Assignment 1 — Column Generation

Koen Moors and Jeroen van Riel

May 2023

1 – Compact Model

Allow us to refer to inequalities $a_i^\top z \leq \beta_i$ as *boundaries* to avoid confusion with inequalities of MIP formulations. We say a point $x \in \mathbb{R}^n$ *satisfies* a boundary if $a_i^\top x \leq \beta_i$ and a point $y \in \mathbb{R}^n$ *violates* a boundary if $a_i^\top y \geq \beta_i + \epsilon$ for some small value of $\epsilon > 0$.

For minimizing the number of boundaries used by the polyhedral classifier, we propose the optimization problem

$$
\begin{aligned}
\min \sum_{i=1}^{K} u_i, & & \text{(1a)}\\
\text{s.t. } a_i^\top x \leq \beta_i, & \quad \text{for all } x \in X, 1 \leq i \leq K, & \text{(1b)}\\
a_i^\top y + s_{yi} M \geq \beta_i + \epsilon, & \quad \text{for all } y \in Y, 1 \leq i \leq K, & \text{(1c)}\\
\sum_{i=1}^{K} s_{yi} < \sum_{i=1}^{K} u_i, & \quad \text{for all } y \in Y, & \text{(1d)}\\
s_{yi} \in \{0, 1\}, & \quad \text{for all } y \in Y, 1 \leq i \leq K, & \text{(1e)}\\
a_i \in \mathbb{R}^n, \beta_i \in \mathbb{R}, u_i \in \{0, 1\}, & \quad \text{for all } 1 \leq i \leq K. & \text{(1f)}
\end{aligned}
$$

Here, K is an upper bound on the minimum order of the polyhedral classifier, u_i an indicator whether boundary i is used by the polyhedral classifier, s_{yi} an indicator whether boundary i is violated by point y and $M \in \mathbb{R}$ a predetermined sufficiently large number.

First of all, the objective is to minimize the number of boundaries that are used by the polyhedral classifier. Inequality (1b) requires that each $x \in X$ satisfies all boundaries and thus fall inside the polyhedron. Similarly, for each combination of $y \in Y$ and $1 \leq i \leq K$ such that $s_{yi} = 0$, inequality (1c) requires that y violates boundary i. However, in case $s_{yi} = 1$, we have that this inequality is always satisfied by setting $M = \infty$. Inequality (1d) requires that each point $y \in Y$ satisfies strictly less boundaries than the total number of used boundaries, which implies that y must violate at least one active boundary and thus fall outside the polyhedron.

2 – Weighted SVM

Suppose every point $y \in Y$ has weight c_y. To find a boundary that separates X from the points of Y with the largest weight, we propose the following MIP.

$$\max \sum_{y \in Y} s_y c_y, \tag{2a}$$

$$\text{s.t. } a^\top x \leq \beta, \quad \text{for all } x \in X, \tag{2b}$$

$$a^\top y + (1 - s_y)M \geq \beta + \epsilon, \quad \text{for all } y \in Y, \tag{2c}$$

$$s_y \in \{0, 1\}, \quad \text{for all } y \in Y, \tag{2d}$$

$$a \in \mathbb{R}^n, \beta \in \mathbb{R}. \tag{2e}$$

Here a and β describe the boundary $a^\top z \leq \beta$ for $z \in \mathbb{R}^n$ and s_y is an indicator whether y violates the boundary. The objective is to maximize the total weight of all points $y \in Y$ that violate the boundary.

Inequality (2b) guarantees that all $x \in X$ satisfy the boundary. Inequality (2c) requires that for each y with $s_y = 1$, we have that y violates the boundary. If $s_y = 0$, then inequality (2c) is always satisfied when the constant M is chosen large enough.

5 – Comparison

Part 1. Suppose there exists a special subset $S \subseteq Y$ with two distinct $y_1, y_2 \in S$ such that $a^\top y_1 > \beta$ and $a^\top y_2 > \beta$. Note that $\text{conv}(\{y_1, y_2\}) = \{\lambda y_1 + (1-\lambda) y_2 : \lambda \in [0, 1]\}$. For every $\lambda \in [0, 1]$, we have

$$a^\top(\lambda y_1 + (1 - \lambda) y_2) = \lambda a^\top y_1 + (1 - \lambda) a^\top y_2 > \lambda\beta + (1 - \lambda)\beta = \beta.$$

Furthermore, we have $a^\top x \leq \beta$ for each $x \in X$. Hence, for each convex combination $x^* = \lambda_1 x_1 + \cdots + \lambda_r x_r \in \text{conv}(X)$ of elements $x_1, \ldots, x_r \in X$, with $\lambda \in \mathbb{R}^n_+$ and $\mathbb{1}^\top \lambda = 1$, we similarly have

$$a^\top(\lambda_1 x_1 + \cdots + \lambda_r x_r) \leq (\lambda_1 + \cdots + \lambda_r)\beta = \beta.$$

This shows that $\text{conv}(\{y_1, y_2\})$ is separated from $\text{conv}(X)$ by the inequality, i.e., $\text{conv}(\{y_1, y_2\}) \cap \text{conv}(X) = \emptyset$, contradicting the assumption that S was special.

Part 2. From the previous part, we know distinct elements $y_1, y_2 \in S$ cannot be separated from X simultaneously, so $\mathcal{I}_{y_1} \cap \mathcal{I}_{y_2} = \emptyset$. Hence, we have

$$\sum_{I \in \mathcal{I}} w_I \geq \sum_{y \in S} \sum_{I \in \mathcal{I}_y} w_I \geq |S|,$$

because we have

$$\sum_{I \in \mathcal{I}_y} w_I \geq 1 \text{ for all } y \in S.$$

This shows that $v^\star_{\text{CG}} \geq |S|$ for any special subsets S, as required.

Part 3. Take $T = \{(2, 0, 0, \ldots), (0, 2, 0, \ldots), (0, 0, 2, \ldots), \ldots\} \subseteq Y$ with $|T| = n$. Take $y_1, y_2 \in T$ such that $y_1 \neq y_2$. Then $\frac{1}{2} y_1 + \frac{1}{2} y_2 \in X$. Therefore $\frac{1}{2} y_1 + \frac{1}{2} y_2 \in \text{conv}(\{y_1, y_2\}) \cap \text{conv}(X)$, so T is special. Hence, by the result of the previous part, we have

$$v^\star_{\text{CG}} \geq \max_{\substack{S \subseteq Y \\ S \text{ special}}} |S| \geq |T| = n.$$