Mathematical Summary of Permutation and Distance Computation

1. Problem Representation and Permutation

A **Mixed-Integer Programming (MIP)** problem is defined as:

$$Ax \leq b, \quad x \in \mathbb{Z}^p \times \mathbb{R}^{n-p}$$

where:

- $A \in \mathbb{R}^{m \times n}$ is the constraint matrix,
- $x \in \mathbb{R}^n$ is the decision variable vector,
- $b \in \mathbb{R}^m$ is the right-hand side vector.

We apply **two independent permutations**:

- A **row permutation** P_{row}, which reorders the constraints.
- A **column permutation** P_{col}, which reorders the variables.

The **permuted problem** becomes:

$$P_{\text{row}} A P_{\text{col}} x \leq P_{\text{row}} b$$

where:

- $P_{\text{row}} \in \mathbb{R}^{m \times m}$ is a **permutation matrix** representing the reordering of constraints.
- $P_{\text{col}} \in \mathbb{R}^{n \times n}$ is a **permutation matrix** representing the reordering of variables.

Each permutation matrix is **orthogonal**:

$$PP^\top = P^\top P = I$$

and contains exactly one "1" in each row and each column.

2. Distance Between Permutations

To quantify the difference between two permutations, we use **separate distance measures** for rows and columns.

2.1 Row and Column Permutations

Given two permutations:

- **Row permutation** π^1_{row} vs. π^2_{row}, each a bijection of $\{1, \dots, m\}$.
- **Column permutation** π^1_{col} vs. π^2_{col}, each a bijection of $\{1, \dots, n\}$.

We define distances d_{rows} and d_{cols} separately.

2.2 Hamming Distance

The **Hamming distance** measures how many positions differ:

$$d_{\text{Hamming}}(\pi^1, \pi^2) = \sum_{i=1}^{k} \mathbf{1}(\pi^1(i) \neq \pi^2(i))$$

where $\mathbf{1}(\cdot)$ is the indicator function.

2.3 Kendall Tau Distance

The **Kendall Tau distance** counts the number of pairwise **inversions**:

$$d_{\text{Kendall}}(\pi^1, \pi^2) = \sum_{1 \leq i < j \leq k} \mathbf{1}\Big((\pi^1(i) < \pi^1(j) \text{ and } \pi^2(i) > \pi^2(j)) \text{ or } (\pi^1(i) > \pi^1(j) \text{ and } \pi^2(i) < \pi^2(j))\Big)$$

where an **inversion** occurs if the relative ordering of two elements differs between the two permutations.

This distance is useful when assessing how much a permutation **disrupts order**.

3. Combined Distance Measure

Given **row** and **column** distances, we define an aggregated distance:

$$d_{\text{total}} = \alpha \cdot d_{\text{rows}} + \beta \cdot d_{\text{cols}}$$

where:

- d_{rows} is computed using one of the above distances on $\pi^1_{\text{row}}, \pi^2_{\text{row}}$.
- d_{cols} is computed using one of the above distances on $\pi^1_{\text{col}}, \pi^2_{\text{col}}$.
- α, β are weighting parameters (default: $\alpha = \beta = 1$).

This scalar d_{total} summarizes how much **both rows and columns** are permuted.