Combinatorial Bayesian Optimization

Table of contents

0.1 Introduction

We consider a scheduling problem where the schedule is represented by a vector $\mathbf{x} = (x_0, x_1, \dots, x_{T-1})^T$. This vector comprises T components, where x_j denotes the non-negative allocation (e.g., number of patients or tasks) to time slot j, for $j = 0, \dots, T-1$. A fundamental constraint is that the total allocation across all time slots must equal a fixed constant N:

$$\sum_{j=0}^{T-1} x_j = N$$

We require $x_j \geq 0$ for all $j = 0, \dots, T-1$. Consequently, a valid schedule $\mathbf{x}$ belongs to the feasible set $\mathcal{F} = \{\mathbf{z} \in \mathbb{D}^T \mid \sum_{j=0}^{T-1} z_j = N, z_j \geq 0 \text{ for all } j\}$, where $\mathbb{D}$ is the set of non-negative integers ($\mathbb{Z}_{\geq 0}$).

We define a neighborhood structure for local search based on perturbation vectors derived from a set of T basis change vectors, $v_i \in \mathbb{D}^T$, for $i = 0, \dots, T-1$. These basis vectors represent elementary shifts of allocation between time slots:

- $v_0 = (-1, 0, \dots, 0, 1)$ (Shift unit *from* slot 0 *to* slot $T-1$)
- $v_1 = (1, -1, 0, \dots, 0)$ (Shift unit *from* slot 1 *to* slot 0)
- $v_i = (0, \dots, 0, \underbrace{1}_{\text{pos } i-1}, \underbrace{-1}_{\text{pos } i}, 0, \dots, 0)$ for $i = 2, \dots, T-1$ (Shift unit *from* slot i *to* slot $i-1$)

A key property of these basis vectors is that the sum of components for each vector is zero: $\sum_{j=0}^{T-1} v_{ij} = 0$ for all $i = 0, \dots, T-1$.

Perturbations are constructed using a binary selection vector $\mathbf{U} = (u_0, u_1, \dots, u_{T-1})$, where $u_i \in \{0, 1\}$. Each u_i indicates whether the basis change v_i is included in the perturbation. The resulting perturbation vector $\mathbf{r}(\mathbf{U}) \in \mathbb{D}^T$ is the linear combination:

$$\mathbf{r}(\mathbf{U}) := \sum_{i=0}^{T-1} u_i v_i$$

Since each v_i sums to zero, any perturbation $\mathbf{r}(\mathbf{U})$ also sums to zero: $\sum_{j=0}^{T-1} r_j(\mathbf{U}) = 0$. This ensures that applying such a perturbation to a valid schedule $\mathbf{x}$ preserves the total allocation N.

Two specific selection vectors result in a zero perturbation:

1. If $\mathbf{U} = \mathbf{0} = (0, \dots, 0)^T$, then $\mathbf{r}(\mathbf{0}) = \mathbf{0}$.
2. If $\mathbf{U} = \mathbf{1} = (1, \dots, 1)^T$, then $\mathbf{r}(\mathbf{1}) = \sum_{i=0}^{T-1} v_i = \mathbf{0}$, as demonstrated by summing the components of the basis vectors.

The neighborhood of a schedule $\mathbf{x} \in \mathcal{F}$, denoted by $\mathcal{N}(\mathbf{x})$, comprises all distinct, feasible schedules $\mathbf{x}'$ reachable by applying a non-zero perturbation $\mathbf{r}(\mathbf{U})$:

$$\mathcal{N}(\mathbf{x}) := \{\mathbf{x}' \mid \mathbf{x}' = \mathbf{x} + \mathbf{r}(\mathbf{U}), \mathbf{U} \in \{0,1\}^T, \mathbf{r}(\mathbf{U}) \neq \mathbf{0}, \text{ and } x'_j \geq 0 \text{ for all } j = 0, \dots, T-1\}$$

Note that because $\sum r_j(\mathbf{U}) = 0$, any $\mathbf{x}'$ generated from $\mathbf{x} \in \mathcal{F}$ automatically satisfies $\sum x'_j = N$. The condition $\mathbf{r}(\mathbf{U}) \neq \mathbf{0}$ explicitly excludes the transformations resulting from $\mathbf{U} = \mathbf{0}$ and $\mathbf{U} = \mathbf{1}$.

There are 2^T possible selection vectors $\mathbf{U}$. Since $\mathbf{U} = \mathbf{0}$ and $\mathbf{U} = \mathbf{1}$ both yield $\mathbf{r}(\mathbf{U}) = \mathbf{0}$ (assuming $T \geq 1$ so $\mathbf{0} \neq \mathbf{1}$), there are $2^T - 2$ distinct selection vectors that generate non-zero perturbations. This establishes an upper bound on the number of candidate neighbors generated by unique non-zero perturbations:

$$|\mathcal{N}(\mathbf{x})| \leq 2^T - 2$$

The actual size of the neighborhood may be smaller than this bound due to the non-negativity constraint ($\mathbf{x}'_j \geq 0$) rendering some potential neighbors unfeasible.

The local search algorithm aims to iteratively improve the schedule based on an objective function $C(\mathbf{x})$. Given the following constants:

- $\mathbf{x}$: The current feasible schedule vector ($\mathbf{x} \in \mathcal{F}$).
- N: The total number of patients/tasks to be scheduled.
- T: The number of time slots.