x_2 \\ x_1 & x_2 \end{pmatrix} = (x_1)\circ(x_2) \qquad \begin{pmatrix} x_1 & x_2 \\ x_2 & x_1 \end{pmatrix} = (x_1 \quad x_2)$$

For case N = 3

$$\begin{pmatrix} x_1 & x_2 & x_3 \\ x_1 & x_2 & x_3 \end{pmatrix} = (x_1)\circ(x_2)\circ(x_3) \qquad \begin{pmatrix} x_1 & x_2 & x_3 \\ x_1 & x_3 & x_2 \end{pmatrix} = (x_1)\circ(x_2 \quad x_3)$$

$$\begin{pmatrix} x_1 & x_2 & x_3 \\ x_2 & x_3 & x_1 \end{pmatrix} = (x_1 \quad x_2 \quad x_3) \qquad \begin{pmatrix} x_1 & x_2 & x_3 \\ x_2 & x_1 & x_3 \end{pmatrix} = (x_3)\circ(x_1 \quad x_2)$$

$$\begin{pmatrix} x_1 & x_2 & x_3 \\ x_3 & x_1 & x_2 \end{pmatrix} = (x_1 \quad x_3 \quad x_2) \qquad \begin{pmatrix} x_1 & x_2 & x_3 \\ x_3 & x_2 & x_1 \end{pmatrix} = (x_2)\circ(x_3 \quad x_1)$$

Step 2

Assuming case N–1 is true, we consider case N.

$$\sigma = \begin{pmatrix} x_1 & x_2 & \dots & x_i & \dots & x_j & \dots & x_{N-1} & x_N \\ \sigma(x_1) & \sigma(x_2) & \dots & \sigma(x_i) & \dots & \sigma(x_j) & \dots & \sigma(x_{N-1}) & \sigma(x_N) \end{pmatrix} \quad \text{where } \sigma \in Perm(X_N)$$

There are three cases.

(case 1) when $\sigma(x_N) = x_N$, we can always find a permutation $\pi \in Perm(X_{N-1})$, such that :

$$\pi(x_n) = \sigma(x_n) \quad \forall n \in [1, N-1]$$

$$\Rightarrow \sigma = \begin{pmatrix} x_1 & x_2 & \dots & x_i & \dots & x_j & \dots & x_{N-1} & x_N \\ \sigma(x_1) & \sigma(x_2) & \dots & \sigma(x_i) & \dots & \sigma(x_j) & \dots & \sigma(x_{N-1}) & x_N \end{pmatrix}$$

$$= \begin{pmatrix} x_1 & x_2 & \dots & x_i & \dots & x_j & \dots & x_{N-2} & x_{N-1} \\ \pi(x_1) & \pi(x_2) & \dots & \pi(x_i) & \dots & \pi(x_j) & \dots & \pi(x_{N-2}) & \pi(x_{N-1}) \end{pmatrix}$$

$$= \pi$$

$$= \pi_1 \circ \pi_2 \circ \dots \circ \pi_{K-1} \circ \pi_K \quad \text{since } \pi \text{ can be decomposed into disjoint orbits}$$

(case 2) when $\sigma(x_N) = x_i$ and $\sigma(x_i) = x_N$, we can always find a permutation $\pi \in Perm(X_{N-1})$, such that :

$$\pi(x_n) = \begin{cases} \sigma(x_n) & n \in [1, N-1], n \neq i \\ x_i & n = i \end{cases}$$

$$\Rightarrow \sigma = \begin{pmatrix} x_1 & x_2 & \dots & x_i & \dots & x_j & \dots & x_{N-1} & x_N \\ \sigma(x_1) & \sigma(x_2) & \dots & x_N & \dots & \sigma(x_j) & \dots & \sigma(x_{N-1}) & x_i \end{pmatrix}$$

$$= \begin{pmatrix} x_1 & x_2 & \dots & x_i & \dots & x_j & \dots & x_{N-1} & x_N \\ \pi(x_1) & \pi(x_2) & \dots & x_i & \dots & \pi(x_j) & \dots & \pi(x_{N-1}) & x_N \end{pmatrix} \circ (x_i \quad x_N)$$

$$= \pi \circ (x_i \quad x_N)$$

$$= \pi_1 \circ \pi_2 \circ \dots \circ \pi_{K-1} \circ \pi_K \circ (x_i \quad x_N) \quad \text{since } \pi \text{ can be decomposed into disjoint orbits}$$

(case 3) when $\sigma(x_N) = x_j$ and $\sigma(x_i) = x_N$, we can always find a permutation $\pi \in Perm(X_{N-1})$, such that :

$$\pi(x_n) = \begin{cases} \sigma(x_n) & n \in [1, N-1], n \neq i \\ x_j & n = i \end{cases}$$

$$\Rightarrow \sigma = \begin{pmatrix} x_1 & x_2 & \dots & x_i & \dots & x_j & \dots & x_{N-1} & x_N \\ \sigma(x_1) & \sigma(x_2) & \dots & x_N & \dots & \sigma(x_j) & \dots & \sigma(x_{N-1}) & x_j \end{pmatrix}$$

$$= \begin{pmatrix} x_1 & x_2 & \dots & x_i & \dots & x_j & \dots & x_{N-1} & x_N \\ \pi(x_1) & \pi(x_2) & \dots & x_j & \dots & \pi(x_j) & \dots & \pi(x_{N-1}) & x_N \end{pmatrix} \circ (x_i \quad x_N)$$

$$= \pi \circ (x_i \quad x_N)$$

$$= \pi_1 \circ \pi_2 \circ \dots \circ \pi_{K-1} \circ \pi_K \circ (x_i \quad x_N) \quad \text{since } \pi \text{ can be decomposed into disjoint orbits}$$

The sight difference between case 2 and case 3 is that : in case 2, $x_i$ itself form an orbit in the decomposition of π, in case 3, $x_i$ belongs to one of the orbit in $\pi_k$, where $k \in [1,K]$, in the decomposition of π. In the decomposition of σ, $x_N$ joins $x_i$ to form a larger orbit.

About cycle notation

As all permutations can be decomposed into disjoint orbits, permutations can be represented in cycle notation, which is simply a single line notaton, representing a bijective mapping from an element to the adjacent element on the right hand side. Here are two examples, $\sigma_1$ is a cyclic permutation, while $\sigma_2$ can be decomposed into two disjoint orbits.

$$\sigma_1 = \begin{pmatrix} A & B & C & D & E \\ D & C & A & E & B \end{pmatrix} = \begin{pmatrix} A & D & E & B & C \\ D & E & B & C & A \end{pmatrix} = (A \quad D \quad E \quad B \quad C)$$

$$\sigma_2 = \begin{pmatrix} A & B & C & D & E \\ D & C & E & A & B \end{pmatrix} = \begin{pmatrix} A & D & B & C & E \\ D & A & C & E & B \end{pmatrix} = \begin{pmatrix} A & D \\ D & A \end{pmatrix} \circ \begin{pmatrix} B & C & E \\ C & E & B \end{pmatrix} = (A \quad D) \circ (B \quad C \quad E)$$

$$= (B \quad C \quad E) \circ (A \quad D)$$

**Property 2 : Decomposition into adjacent transposition**

All permutations can be decomposed into a sequence of (joint) adjacent transpositions. Since adjacent transpositions are joint, the decomposition is not commutative. Lets prove it. All orbits can be decomposed into a smaller orbit and an adjacent transposition, as shown in the following (since columns of a permutation can be interchanged, without loss of generality, I represent an orbit as $\sigma(x_i) = x_{i+1}, \forall i$).

$$\sigma = \begin{pmatrix} x_1 & x_2 & \cdots & x_{N-2} & x_{N-1} & x_N \\ x_2 & x_3 & \cdots & x_{N-1} & x_N & x_1 \end{pmatrix}$$

$$= \begin{pmatrix} x_1 & x_2 & \cdots & x_{N-2} & x_{N-1} \\ x_2 & x_3 & \cdots & x_{N-1} & x_1 \end{pmatrix} \circ (x_{N-1} \quad x_N)$$

$$= \sigma' \circ (x_{N-1} \quad x_N)$$

This procedure can be applied to orbit σ' repeatedly until orbit σ is completely decomposed into a sequence of adjacent transpositions. Furthermore from property 1, since all permutations can be decomposed into disjoint orbits, together with the fact we derived above, we can conclude that all permutations can be decomposed into a sequence of adjacent transpositions. This property can be used to prove why all random permutation can be sorted by bubble sort, which is a sequence of adjacent transposition.

**Parity of permutation**

Parity of permutation (sgn) is a mapping from a permutation to either −1 or +1, respectively representing an odd or even number of adjacent transpositions in the decomposition of the permutation.

$$\text{sgn} : Perm(X) \to \{-1, +1\}$$

Sgn of an identity permutation is +1 (i.e. even), since its decomposition contains no adjacent transposition. Sgn of an adjacent transposition is −1 (i.e. odd), since its decomposition contains exactly one adjacent transposition. Sgn of cyclic permutation is $(-1)^{(\text{size}-1)}$, since its decomposition contains size − 1 adjacent transpositions. Sgn of composition of two permutations equals to the product of individual permutation's sgn.

$$\begin{aligned} \text{sgn}(\sigma \circ \pi) &= \text{sgn}(\sigma_1 \circ \sigma_2 \circ \ldots \circ \sigma_N \circ \pi_1 \circ \pi_2 \circ \ldots \circ \pi_M) && \text{decompose } \sigma \text{ and } \pi \text{ into adjacent transposition} \\ &= (-1)^{N+M} \\ &= (-1)^N \times (-1)^M \\ &= \text{sgn}(\sigma) \times \text{sgn}(\pi) \end{aligned}$$

If σ and π are inverse of each other, we have :

$$\begin{aligned} \text{sgn}(\sigma \circ \pi) &= \text{sgn}(\sigma) \times \text{sgn}(\pi) \\ \text{sgn}(e) &= \text{sgn}(\sigma) \times \text{sgn}(\pi) \\ +1 &= \text{sgn}(\sigma) \times \text{sgn}(\pi) \quad \Rightarrow \quad \text{sgn}(\sigma) = \text{sgn}(\pi) = \text{sgn}(\sigma^{-1}) && \text{hence parity of a permutation is the same as its inverse.} \end{aligned}$$

A permutation can be decomposed into sequence of adjacent transposition in multiple ways, but all decompositions of the same permutation should have the same parity, i.e. the parity of a permutation is unique. Lets prove it. Suppose a permutation σ can be decomposed into two different adjacent transposition sequences with parity respectively equals to $(-1)^K$ and $(-1)^M$. We have to show that $(-1)^K \times (-1)^M = +1$ or equivalently K + M is even.

$$\begin{aligned} \sigma &= \pi_1 \circ \pi_2 \circ \ldots \circ \pi_K = \eta_1 \circ \eta_2 \circ \ldots \circ \eta_M \\ &\pi_1 \circ \pi_2 \circ \ldots \circ \pi_{K-1} = \eta_1 \circ \eta_2 \circ \ldots \circ \eta_M \circ \pi_K && \text{since inverse of a transposition is itself} \\ &\pi_1 \circ \pi_2 \circ \ldots \circ \pi_{K-2} = \eta_1 \circ \eta_2 \circ \ldots \circ \eta_M \circ \pi_K \circ \pi_{K-1} \\ \Rightarrow \quad e &= \eta_1 \circ \eta_2 \circ \ldots \circ \eta_M \circ \pi_K \circ \pi_{K-1} \circ \ldots \circ \pi_2 \circ \pi_1 && \text{where e is identity permutation} \\ \text{sgn}(e) &= \text{sgn}(\eta_1 \circ \eta_2 \circ \ldots \circ \eta_M \circ \pi_K \circ \pi_{K-1} \circ \ldots \circ \pi_2 \circ \pi_1) \\ +1 &= (-1)^{K+M} && \text{hence parity of a permutation is unique.} \end{aligned}$$

**Permutation matrix**

Permutation matrix is a binary matrix that can permute the rows of a N×M matrix A or the columns of a M×N matrix B, when it is multiplied to matrix A or matrix B. Suppose we want to perform permutation $\sigma \in \mathrm{Perm}(S_N)$ on the rows of A or columns of B, we firstly define a N×N permutation matrix P as :

$$P_\sigma = (p_{n,m})_{n,m\in[1,N]} \quad \text{such that } p_{n,m} = \begin{bmatrix} 1 & if & m=\sigma(n) \\ 0 & if & m\neq\sigma(n) \end{bmatrix}$$

or $$P_\sigma = \begin{bmatrix} P_{\sigma(1)} \\ P_{\sigma(2)} \\ P_{\sigma(3)} \\ \dots \\ P_{\sigma(N)} \end{bmatrix}$$ where $P_k$ is a 1×N row matrix, with 1 at position k, and 0 otherwise

Suppose matrix A and B are :

$$A = \begin{bmatrix} A_1 \\ A_2 \\ A_3 \\ \dots \\ A_N \end{bmatrix}$$ where $A_n$ is 1×M row matrix, $\forall n\in[1,N]$

$$B = \begin{bmatrix} B_1 & B_2 & B_3 & \dots & B_N \end{bmatrix}$$ where $B_n$ is M×1 column matrix, $\forall n\in[1,N]$

We can then apply permutation σ on the rows of A to form a new matrix A' and apply permutation σ on the columns of B to form a new matrix B' as the following.

$$A' = P_\sigma A \quad \text{(equation 4)}$$

$$B' = BP_\sigma^T \quad \text{(equation 5)}$$

Equation 4 can be derived simply by noticing that :

$$A' = \begin{bmatrix} P_{\sigma(1)} \\ P_{\sigma(2)} \\ P_{\sigma(3)} \\ \dots \\ P_{\sigma(N)} \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \\ A_3 \\ \dots \\ A_N \end{bmatrix} = \begin{bmatrix} A_{\sigma(1)} \\ A_{\sigma(2)} \\ A_{\sigma(3)} \\ \dots \\ A_{\sigma(N)} \end{bmatrix}$$

Equation 5 can be derived from equation 4 by noticing that : for every matrix B, there always exists matrix $A^T = B$, and hence the column permutation in B is corresponding to the row permutation in A, i.e. $A'^T = B'$.

$$A' = P_\sigma A$$
$$\Rightarrow \quad A'^T = A^T P_\sigma^T$$
$$\Rightarrow \quad B' = BP_\sigma^T$$

Multiplication of permutation matrix

Suppose $\sigma,\pi\in\mathrm{Perm}(S_N)$, the product of permutation matrices $P_\sigma P_\pi$ gives Q.

$$P_\sigma P_\pi = Q$$
$$= (q_{n,m})_{n,m\in[1,N]}$$
$$q_{n,m} = \sum_{n=1}^{N} p_{\sigma,n,k}\, p_{\pi,k,m} \quad \text{where } p_{\sigma,n,k} = \begin{bmatrix} 1 & if & k=\sigma(n) \\ 0 & if & k\neq\sigma(n) \end{bmatrix} \text{ and } p_{\pi,k,m} = \begin{bmatrix} 1 & if & m=\pi(k) \\ 0 & if & m\neq\pi(k) \end{bmatrix}$$
$$= \sum_{n=1}^{N} \delta_{\sigma(n),k}\delta_{\pi(k),m}$$
$$= \delta_{\sigma(n),\pi^{-1}(m)}$$
$$= \delta_{\pi(\sigma(n)),m}$$
$$= \delta_{(\pi\circ\sigma)(n),m}$$

thus $$P_\sigma P_\pi = P_{\pi\circ\sigma}$$

This is not a good looking result, as the ordering of permutation is reversed. Please note : composition of permutation is right associative.

Transpose of permutation matrix

Suppose $\sigma \in Perm(S_N)$, the transpose of permutation matrices $P_\sigma$ gives Q.

$$
\begin{aligned}
P_\sigma{}^T &= Q \\
&= (q_{n,m})_{n,m \in [1,N]} \\
q_{n,m} &= p_{\sigma,m,n} \qquad \text{where } p_{\sigma,m,n} = \begin{cases} 1 & if \quad n = \sigma(m) \\ 0 & if \quad n \neq \sigma(m) \end{cases} \\
&= \delta_{\sigma(m),n} \\
&= \delta_{\sigma^{-1}(n),m}
\end{aligned}
$$

thus $P_\sigma{}^T = P_{\sigma^{-1}}$

Inverse of permutation matrix

Suppose $\sigma, \pi \in Perm(S_N)$, such that $\pi = \sigma^{-1}$, then we have :

$$
\begin{aligned}
P_\pi P_\sigma &= P_{\sigma \circ \pi} \qquad \text{where } \pi \in Perm(S_N) \text{ such that } \pi = \sigma^{-1} \\
&= P_{\sigma \circ \sigma^{-1}} \\
&= P_e \qquad \text{where } e \in Perm(S_N) \text{ such that } e(n) = n, \forall n \in [1, N] \\
&= I \\
P_\sigma^{-1} &= P_{\sigma^{-1}}
\end{aligned}
$$

Combining the result of transpose and the result of inverse, we have : (1) transpose of permutation matrix, (2) inverse of permutation matrix and (3) matrix of inverse permutation being equivalent.

$$P_\sigma^T = P_\sigma^{-1} = P_{\sigma^{-1}}$$