Logistic Regression with `Rmosek`

September 5, 2019

Consider the simplest logistic regression.

$$\max_{\beta} \sum_{i=1}^{n} \left\{ y_i \left(x_i' \beta \right) - \log \left(1 + \exp \left(x_i' \beta \right) \right) \right\} \tag{1}$$

Introduce t_i to replace the softplus function in the objective and $\phi_i = x_i'\beta$, we turned (1) to

$$\begin{aligned} &\max_{t_i, \phi_i, \beta} \sum_{i=1}^{n} \left\{ y_i \phi_i + t_i \right\} \\ &\text{s.t.} - \log \left(1 + \exp \left(\phi_i \right) \right) \geq t_i \\ &\phi_i = x_i' \beta \end{aligned}$$

Further deal with the softplus function in the constraint by noting that

$$-\log \left(1 + \exp \left(\phi_i \right) \right) \geq t_i \Leftrightarrow \exp \left(\phi_i + t_i \right) + \exp \left(t_i \right) \leq 1 \tag{2}$$

Introduce u_i and v_i, (2) can be further transformed to

$$\begin{aligned} u_i + v_i &\leq 1 \\ (u_i, 1, \phi_i + t_i) &\in \mathcal{K}_{\exp} \\ (v_i, 1, t_i) &\in \mathcal{K}_{\exp} \end{aligned}$$

where $\mathcal{K}_{\exp} = \{(x_1, x_2, x_3) : x_1 \geq x_2 \exp \left(x_3 / x_2 \right), \, x_2 > 0\} \cup \{(x_1, 0, x_3) : x_1 \geq 0, x_3 \leq 0\}$. We reach the standard format

$$\begin{aligned} &\max_{\theta} \begin{bmatrix} \mathbf{1}_n & y & 0_{1 \times (2n+p)} \end{bmatrix} \theta \\ \text{s.t.} \quad &\begin{bmatrix} 0_{n \times n} & I_n & 0_{n \times 2n} & -X \\ 0_{n \times 2n} & I_n & I_n & 0_{n \times p} \end{bmatrix} \theta \begin{matrix} = \\ \leq \end{matrix} \begin{bmatrix} 0_{n \times 1} \\ \mathbf{1}_n \end{bmatrix} \\ &(u_i, 1, \phi_i + t_i) \in \mathcal{K}_{\exp} \; i = 1, 2, \cdots, n \\ &(v_i, 1, t_i) \in \mathcal{K}_{\exp} \; i = 1, 2, \cdots, n \end{aligned}$$

where $\theta = (t, \phi, u, v, \beta)$. However it seems that `Rmosek` accepts only the exact variables in the arguments of the exponential cones, which means we cannot have 1 and $\phi_i + t_i$ in the conic constraints as above. We need to further introduce $\alpha = \mathbf{1}_{2n}$ and $\gamma = \phi + t$ to reach the final formulation with $\theta = (t, \phi, u, v, \beta, \alpha, \gamma)$.

$$
\begin{aligned}
&\max_{\theta} \begin{bmatrix} \mathbf{1}_n & y & 0_{1\times(2n+p)} \end{bmatrix} \theta \\
\text{s.t.} \quad & \begin{bmatrix} 0_{n\times n} & I_n & 0_{n\times 2n} & -X & 0_{n\times(3n)} \\ 0_{n\times 2n} & I_n & I_n & & 0_{n\times(p+3n)} \\ I_n & I_n & 0_{n\times(4n+p)} & & -I_n \end{bmatrix} \theta \begin{matrix} = \\ \leq \\ = \end{matrix} \begin{bmatrix} 0_{n\times 1} \\ \mathbf{1}_n \\ 0_{n\times 1} \end{bmatrix} \\
& (u_i, \alpha_i, \gamma_i) \in \mathcal{K}_{\exp} \; i = 1, 2, \cdots, n \\
& (v_j, \alpha_{n+j}, t_j) \in \mathcal{K}_{\exp} \; j = 1, 2, \cdots, n \\
& \alpha = \mathbf{1}_{2n}
\end{aligned}
$$