5 A quadratic programming problem

The question at hand is:

$$\text{Minimize}\beta^t A\beta + \|\beta\|_1 \quad \text{subject to } B\beta = b. \tag{5.1}$$

Here $A \in \mathbb{R}^{6\times 6}$ positive semidefinite, $B \in \mathbb{R}^{6\times 6}$, and $b \in \mathbb{R}^6$ are given. Say

$$A = \operatorname{diag}(1/2, 1/3, 1/4, 1, 2, 3) \quad \text{and} \quad B = \operatorname{diag}(2, 3, 6, 0, 0, 0). \tag{5.2}$$

We are solving for $\beta = (\beta_1, \cdots, \beta_6) \in \mathbb{R}^6$, and

$$\|\beta\|_1 = |\beta_1| + \cdots + |\beta_6|\,. \tag{5.3}$$

To deal with the absolute value, we augment the variable β into $(\beta^+, \beta^-) \in \mathbb{R}^{6+6}$, where we would like

$$\beta_i^+ = \max\{0, \beta_i\} \quad \text{and} \quad \beta_i^- = \max\{0, -\beta_i\} \tag{5.4}$$

so that $\beta = \beta^+ - \beta^-$. Note that in this case we have

$$\|\beta\|_1 = \underbrace{(1, \cdots, 1)^t}_{\text{twelve copies}} \begin{pmatrix} \beta^+ \\ \beta^- \end{pmatrix}. \tag{5.5}$$

Consider the new equivalent problem:

$$\begin{aligned} &\text{Minimize } \begin{pmatrix} \beta^+ \\ \beta^- \end{pmatrix}^t \begin{pmatrix} A & -A \\ -A & A \end{pmatrix} \begin{pmatrix} \beta^+ \\ \beta^- \end{pmatrix} + \mathbf{1}^t \begin{pmatrix} \beta^+ \\ \beta^- \end{pmatrix}. \\ &\text{Subject to} \\ &\quad \begin{pmatrix} B & -B \end{pmatrix} \begin{pmatrix} \beta^+ \\ \beta^- \end{pmatrix} = b \\ &\quad \begin{pmatrix} \beta^+ \\ \beta^- \end{pmatrix} \geq 0 \\ &\quad \beta_i^+ \beta_i^- = 0 \text{ for } i = 1, \cdots, 6. \end{aligned} \tag{5.6}$$

The system is almost easy to work with except for the last condition, which we can reformulate into the following:

- For some $J \subset \{1, \cdots, 6\}$, we have

$$e_j^t \beta^+ = 0 \text{ for } j \in J, \text{ and } e_j^t \beta^- = 0 \text{ for } j \in \{1, \cdots, 6\} \setminus J. \tag{5.7}$$

We will exhaust all these possible J in our implementation of the algorithm. There are $2^6 = 64$ possibilities. We can store J as a vector with 6 entries, each of which is either 0 or 1.

For convenience, set

$$\hat{\beta} = \begin{pmatrix} \beta^+ \\ \beta^- \end{pmatrix}, \quad \hat{A} = \begin{pmatrix} A & -A \\ -A & A \end{pmatrix}, \quad \hat{B} = \begin{pmatrix} B & -B \end{pmatrix} \in \mathbb{R}^{6\times 12}. \tag{5.8}$$

Let $\{\hat{e}_i : i = 1, \cdots, 12\}$ be the standard basis for $\mathbb{R}^{6+6}$.

Fix $J_* \subset \{1, \cdots, 6\}$. The Karush-Kuhn-Tucker (KKT) conditions associated with this particular J_* are specified by:

$$\begin{aligned}
2\hat{A}\hat{\beta} - \sum_{i=1}^{12} \mu_i \hat{e}_i + \sum_{k=1}^{6} \lambda_k \hat{b}_k^t + \sum_{j\in J_*} v_j \hat{e}_j + \sum_{j\in\{1,\cdots,6\}\setminus J_*} v_j \hat{e}_{j+6} = \mathbf{0} \in \mathbb{R}^{12}. \\
\hat{\beta} \geq \mathbf{0} \\
\hat{B}\hat{\beta} - b = \mathbf{0}_6 \\
\hat{e}_j^t \hat{\beta} = 0 \text{ for } j \in J_* \\
\hat{e}_{j+6}^t \hat{\beta} = 0 \text{ for } j \in \{1, \cdots, 6\} \setminus J_* \\
\mu_k \geq 0 \text{ for } k = 1, \cdots, 12 \\
\sum_{k=1}^{12} \mu_k (\hat{e}_k^t \hat{\beta}) = 0.
\end{aligned} \tag{5.9}$$

Here, we are simultaneously solving for

$$(\hat{\beta}; \vec{\mu}; \vec{\lambda}; \vec{v}) := (\hat{\beta} \quad ; \quad \mu_1, \cdots, \mu_{12} \quad ; \quad \lambda_1, \cdots, \lambda_6 \quad ; \quad v_1, \cdots, v_6) \in \mathbb{R}^{36}.$$

Using a loop, we solve the above system for all the possible $J \subset \{1, \cdots, 6\}$ and compare the answer. We find the $\hat{\beta}_*$ the corresponds to the leave value and the corresponding J_*. This will give us the best β in the original problem.

Let us put (5.9) in an equivalent matrix form:

$$\begin{aligned}
\begin{pmatrix} 2\hat{A} & -I_{12} & \hat{B}^t & E_* \end{pmatrix} \cdot (\hat{\beta}; \vec{\mu}; \vec{\lambda}; \vec{v}) = \mathbf{0} \in \mathbb{R}^{12} \\
\begin{pmatrix} I_{12} & 0 & 0 & 0 \end{pmatrix} \cdot (\hat{\beta}; \vec{\mu}; \vec{\lambda}; \vec{v}) \geq \mathbf{0} \in \mathbb{R}^{12}. \\
\begin{pmatrix} \hat{B} & 0 & 0 & 0 \end{pmatrix} \cdot (\hat{\beta}; \vec{\mu}; \vec{\lambda}; \vec{v}) = b \in \mathbb{R}^{6} \\
\begin{pmatrix} E_* & 0 & 0 & 0 \end{pmatrix} \cdot (\hat{\beta}; \vec{\mu}; \vec{\lambda}; \vec{v}) = \mathbf{0} \in \mathbb{R}^{12} \\
\begin{pmatrix} 0 & I_{12} & 0 & 0 \end{pmatrix} \cdot (\hat{\beta}; \vec{\mu}; \vec{\lambda}; \vec{v}) \geq \mathbf{0} \in \mathbb{R}^{12} \\
\begin{pmatrix} 0 & 0 & E_* & 0 \end{pmatrix} \cdot (\hat{\beta}; \vec{\mu}; \vec{\lambda}; \vec{v}) = \mathbf{0} \in \mathbb{R}^{12}.
\end{aligned} \tag{5.10}$$

In the system above, $E_* \in \mathbb{R}^{12\times 12}$ is a diagonal matrix corresponding to the fixed index set $J_* \subset \{1, \cdots, 6\}$, given by

$$E_* = \operatorname{diag}(j_1, \cdots, j_6, j_7, \cdots, j_{12}) \text{ where } j_\ell = \begin{cases} 1 & \text{if } 1 \leq \ell \leq 6 \text{ and } \ell \in J_* \\ 1 & \text{if } 7 \leq \ell \leq 12 \text{ and } \ell - 6 \notin J_* \\ 0 & \text{otherwise} \end{cases}. \tag{5.11}$$

For instance, if $J_* = \{1, 3, 4\}$, then

$$E_* = \mathrm{diag}(1, 0, 1, 1, 0, 0; 0, 1, 0, 0, 1, 1).$$

Notice that the fourth and the fifth conditions of (5.9) simply say certain entries of $\hat{\beta}$ must vanish (as expected). Combined with the second and the sixth conditions, we see that the last condition of (5.9) says that each corresponding entries of $\vec{\mu}$ and $\hat{\beta}$ cannot be both nonzero. Hence the last condition in (5.11).