Nonlinear Programming
3rd Edition

Theoretical Solutions Manual

Chapter 6

Dimitri P. Bertsekas

Massachusetts Institute of Technology

Athena Scientific, Belmont, Massachusetts

NOTE

This manual contains solutions of the theoretical problems, marked in the book by   It is continuously updated and improved, and it is posted on the internet at the book's www page

http://www.athenasc.com/nonlinbook.html

Many thanks are due to several people who have contributed solutions, and particularly to David Brown, Angelia Nedic, Asuman Ozdaglar, Cynara Wu.

Last Updated: May 2016

Solutions Chapter 6

SECTION 6.1

6.1.4

Throughout this exercise we will use the fact that strong duality holds for convex quadratic problems with linear constraints (cf. Section 4.4).

The problem of finding the minimum distance from the origin to a line is written as

$$\min \frac{1}{2}\|x\|^2$$

$$\text{subject to } Ax = b$$

where A is a 2×3 matrix with full rank, and $b \in \Re^2$. Let f^* be the optimal value and consider the dual function

$$q(\lambda) = \min_x \left\{ \frac{1}{2}\|x\|^2 + \lambda'(Ax - b) \right\}.$$

Let V^* be the supremum over all distances of the origin from planes that contain the line $\{x \mid Ax = b\}$. Clearly, we have $V^* \leq f^*$, since the distance to the line $\{x \mid Ax = b\}$ cannot be smaller that the distance to the plane that contains the line.

We now note that any plane of the form $\{x \mid p'Ax = p'b\}$, where $p \in \Re^2$, contains the line $\{x \mid Ax = b\}$, so we have for all $p \in \Re^2$,

$$V(p) \equiv \min_{p'Ax = p'x} \frac{1}{2}\|x\|^2 \leq V^*.$$

On the other hand, by duality in the minimization of the preceding equation, we have

$$U(p, \gamma) \equiv \min_x \left\{ \frac{1}{2}\|x\|^2 + \gamma(p'Ax - p'x) \right\} \leq V(p), \qquad \forall\, p \in \Re^2,\ \gamma \in \Re.$$

Combining the preceding relations, it follows that

$$\sup_\lambda q(\lambda) = \sup_{p,\gamma} U(p, \gamma) \leq \sup_p U(p, 1) \leq \sup_p V(p) \leq V^* \leq f^*.$$

Since by duality in the original problem, we have $\sup_\lambda q(\lambda) = f^*$, it follows that equality holds throughout above. Hence $V^* = f^*$, which was to be proved.

6.1.7 (Duality Gap of the Knapsack Problem)

(a) Let

$$x_i = \begin{cases} 1 & \text{if the } i\text{th object is included in the subset,} \\ 0 & \text{otherwise.} \end{cases}$$

Then the total weight and value of objects included is $\sum_i w_i x_i$ and $\sum_i v_i x_i$, respectively, and the problem can be written as

$$\begin{aligned} &\text{maximize} \quad f(x) = \sum_{i=1}^n v_i x_i \\ &\text{subject to} \ \sum_{i=1}^n w_i x_i \leq A, \quad x_i \in \{0,1\}, \quad i = 1,\ldots,n. \end{aligned}$$

(b) Let $\tilde{f}(x) = -f(x)$ and consider the equivalent problem of minimizing $\tilde{f}(x)$ subject to the constraints given above. Then

$$L(x,\mu) = -\sum_{i=1}^n v_i x_i + \mu\left(\sum_{i=1}^n w_i x_i - A\right),$$

and

$$\tilde{q}(\mu) = \inf_{x_i \in \{0,1\}} \left\{ \sum_{i=1}^n (\mu w_i - v_i) x_i - \mu A \right\}.$$

Note that the minimization above is a separable problem, and the infimum is attained at

$$\tilde{x}_i(\mu) = \begin{cases} 0 & \text{if } \mu > v_i/w_i, \\ 1 & \text{if } \mu < v_i/w_i, \\ 0 \text{ or } 1 & \text{if } \mu = v_i/w_i. \end{cases}$$

Without loss of generality, assume that the objects are ordered such that $\frac{v_1}{w_1} \leq \frac{v_2}{w_2} \leq \ldots \leq \frac{v_n}{w_n}$. When $\mu \in \left(\frac{v_{j-1}}{w_{j-1}}, \frac{v_j}{w_j}\right]$ for some j with $1 \leq j \leq n$, then $\tilde{x}_i(\mu) = 1$ for all $i \geq j$ and $\tilde{x}_i(\mu) = 0$ otherwise, and

$$\tilde{q}(\mu) = \mu\left(\sum_{i=j}^n w_i - A\right) - \sum_{i=j}^n v_i.$$

From this relation, we see that, as μ increases, the slope of $\tilde{q}(\mu)$ decreases from $\sum_{i=1}^n w_i - A$ to $-A$. Therefore, if $\sum_{i=1}^n w_i - A > 0$, $\tilde{q}(\mu)$ is maximized when the slope of the curve goes from positive to negative. In this case, the dual optimal value $\tilde{q}^*$ is attained at $\mu^* = \frac{v_{i^*}}{w_{i^*}}$, where i^* is the largest i such that

$$w_i + \ldots + w_n \geq A.$$

If $\sum_{i=1}^n w_i - A \leq 0$, then the dual optimal value $\tilde{q}^*$ is attained at $\mu^* = 0$.

(c) Consider a relaxed version of the problem of minimizing $\tilde{f}$:

$$\begin{aligned} &\text{minimize} \quad f_R(x) = -\sum_{i=1}^n v_i x_i \\ &\text{subject to} \quad \sum_{i=1}^n w_i x_i \le A, \qquad x_i \in [0,1], \quad i = 1, \ldots, n. \end{aligned}$$

Let f_R^* and q_R^* be the optimal values of the relaxed problem and its dual, respectively. In the relaxed problem, the cost function is convex over $\Re^n$ (in fact it is linear), and the constraint set is polyhedral. Thus, according to Prop. 6.2.1, there is no duality gap, and $f_R^* = q_R^*$. The dual function of the relaxed problem is

$$q_R(\mu) = \inf_{x_i \in [0,1]} \left\{ \sum_{i=1}^n (\mu w_i - v_i) x_i - \mu A + \right\}.$$

Again, $q_R(\mu)$ is separable and the infimum is attained at

$$x_i(\mu) = \begin{cases} 0 & \text{if } \mu > v_i/w_i, \\ 1 & \text{if } \mu < v_i/w_i, \\ \text{anything in } [0,1] & \text{if } \mu = v_i/w_i. \end{cases}$$

Thus the solution is the same as the $\{0,1\}$ constrained problem for all i with $\mu \neq v_i/w_i$. For i with $\mu = v_i/w_i$, the value of x_i is irrelevant to the dual function value. Therefore, $\tilde{q}(\mu) = q_R(\mu)$ for all μ, and thus $\tilde{q}^* = q_R^*$.

Following Example 6.1.2, it can be seen that the optimal primal and dual solution pair (x^*, μ^*) of the relaxed problem satisfies

$$\mu^* w_i = v_i, \quad \text{if } 0 < x_i^* < 1,$$

$$\mu^* w_i \ge v_i, \quad \text{if } x_i^* = 0,$$

$$\mu^* w_i \le v_i, \quad \text{if } x_i^* = 1.$$

In fact, it is straightforward to show that there is an optimal solution of the relaxed problem such that at most one x_i^* satisfies $0 < x_i^* < 1$. Consider a solution $\bar{x}$ equivalent to this optimal solution with the exception that $\bar{x}_i = 0$ if $0 < x_i^* < 1$. This solution is clearly feasible for the $\{0,1\}$ constrained problem, so that we have

$$\tilde{f}^* \le \tilde{f}(\bar{x}) \le f_R^* + \max_{1 \le i \le n} v_i.$$

Combining with earlier results,

$$\tilde{f}^* \le q_R^* + \max_{1 \le i \le n} v_i = \tilde{q}^* + \max_{1 \le i \le n} v_i.$$

Since $\tilde{f}^* = -f^*$ and $\tilde{q}^* = -q^*$, we have the desired result.

(d) Since the object weights and values remain the same we have from (c) that $0 \leq q^*(k) - f^*(k) \leq \max_{1 \leq i \leq n} v_i$. By including in the subset each replica of an object assembled in the optimal solution to the original problem, we see that $f^*(k) \geq k f^*$. It then follows that

$$\lim_{k \to \infty} \frac{q^*(k) - f^*(k)}{f^*(k)} = 0.$$

6.1.8 (Sensitivity)

We have

$$\overline{f} = \inf_{x \in X} \big\{ f(x) + \overline{\mu}'\big(g(x) - \overline{u}\big) \big\},$$

$$\tilde{f} = \inf_{x \in X} \big\{ f(x) + \tilde{\mu}'\big(g(x) - \tilde{u}\big) \big\},$$

from which

$$\begin{aligned} \overline{f} - \tilde{f} &= \inf_{x \in X} \big\{ f(x) + \overline{\mu}'\big(g(x) - \overline{u}\big) \big\} - \inf_{x \in X} \big\{ f(x) + \tilde{\mu}'\big(g(x) - \overline{u}\big) \big\} + \tilde{\mu}'(\tilde{u} - \overline{u}) \\ &\geq \tilde{\mu}'(\tilde{u} - \overline{u}), \end{aligned}$$

where the last inequality holds because $\overline{\mu}$ is a dual-optimal solution of the problem

$$\begin{aligned} &\text{minimize} \quad f(x) \\ &\text{subject to} \quad x \in X, \qquad g(x) \leq \overline{u}, \end{aligned}$$

so that it maximizes over $\mu \geq 0$ the dual value $\inf_{x \in X} \big\{ f(x) + \mu'\big(g(x) - \overline{u}\big) \big\}$.

This proves the left-hand side of the desired inequality. Interchanging the roles of $\overline{f}$, $\overline{u}$, $\overline{\mu}$, and $\tilde{f}$, $\tilde{u}$, $\tilde{\mu}$, shows the desired right-hand side.

6.1.9 (Hoffman's Bound)

(a) Since the feasible set is closed, the projection problem given in the exercise has an optimal solution. Therefore by Proposition 3.4.2, there exists an optimal primal solution and geometric multiplier pair $(x^*(y,z), \mu^*(y,z))$ for each $y \in Y, z \in \Re^n$. By the optimality condition, for μ to be a geometric multiplier, it suffices that

$$-\frac{z - x^*}{\|z - x^*\|} = \sum_{i \in I(x^*(y,z))} \mu_i a_i,$$

where $I(x^*(y,z))$ is the set of indexes corresponding to the active constraints at $x^*(y,z)$, and a_i is the ith column vector of A'. Since the vector in the left-hand-side of the equation has norm

1, we can pick $\mu^*(y,z)$ to be the minimum norm solution for this linear equation. Since there are only a finite number of such equations, the set $\{\mu^*(y) \mid y \in Y\}$ is bounded. Finally, by the optimality of a geometric multiplier, we have

$$f^*(y,z) \le \|z - z\| + \mu^*(y,z)'(Az - b - y) \le \mu^*(y,z)'(Az - b - y)^+.$$

(b) Using the last inequality in part (a), we have

$$f^*(y,z) \le \sum_i \mu_i^*(y,z)\|(Az - b - y)^+\|, \quad \forall y \in Y, z \in Z,$$

and

$$f^*(y,z) \le \max_i \mu_i^*(y,z)\, \|(Az - b - y)^+\|, \quad \forall y \in Y, z \in \Re^n.$$

Let c be a constant such that $c > \{\mu^*(y) \mid y \in Y\}$. By the boundedness of $\{\mu^*(y) \mid y \in Y\}$ shown in part (a), we can choose c so that the bound $f^*(y,z) \le c\|(Ax - b - y)^+\|$ holds.

6.1.10 (Upper Bounds to the Optimal Dual Value [Ber99])

We consider the subset of $\Re^{r+1}$

$$A = \big\{(z,w) \mid \text{ there exists } x \in X \text{ such that } \ g(x) \le z,\ f(x) \le w\big\},$$

and its convex hull $\mathrm{Conv}(A)$. The vectors $\big(g(x_F), f(x_F)\big)$ and $\big(g(x_I), f(x_I)\big)$ belong to A. In addition, the vector $(0, \tilde{f})$, where

$$\tilde{f} = \inf\big\{w \mid (z,w) \in \mathrm{Conv}(A),\ z \le 0\big\},$$

is in the closure of $\mathrm{Conv}(A)$. Let us now show that $q^* \le \tilde{f}$, as indicated by Fig. 1.

Indeed, for each $(z,w) \in \mathrm{Conv}(A)$, there exist $\xi_1 \ge 0$ and $\xi_2 \ge 0$ with $\xi_1 + \xi_2 = 1$, and $x_1 \in X$, $x_2 \in X$ such that

$$\xi_1 g(x_1) + \xi_2 g(x_2) \le z,$$

$$\xi_1 f(x_1) + \xi_2 f(x_2) \le w.$$

Furthermore, by the definition of the dual function q, we have for all $\mu \in \Re^r$,

$$q(\mu) \le f(x_1) + \mu' g(x_1),$$

$$q(\mu) \le f(x_2) + \mu' g(x_2).$$

Combining the preceding four inequalities, we obtain

$$q(\mu) \leq w + \mu' z, \qquad \forall\ (z, w) \in \mathrm{Conv}(A), \quad \mu \geq 0.$$

The above inequality holds also for all (z, w) that are in the closure of $\mathrm{Conv}(A)$, and in particular, for $(z, w) = (0, \tilde{f})$. It follows that

$$q(\mu) \leq \tilde{f}, \qquad \forall\ \mu \geq 0,$$

from which, by taking the maximum over $\mu \geq 0$, we obtain $q^* \leq \tilde{f}$.

Let γ be any nonnegative scalar such that $g(x_I) \leq -\gamma g(x_F)$, and consider the vector

$$\Delta = -\gamma g(x_F) - g(x_I).$$

Since $\Delta \geq 0$, it follows that the vector

$$\big(-\gamma g(x_F), f(x_I)\big) = \big(g(x_I) + \Delta, f(x_I)\big)$$

also belongs to the set A. Thus the three vectors

$$\big(g(x_F), f(x_F)\big), \quad (0, \tilde{f}), \quad \big(-\gamma g(x_F), f(x_I)\big)$$

belong to the closure of $\mathrm{Conv}(A)$, and form a triangle in the plane spanned by the "vertical" vector $(0, 1)$ and the "horizontal" vector $\big(g(x_F), 0\big)$.

Let $(0, \hat{f})$ be the intersection of the vertical axis with the line segment connecting the vectors $\big(g(x_F), f(x_F)\big)$ and $\big(-\gamma g(x_F), f(x_I)\big)$ (there is a point of intersection because $\gamma \geq 0$). We have by Euclidean triangle geometry (cf. Fig. 1)

$$\frac{\hat{f} - f(x_I)}{f(x_F) - f(x_I)} = \frac{\gamma}{\gamma + 1}. \tag{1}$$

Since the vectors $\big(g(x_F), f(x_F)\big)$ and $\big(-\gamma g(x_F), f(x_I)\big)$ both belong to $\mathrm{Conv}(A)$, we also have $(0, \hat{f}) \in \mathrm{Conv}(A)$. By the definition of $\tilde{f}$, we obtain $\tilde{f} \leq \hat{f}$, and since $q^* \leq \tilde{f}$, as shown earlier, from Eq. (1) we have

$$\frac{q^* - f(x_I)}{f(x_F) - f(x_I)} \leq \frac{\tilde{f} - f(x_I)}{f(x_F) - f(x_I)} \leq \frac{\gamma}{\gamma + 1}.$$

Taking the infimum over $\gamma \geq 0$, the desired error bound follows.

Figure for Exercise 6.1.10 Geometrical interpretation of the bound of Exercise 6.1.10 in the case where there is only one constraint. We consider the convex hull of the subset A of $\Re^2$ given by

$$A = \Big\{(z, w) \mid \text{ there exists } x \in X \text{ such that } \ g(x) \le z,\ f(x) \le w\Big\}.$$

Let $\hat{f}$ be the point of intersection of the vertical axis of $\Re^2$ with the line segment connecting the vectors $\big(g(x_F), f(x_F)\big)$ and $\big(g(x_I), f(x_I)\big)$. The vector $(0, \hat{f})$ belongs to $\text{Conv}(A)$. Also, by Euclidean geometry, we have

$$\frac{\hat{f} - f(x_I)}{f(x_F) - f(x_I)} = \frac{g(x_I)}{g(x_I) - g(x_F)},$$

and by the definition of q^* we have

$$q^* \le \tilde{f} \le \hat{f},$$

where

$$\tilde{f} = \inf\Big\{w \mid (z, w) \in \text{Conv}(A),\ z \le 0\Big\}.$$

Combining these two relations, the bound follows.

6.1.11

Let f^* be the (common) optimal value of the two problems

$$\begin{aligned} &\text{minimize} \ \ f(x) \\ &\text{subject to} \ \ x \in X, \qquad g_j(x) \le 0, \quad j = 1, \ldots, r, \end{aligned} \tag{1.1}$$

and

$$\begin{aligned} &\text{minimize} \ \ f(x) \\ &\text{subject to} \ \ x \in \overline{X}, \qquad g_j(x) \le 0, \quad j \in J, \end{aligned} \tag{1.2}$$

where

$$\overline{X} = \Big\{x \in X \mid g_j(x) \le 0,\ j \in \overline{J}\Big\}.$$

Since $\{\mu_j^* \mid j \in J\}$ is a geometric multiplier of problem (1.2), we have

$$f^* = \inf_{x \in \overline{X}} \left\{ f(x) + \sum_{j \in J} \mu_j^* g_j(x) \right\}. \tag{1.3}$$

Since the problem

$$\begin{aligned} &\text{minimize} \quad \left\{ f(x) + \sum_{j \in J} \mu_j^* g_j(x) \right\} \\ &\text{subject to} \;\; x \in X, \qquad g_j(x) \le 0, \quad j \in \overline{J}, \end{aligned} \tag{1.4}$$

has no duality gap, we have

$$\inf_{x \in \overline{X}} \left\{ f(x) + \sum_{j \in J} \mu_j^* g_j(x) \right\} = \sup_{\mu_j \ge 0,\; j \in \overline{J}} \inf_{x \in X} \left\{ f(x) + \sum_{j \in J} \mu_j^* g_j(x) + \sum_{j \in \overline{J}} \mu_j g_j(x) \right\}. \tag{1.5}$$

Combining Eqs. (1.3) and (1.5), we have

$$f^* = \sup_{\mu_j \ge 0,\; j \in \overline{J}} \inf_{x \in X} \left\{ f(x) + \sum_{j \in J} \mu_j^* g_j(x) + \sum_{j \in \overline{J}} \mu_j g_j(x) \right\},$$

which can be written as

$$f^* = \sup_{\mu_j \ge 0,\; j \in \overline{J}} q\big(\{\mu_j^* \mid j \in J\}, \{\mu_j \mid j \in \overline{J}\}\big),$$

where q is the dual function of problem (1.1). It follows that problem (1.1) has no duality gap.

If $\{\mu_j^* \mid j \in \overline{J}\}$ is a geometric multiplier for problem (1.4), we have

$$\inf_{x \in \overline{X}} \left\{ f(x) + \sum_{j \in J} \mu_j^* g_j(x) \right\} = \inf_{x \in X} \left\{ f(x) + \sum_{j \in J} \mu_j^* g_j(x) + \sum_{j \in \overline{J}} \mu_j^* g_j(x) \right\},$$

which together with Eq. (1.3), implies that

$$f^* = \inf_{x \in X} \left\{ f(x) + \sum_{j \in J} \mu_j^* g_j(x) + \sum_{j \in \overline{J}} \mu_j^* g_j(x) \right\},$$

or that $\{\mu_j^* \mid j = 1, \ldots, r\}$ is a geometric multiplier for the original problem (1.1).

6.1.12 (Extended Representation)

(a) Consider the problems

$$\begin{aligned} &\text{minimize} \quad f(x) \\ &\text{subject to} \;\; x \in X, \qquad g_j(x) \le 0, \quad j = 1, \ldots, r, \end{aligned} \tag{1.6}$$

and

$$
\begin{aligned}
&\text{minimize} \quad f(x) \\
&\text{subject to} \quad x \in \tilde{X}, \qquad g_j(x) \le 0, \quad j = 1, \ldots, r, \qquad \tilde{g}_j(x) \le 0, \quad j = 1, \ldots, \tilde{r},
\end{aligned}
\tag{1.7}
$$

where

$$X = \tilde{X} \cap \big\{x \mid \tilde{g}_j(x) \le 0,\, j = 1, \ldots, \tilde{r}\big\},$$

and let f^* be their (common) optimal value.

If problem (1.7) has no duality gap, we have

$$
\begin{aligned}
f^* &= \sup_{\mu \ge 0,\, \tilde{\mu} \ge 0} \inf_{x \in \tilde{X}} \left\{ f(x) + \sum_{j=1}^{r} \mu_j g_j(x) + \sum_{j=1}^{\tilde{r}} \tilde{\mu}_j \tilde{g}_j(x) \right\} \\
&\le \sup_{\mu \ge 0,\, \tilde{\mu} \ge 0} \inf_{\substack{x \in \tilde{X} \\ \tilde{g}_j(x) \le 0,\, j=1,\ldots,\tilde{r}}} \left\{ f(x) + \sum_{j=1}^{r} \mu_j g_j(x) + \sum_{j=1}^{\tilde{r}} \tilde{\mu}_j \tilde{g}_j(x) \right\} \\
&= \sup_{\mu \ge 0,\, \tilde{\mu} \ge 0} \inf_{x \in X} \left\{ f(x) + \sum_{j=1}^{r} \mu_j g_j(x) + \sum_{j=1}^{\tilde{r}} \tilde{\mu}_j \tilde{g}_j(x) \right\} \\
&\le \sup_{\mu \ge 0,\, \tilde{\mu} \ge 0} \inf_{x \in X} \left\{ f(x) + \sum_{j=1}^{r} \mu_j g_j(x) \right\} \\
&= \sup_{\mu \ge 0} \inf_{x \in X} \left\{ f(x) + \sum_{j=1}^{r} \mu_j g_j(x) \right\} \\
&= \sup_{\mu \ge 0} q(\mu),
\end{aligned}
$$

where q is the dual function of problem (1.6). Therefore, problem (1.6) has no duality gap.

(b) If $\mu^* = \{\mu_j^* \mid j = 1, \ldots, r\}$, and $\tilde{\mu}^* = \{\mu_j^* \mid j = 1, \ldots, \tilde{r}\}$, are geometric multipliers for problem (1.7), we have

$$
\begin{aligned}
f^* &= \inf_{x \in \tilde{X}} \left\{ f(x) + \sum_{j=1}^{r} \mu_j^* g_j(x) + \sum_{j=1}^{\tilde{r}} \tilde{\mu}_j^* \tilde{g}_j(x) \right\} \\
&\le \inf_{\substack{x \in \tilde{X} \\ \tilde{g}_j(x) \le 0,\, j=1,\ldots,\tilde{r}}} \left\{ f(x) + \sum_{j=1}^{r} \mu_j^* g_j(x) + \sum_{j=1}^{\tilde{r}} \tilde{\mu}_j^* \tilde{g}_j(x) \right\} \\
&= \inf_{x \in X} \left\{ f(x) + \sum_{j=1}^{r} \mu_j^* g_j(x) + \sum_{j=1}^{\tilde{r}} \tilde{\mu}_j^* \tilde{g}_j(x) \right\} \\
&\le \inf_{x \in X} \left\{ f(x) + \sum_{j=1}^{r} \mu_j^* g_j(x) \right\} \\
&= \inf_{x \in X} \left\{ f(x) + \sum_{j=1}^{r} \mu_j^* g_j(x) \right\} \\
&= q(\mu^*).
\end{aligned}
$$

It follows that μ^* is a geometric multiplier for problem (1.6).

SECTION 6.2

6.2.2 (A Stronger Version of the Duality Theorem)

Without loss of generality, we may assume that there are no equality constraints, so that the problem is

$$\begin{aligned}&\text{minimize} \;\; f(x)\\ &\text{subject to} \;\; x \in X, \quad a_j'x - b_j \le 0,\ j = 1, \ldots, r.\end{aligned}$$

Let $X = C \cap P$, and let the polyhedron P be described in terms of linear inequalities as

$$P = \{x \in \Re^n \mid a_j'x - b_j \le 0,\ j = r+1, \ldots, p\},$$

where p is an integer with $p > r$. By applying Lemma 6.2.2 with

$$S = \{x \in \Re^n \mid a_j'x - b_j \le 0,\ j = 1, \ldots, p\},$$

and $F(x) = f(x) - f^*$, we have that there exist scalars $\mu_i \ge 0$, $j = 1, \ldots, p$, such that

$$f^* \le f(x) + \sum_{j=1}^{p} \mu_j(a_j'x - b_j), \qquad \forall\ x \in C.$$

For any $x \in X$ we have $\mu_j(a_j'x - b_j) \le 0$ for all $j = r+1, \ldots, p$, so the above relation implies that

$$f^* \le f(x) + \sum_{j=1}^{r} \mu_j(a_j'x - b_j), \qquad \forall\ x \in X,$$

or equivalently

$$f^* \le \inf_{x \in X}\{f(x) + \sum_{j=1}^{r} \mu_j(a_j'x - b_j)\} = q(\mu) \le q^*.$$

By using the weak duality theorem (Prop. 6.1.3), it follows that μ is a Lagrange multiplier and that there is no duality gap.

In Example 6.2.1, we can set $C = \{x \in \Re^2 \mid x \ge 0\}$ and $P = \{x \in \Re^2 \mid x_1 \ge 0\}$. Then evidently $X = C$ and f is convex over C. However, $ri(C) = int(C) = \{x \in \Re^2 \mid x > 0\}$, while every feasible point x must have $x_1 = 0$. Hence no feasible point belongs to the relative interior of C, and as seen in Example 6.2.1, there is a duality gap.

SECTION 6.3

6.3.1 (Boundedness of the Set of Geometric Multipliers)

Assume that there exists an $\overline{x} \in X$ such that $g_j(\overline{x}) < 0$ for all j. By Prop. 6.3.1, the set of Lagrange multipliers is nonempty. Let μ be any Lagrange multiplier. By assumption, $-\infty < f^*$, and we have

$$-\infty < f^* \leq L(\bar{x}, \mu) = f(\bar{x}) + \sum_{j=1}^{r} \mu_j g_j(\bar{x}),$$

or

$$-\sum_{j=1}^{r} \mu_j g_j(\bar{x}) \leq f(\bar{x}) - f^*.$$

We have

$$\min_{i=1,\ldots,r} \{-g_i(\bar{x})\} \leq -g_j(\bar{x}), \qquad \forall\ j,$$

so by combining the last two relations, we obtain

$$\left(\sum_{j=1}^{r} \mu_j\right) \min_{i=1,\ldots,r} \{-g_i(\bar{x})\} \leq f(\bar{x}) - f^*.$$

Since $\bar{x}$ satisfies $g_j(\bar{x}) < 0$ for all j, we have

$$\sum_{j=1}^{r} \mu_j \leq \frac{f(\bar{x}) - f^*}{\min_{j=1,\ldots,r}\{-g_j(\bar{x})\}}.$$

Hence the set of Lagrange multipliers is bounded.

Conversely, let the set of Lagrange multipliers be nonempty and bounded. Consider the set

$$B = \big\{z \mid \text{there exists } x \in X \text{ such that } g(x) \leq z\big\}.$$

Assume, to arrive at a contradiction, that there is no $\overline{x} \in X$ such that $g(\overline{x}) < 0$. Then the origin is not an interior point of B, and similar to the proof of Prop. 6.3.1, we can show that B is convex, and that there exists a hyperplane whose normal γ satisfies $\gamma \neq 0$, $\gamma \geq 0$, and

$$\gamma' g(x) \geq 0, \qquad \forall\ x \in X. \tag{1}$$

Let now μ be a Lagrange multiplier. Using Eq. (1), we have for all $\beta \geq 0$

$$f^* = \inf_{x \in X} L(x, \mu) \leq \inf_{x \in X} L(x, \mu + \beta\gamma) \leq \inf_{x \in X,\, g(x) \leq 0} L(x, \mu + \beta\gamma) \leq \inf_{x \in X,\, g(x) \leq 0} f(x) = f^*,$$

where the last inequality holds because $\mu + \beta\gamma \geq 0$, and hence $(\mu + \beta\gamma)' g(x) \leq 0$ if $g(x) \leq 0$. Hence, equality holds throughout in the above relation, so $\mu + \beta\gamma$ is a Lagrange multiplier for all $\beta \geq 0$. Since $\gamma \neq 0$, it follows that the set of Lagrange multipliers is unbounded – a contradiction.

6.3.2 (Optimality Conditions for Nonconvex Cost)

(a) Since the constraint set $\overline{X} = \{x \mid x \in X,\ g_j(x) \leq 0,\ j = 1, \ldots, r\}$ is convex, and x^* is a local minimum, we have

$$\nabla f(x^*)'(x - x^*) \geq 0, \qquad \forall\ x \in \overline{X}$$

(see Prop. 2.1.2 of Chapter 2). Hence x^* is a local minimum of the problem

$$\begin{aligned} &\text{minimize } \nabla f(x^*)'x \\ &\text{subject to } x \in X, \quad g_j(x) \leq 0, \quad j = 1, \ldots, r. \end{aligned} \tag{1}$$

he Assumption 6.3.1 holds for problem (1), so that we can apply Prop. 6.3.1. Thus we have that there is no duality gap, and there exists a Lagrange multiplier $\mu^* \geq 0$ for problem (1), i.e.

$$\sup_{\mu \geq 0} q(\mu) = q(\mu^*) = \inf_{x \in X} \left\{ \nabla f(x^*)'x + \sum_{j=1}^{r} \mu_j^* g_j(x) \right\} = \nabla f(x^*)'x^*.$$

From Prop. 6.1.1, we also obtain

$$\mu_j^* g_j(x^*) = 0, \qquad \forall\ j.$$

The last two relations imply that

$$x^* \in \arg\min_{x \in X} \left\{ \nabla f(x^*)'x + \sum_{j=1}^{r} \mu_j^* g_j(x) \right\}. \tag{2}$$

(b) We use Prop. 4.3.12 to assert that there exist $\mu_j^* \geq 0$, $j = 1, \ldots, r$, such that $\mu_j^* g_j(x^*) = 0$ for all j and

$$\nabla_x L(x^*, \mu^*)'(x - x^*) \geq 0, \qquad \forall\ x \in X.$$

The last relation implies that

$$x^* = \arg\min_{x \in X} \nabla_x L(x^*, \mu^*)'x.$$

6.3.3 (Boundedness of the Set of Lagrange Multipliers for Nonconvex Constraints)

For simplicity and without loss of generality, assume that $A(x^*) = \{1, \ldots, r\}$, and denote

$$h_j(x) = \nabla g_j(x^*)'(x - x^*), \qquad \forall\ j.$$

By Prop. 6.1.1, $\mu \in M^*$ if and only if x^* is a global minimum of the convex problem

$$\begin{aligned} &\text{minimize } \nabla f(x^*)'(x - x^*) \\ &\text{subject to } x \in X, \quad h_j(x) \le 0, \quad j = 1, \ldots, r, \end{aligned} \tag{1}$$

while μ is a Lagrange multiplier. The feasible directions of X at x^* are the vectors of the form $d = x - x^*$ where $x \in X$. Hence the assumption that there exists a feasible direction d with the property described is equivalent to the existence of an $\overline{x} \in X$ such that $h_j(\overline{x}) < 0$ for all j.

If there exists a feasible direction d with $\nabla g_j(x^*)'d < 0$ for all j, then by Prop. 4.3.12, the set M^* is nonempty. Applying the result of Exercise 6.3.1 to problem (1), we see that the set M^* is bounded. Conversely, if M^* is nonempty and bounded, again applying the result of Exercise 6.3.1, we see that there exists $\overline{x} \in X$ such that $h_j(\overline{x}) < 0$ for all j, and hence also there exists a feasible direction with the required property.

6.3.4 (Characterization of Pareto Optimality)

(a) Assume that x^* is not a Pareto optimal solution. Then there is a vector $\overline{x} \in X$ such that either

$$f_1(\bar{x}) \le f_1(x^*), \quad f_2(\bar{x}) < f_2(x^*),$$

or

$$f_1(\bar{x}) < f_1(x^*), \quad f_2(\bar{x}) \le f_2(x^*).$$

In either case, by using the facts $\lambda_1^* > 0$ and $\lambda_2^* > 0$, we have

$$\lambda_1^* f_1(\bar{x}) + \lambda_2^* f_2(\bar{x}) < \lambda_1^* f_1(x^*) + \lambda_2^* f_2(x^*),$$

yielding a contradiction. Therefore x^* is a Pareto optimal solution.

(b) Let

$$A = \{(z_1, z_2) \mid \text{ there exists } x \in X \text{ such that } f_1(x) \le z_1,\ f_2(x) \le z_2\}.$$

We first show that A is convex. Indeed, let (a_1, a_2), (b_1, b_2) be elements of A, and let $(c_1, c_2) = \alpha(a_1, a_2) + (1-\alpha)(b_1, b_2)$, for any $\alpha \in [0, 1]$. Then for some $x_a \in X$, $x_b \in X$, we have $f_1(x_a) \le a_1$, $f_2(x_a) \le a_2$, $f_1(x_b) \le b_1$, and $f_2(x_b) \le b_2$. Let $x_c = \alpha x_a + (1 - \alpha)x_b$. Since X is convex, $x_c \in X$, and since f_1 and f_2 are convex, we have

$$f_1(x_c) \le c_1, \qquad f_2(x_c) \le c_2.$$

Hence $(c_1, c_2) \in A$ implying that A is convex.

Note that $\big(f_1(x^*), f_2(x^*)\big)$ is not an interior point of A. [If this were not the case, then for some $\overline{x} \in X$ we would have $f_1(\overline{x}) < f_1(x^*)$ and $f_2(\overline{x}) < f_2(x^*)$, so that x^* would not be Pareto optimal.] By supporting hyperplane theorem, there exist λ_1^* and λ_2^*, not both equal to 0, such that

$$\lambda_1^* z_1 + \lambda_2^* z_2 \geq \lambda_1^* f_1(x^*) + \lambda_2^* f_2(x^*), \qquad \forall\ (z_1, z_2) \in A.$$

Since z_1 and z_2 can be arbitrarily large, we must have $\lambda_1^* \geq 0$ and $\lambda_2^* \geq 0$. Furthermore, by the definition of the set A, from the above equation we obtain

$$\lambda_1^* f_1(x) + \lambda_2^* f_2(x) \geq \lambda_1^* f_1(x^*) + \lambda_2^* f_2(x^*), \qquad \forall\ x \in X,$$

implying that

$$\min_{x \in X} \big\{\lambda_1^* f_1(x) + \lambda_2^* f_2(x)\big\} = \lambda_1^* f_1(x^*) + \lambda_2^* f_2(x^*).$$

(c) Generalization of (a): If x^* is a vector in X, and $\lambda_1^*, \ldots, \lambda_m^*$ are positive scalars such that

$$\sum_{i=1}^m \lambda_i^* f_i(x^*) = \min_{x \in X} \left\{ \sum_{i=1}^m \lambda_i^* f_i(x) \right\},$$

then x^* is a Pareto optimal solution.

Generalization of (b): Assume that X is convex and $f_1, \ldots, f_m$ are convex over X. If x^* is a Pareto optimal solution, then there exist non-negative scalars $\lambda_1^*, \ldots, \lambda_m^*$, not all zero, such that

$$\sum_{i=1}^m \lambda_i^* f_i(x^*) = \min_{x \in X} \left\{ \sum_{i=1}^m \lambda_i^* f_i(x) \right\}.$$

6.3.5 (Directional Convexity)

Let

$$A = \{(z, w) \mid \text{there exists } (x, u) \in \Re^{n+s} \text{ such that}$$
$$h(x, u) = z,\ \ f(x, u) \leq w,\ \ u \in U\}.$$

Suppose that $(0, f^*)$ is an interior point of A. Then the point $(0, f^* - \delta)$ belongs to A for some small enough $\delta > 0$. By definition of the set A, we have that $h(x, u) = 0$ and $f(x, u) \leq f^* - \delta$ for some $x \in \Re^n$ and $u \in U$, which contradicts the fact that f^* is the optimal value. Therefore $(0, f^*)$ must be on the boundary of the set A. Furthermore, there is a supporting hyperplane of the set A that passes through the point $(0, f^*)$. In other words, there exists a nonzero vector (λ, β) such that

$$\beta f^* \leq \lambda' z + \beta w, \qquad \forall (z, w) \in A. \tag{1}$$

By assumption (2), we have that for $z = 0$ there are a vector $u \in U$ and a vector $x \in \Re^n$ such that $h(x,u) = 0$, which implies that $(0,w) \in A$ for all w with $w \geq f(x,u) \geq f^*$. Then from (1) we have

$$0 \leq \beta(w - f^*), \qquad \forall\ w \geq f^*,$$

which holds only if $\beta \geq 0$. Suppose that $\beta = 0$. Then assumption (2) and Eq. (1) imply that

$$\lambda' z \geq 0, \qquad \forall\ z \text{ with } ||z|| < \epsilon,$$

which is possible only if $\lambda = 0$. But this contradicts the fact that $(\lambda, \beta) \neq 0$. Hence, we can take $\beta = 1$ in Eq. (1). From here and the definition of the set A, we obtain

$$f^* \leq f(x,u) + \lambda' h(x,u), \qquad \forall\ x \in \Re^n,\ u \in U.$$

This, combined with weak duality, implies that

$$\inf_{x \in \Re^n, u \in U} \big\{ f(x,u) + \lambda' h(x,u) \big\} = f^*. \tag{2}$$

Suppose that (x^*, u^*) is an optimal solution. Then (x^*, u^*) must be feasible [i.e., it must satisfy $h(x^*, u^*) = 0$], and

$$f^* = f(x^*, u^*) + \lambda' h(x^*, u^*) = \inf_{x \in \Re^n, u \in U} \left\{ f(x,u) + \lambda' h(x,u) \right\},$$

where the last equality follows from Eq. (2). Therefore we must have

$$u^* = \arg\min_{u \in U} \left\{ f(x^*, u) + \lambda' h(x^*, u) \right\}.$$

Similarly, we can argue that

$$x^* = \arg\min_{x \in \Re^n} \left\{ f(x, u^*) + \lambda' h(x, u^*) \right\}.$$

If f and h are continuously differentiable with respect to x for any $u \in U$, the last relation implies that

$$\nabla_x f(x^*, u^*) + \nabla_x h(x^*, u^*) \lambda = 0.$$

6.3.6 (Directional Convexity and Optimal Control)

Similar to Exercise 6.3.5, we can show that there is a vector $(-p^*, 1)$ such that

$$f^* \le -p^{*\prime} z + w, \qquad \forall\ (z, w) \in A,$$

where $p^* \in \Re^N$. This implies

$$\begin{aligned}
f^* &\le \inf_{x_i\in\Re, i=1,\ldots,N} \ \inf_{u_i\in U_i, i=0,\ldots,N-1} \left\{ \sum_{i=0}^{N-1} \left({p_{i+1}^*}' f_i(x_i, u_i) + g_i(x_i, u_i) - {p_{i+1}^*}' x_{i+1} \right) + g_N(x_N) \right\} \\
&= \inf_{x_i\in\Re,\ i=1,\ldots,N} \left\{ g_N(x_N) + \sum_{i=0}^{N-1} \left(\inf_{u_i\in U_i} \left\{ {p_{i+1}^*}' f_i(x_i, u_i) + g_i(x_i, u_i) \right\} - {p_{i+1}^*}' x_{i+1} \right) \right\} \\
&= q(p^*) = q^*.
\end{aligned}$$

From here and the weak duality theorem (Prop. 6.1.3), it follows that p^* is a Lagrange multiplier and that there is no duality gap. Using the same argument as in Exercise 6.3.5, we can show that

$$u_i^* = \arg\min_{u_i\in U_i} H_i(x^*, u_i, p_{i+1}^*), \qquad i = 0, \ldots, N-1,$$

where

$$H_i(x, u_i, p_{i+1}) = p_{i+1}' f_i(x_i, u_i) + g_i(x_i, u_i).$$

Also, we have

$$\begin{aligned}
x^* &= \mathrm{argmin}_{x_i\in\Re,\ i=1,\ldots,N} \left\{ g_N(x_N) + \sum_{i=0}^{N-1} \left({p_{i+1}^*}' f_i(x_i, u_i^*) + g_i(x_i, u_i^*) - {p_{i+1}^*}' x_{i+1} \right) \right\} \\
&= \mathrm{argmin}_{x_i\in\Re,\ i=1,\ldots,N} \left\{ \sum_{i=1}^{N-1} \left({p_{i+1}^*}' f_i(x_i, u_i^*) + g_i(x_i, u_i^*) - p_i^{*\prime} x_i \right) + g_N(x_N) - {p_N^*}' x_N \right\},
\end{aligned}$$

where $x^* = (x_1^*, \ldots, x_N^*)$. By using the separable structure of the expression on the left hand-side in the relation above, we obtain

$$x_i^* = \arg\min_{x_i\in\Re} \left\{ {p_{i+1}^*}' f_i(x_i, u_i^*) + g_i(x_i, u_i^*) - p_i^{*\prime} x_i \right\}, \quad \text{for } i = 1, \ldots, N-1,$$

and

$$x_N^* = \arg\min_{x_N\in\Re} \left\{ g_N(x_N) - {p_N^*}' x_N \right\}.$$

Since the functions f_i and g_i are continuously differentiable with respect to x_i for each $u_i \in U_i$, the last two relations are equivalent to

$$\nabla_{x_i} H_i(x^*, u_i^*, p_{i+1}^*) = p_i^*, \quad \text{for } i = 1, \ldots, N-1$$

and

$$\nabla g_N(x_N^*) = p_N^*,$$

respectively.

6.3.8 (Inconsistent Convex Systems of Inequalities)

The dual function for the problem in the hint is

$$q(\mu) = \inf_{y\in\Re,\, x\in X} \left\{ y + \sum_{j=1}^{r} \mu_j \big(g_j(x) - y\big) \right\} = \begin{cases} \inf_{x\in X} \sum_{j=1}^{r} \mu_j g_j(x) & \text{if } \sum_{j=1}^{r} \mu_j = 1 \\ -\infty & \text{if } \sum_{j=1}^{r} \mu_j \neq 1 \end{cases}$$

The problem in the hint satisfies the interior point Assumption 6.3.1, so by Prop. 6.3.1 the dual problem has an optimal solution μ^* and there is no duality gap.

Clearly the problem in the hint has an optimal value that is greater or equal to 0 if and only if the system of inequalities

$$g_j(x) < 0, \qquad j = 1, \ldots, r,$$

has no solution within X. Since there is no duality gap, we have

$$\max_{\mu\geq 0,\, \sum_{j=1}^{r} \mu_j = 1} q(\mu) \geq 0$$

if and only if the system of inequalities $g_j(x) < 0$, $j = 1, \ldots, r$, has no solution within X. This is equivalent to the statement we want to prove.

6.3.9 (Duality Gap Example)

It can be seen that a vector (x_1, x_2) is feasible if and only if

$$x_1 \geq 0, \qquad x_2 = 0.$$

Furthermore, all feasible points attain the optimal value, which is $f^* = 1$.

Consider now the dual function

$$q(\mu) = \inf_{x\in\Re^2} \big\{ e^{x_2} + \mu\big(\|x\| - x_1\big) \big\}. \tag{1}$$

We will show that $q(\mu) = 0$ for all $\mu \geq 0$ by deriving the set of constraint-cost pairs

$$\big\{ (\|x\| - x_1, e^{x_2}) \mid x \in \Re^2 \big\}.$$

Indeed, for $u < 0$ there is no x such that $\|x\| - x_1 = u$. For $u = 0$, the vectors x such that $\|x\| - x_1 = u$ are of the form $x = (x_1, 0)$, so the set of constraint-cost pairs with constraint value equal to 0 is $(0, 1)$. For $u > 0$, for each x_2, the equation $\|x\| - x_1 = u$ has a solution in x_1:

$$x_1 = \frac{x_2^2 - u^2}{2u}.$$

Thus, if $u > 0$, the set of constraint-cost pairs with constraint value equal to u is

$$\big\{(u, w) \mid w > 0\big\}.$$

Combining the preceding facts, we see that the set of constraint-cost pairs is

$$\big\{(0, 1)\big\} \cup \big\{(u, w) \mid u > 0,\, w > 0\big\},$$

which based on the geometric constructions of Section 6.1, shows that $q(\mu) = 0$ for all $\mu \geq 0$ [this can also be verified using the definition (1) of q]. Thus,

$$q^* = \sup_{\mu \geq 0} q(\mu) = 0,$$

and there is a duality gap, $f^* - q^* = 1$.

The difficulty here is that g is nonlinear and there is no $\overline{x} \in X$ such that $g(\overline{x}) < 0$, so the Slater condition (Assumption 6.3.1) is violated.

SECTION 6.4

6.4.3

(a) Let us apply Fenchel's duality theorem to the functions

$$f_1(x) = \frac{c}{2}\|x\|^2, \qquad f_2(x) = -\gamma(t - x),$$

and the sets

$$X_1 = X_2 = \Re^n.$$

Using Prop. 6.4.1, we have

$$P_c(t) = \min_{x \in \Re^n} \left\{ \frac{c}{2}\|x\|^2 + \gamma(t - x) \right\} = \max_{\lambda \in \Re^n} \big\{ g_2(\lambda) - g_1(\lambda) \big\},$$

where the corresponding conjugates g_1 and g_2 are calculated as follows:

$$g_1(\lambda) = \sup_{x \in \Re^n} \left\{ x'\lambda - \frac{c}{2}\|x\|^2 \right\} = \frac{1}{2c}\|\lambda\|^2,$$

$$\begin{aligned} g_2(\lambda) &= \inf_{x\in\Re^n} \big\{x'\lambda + \gamma(t-x)\big\} \\ &= t'\lambda + \inf_{x\in\Re^n} \big\{\gamma(t-x) - (t-x)'\lambda\big\} \\ &= t'\lambda - \sup_{u\in\Re^n} \big\{u'\lambda - \gamma(u)\big\} \\ &= t'\lambda - g(\lambda). \end{aligned}$$

Thus we have

$$P_c(t) = \max_{\lambda\in\Re^n} \left\{ t'\lambda - g(\lambda) - \frac{1}{2c}\|\lambda\|^2 \right\}.$$

The function $P_c(\cdot)$ is the pointwise maximum of a collection of linear functions, so it is convex. To show that $P_c(t)$ is differentiable, we view it as the primal function of a suitable problem, and we use the fact that the subgradients of the primal function at 0 are the negatives of the corresponding Lagrange multipliers (cf. Section 6.4.4). Consider the convex programming problem

$$\begin{aligned} &\text{minimize} \;\; \frac{c}{2}\|y\|^2 + \gamma(z) \\ &\text{subject to} \;\; t - y - z = 0, \end{aligned} \tag{1}$$

whose primal function is

$$p(w) = \min_{t-y-z=w} \left\{ \frac{c}{2}\|y\|^2 + \gamma(z) \right\} = \min_{x\in\Re^n} \left\{ \frac{c}{2}\|x\|^2 + \gamma(t-x-w) \right\}.$$

We have

$$p(w) = P_c(t-w)$$

and the set of subgradients of P_c at t is the set of the negatives of the subgradients of $p(w)$ at 0, or equivalently (by the theory of Section 6.4.4), the set of dual optimal solutions of problem (1).

The dual function of problem (1) is

$$\begin{aligned} q(\lambda) &= \min_{z,y} \left\{ \frac{c}{2}\|y\|^2 + \gamma(z) + \lambda'(t-y-z) \right\} \\ &= \min_{y} \left\{ \frac{c}{2}\|y\|^2 - \lambda'y \right\} + \min_{z} \big\{\gamma(z) + \lambda'(t-z)\big\} \\ &= -\frac{1}{2c}\|\lambda\|^2 - \max_{z} \big\{(z-t)'\lambda - \gamma(z)\big\} \qquad\qquad . \\ &= -\frac{1}{2c}\|\lambda\|^2 - \max_{z} \big\{z'\lambda - \gamma(z)\big\} + t'\lambda \\ &= t'\lambda - g(\lambda) - \frac{1}{2c}\|\lambda\|^2 \end{aligned}$$

Thus the optimal dual solution is the unique λ attaining the maximum of $q(\lambda)$. As argued earlier, this λ must be equal to $\nabla P_c(t)$.

(b) The formulas and their derivation can be found in [Ber77] and [Ber82a], Section 3.3.

(c) We have

$$P_c(t) = \inf_{u\in\Re^n} \left\{ \gamma(t-u) + \frac{c}{2}\|u\|^2 \right\} \le \gamma(t).$$

Also, if d is a subgradient of γ at t, we have for all $u \in \Re^n$

$$\gamma(t-u) + \frac{c}{2}\|u\|^2 \geq \gamma(t) - d'u + \frac{c}{2}\|u\|^2 \geq \gamma(t) + \min_u \left\{-d'u + \frac{c}{2}\|u\|^2\right\} = \gamma(t) - \frac{1}{2c}\|d\|^2.$$

Thus we have for all t

$$\gamma(t) - \frac{1}{2c}\|d\|^2 \leq P_c(t) \leq \gamma(t),$$

which implies that $\lim_{c\to\infty} P_c(t) = \gamma(t)$.

(d) From the Fenchel duality theorem we obtain, similar to part (a),

$$P_c(t) = \sup_{\lambda\in\Re^n} \left\{t'\lambda - g(\lambda) - \frac{1}{2c}\|\lambda - y\|^2\right\}.$$

6.4.5

Define $X = \{x \mid ||x|| \leq 1\}$, and note that X is convex and compact set. Therefore, according to Minimax Theorem, we have

$$\min_{x\in X}\max_{y\in Y} x'y = \max_{y\in Y}\min_{x\in X} x'y.$$

For a fixed $y \in Y$, the minimum of $x'y$ over X is attained at $x^* = -y/||y||$ if $y \neq 0$ and $x^* = 0$ if $y = 0$ [this can be verified by the first order necessary condition, which is here also sufficient by convexity of $x'y$]. Thus we obtain

$$\min_{x\in X}\max_{y\in Y} x'y = \max_{y\in Y}\big(-||y||\big) = -\min_{y\in Y}||y||.$$

Thus the original problem can be solved by projecting the origin on the set Y.

6.4.6 (Quadratically Constrained Quadratic Problems [LVB98])

Since each P_i is symmetric and positive definite, we have

$$\begin{aligned} x'P_ix + 2q_i'x + r_i &= \left(P_i^{1/2}x\right)' P_i^{1/2}x + 2\left(P_i^{-1/2}q_i\right)' P_i^{1/2}x + r_i \\ &= ||P_i^{1/2}x + P_i^{-1/2}q_i||^2 + r_i - q_i'P_i^{-1}q_i, \end{aligned}$$

for $i = 0, 1, \ldots, p$. This allows us to write the original problem as

$$\begin{aligned} &\text{minimize} \quad ||P_0^{1/2}x + P_0^{-1/2}q_0||^2 + r_0 - q_0'P_0^{-1}q_0 \\ &\text{subject to} \quad ||P_i^{1/2}x + P_i^{-1/2}q_i||^2 + r_i - q_i'P_i^{-1}q_i \leq 0, \;\; i = 1, \ldots, p. \end{aligned}$$

By introducing a new variable x_{n+1}, this problem can be formulated in $\Re^{n+1}$ as

$$\begin{aligned} \text{minimize} \quad & x_{n+1} \\ \text{subject to} \quad & \|P_0^{1/2}x + P_0^{-1/2}q_0\| \leq x_{n+1} \\ & \|P_i^{1/2}x + P_i^{-1/2}q_i\| \leq \left(q_i'P_i^{-1}q_i - r_i\right)^{1/2}, \quad i = 1, \ldots, p. \end{aligned}$$

The optimal values of this problem and the original problem are equal up to a constant and a square root. The above problem is of the type described in Section 6.4.1. To see that define $A_i = \left(P_i^{1/2} \mid 0\right)$, $b_i = P_i^{-1/2}q_i$, $e_i = 0$, $d_i = \left(q_i'P_i^{-1}q_i - r_i\right)^{1/2}$ for $i = 1, \ldots, p$, $A_0 = \left(P_0^{1/2} \mid 0\right)$, $b_0 = P_0^{-1/2}q_0$, $e_0 = (0, \ldots, 0, 1)$, $d_0 = 0$, and $c = (0, \ldots, 0, 1)$. Its dual is given by

$$\begin{aligned} \text{maximize} \quad & -\sum_{i=1}^{p} \left(q_i'P_i^{-1/2}z_i + \left(q_i'P_i^{-1}q_i - r_i\right)^{1/2} w_i\right) - q_0'P_0^{-1/2}z_0 \\ \text{subject to} \quad & \sum_{i=0}^{p} P_i^{1/2}z_i = 0, \quad \|z_0\| \leq 1, \quad \|z_i\| \leq w_i, \quad i = 1, \ldots, p. \end{aligned}$$

6.4.7 (Minimizing the Sum or the Maximum of Norms [LVB98])

Consider the problem

$$\begin{aligned} \text{minimize} \quad & \sum_{i=1}^{p} \|F_ix + g_i\| \\ \text{subject to} \quad & x \in \Re^n. \end{aligned}$$

By introducing variables $t_1, \ldots, t_p$, this problem can be expressed as a second-order cone programming problem (see Exercise 6.4.17):

$$\begin{aligned} \text{minimize} \quad & \sum_{i=1}^{p} t_i \\ \text{subject to} \quad & \|F_ix + g_i\| \leq t_i, \quad i = 1, \ldots, p. \end{aligned}$$

Define

$$X = \{(x, u, t) \mid x \in \Re^n,\ u_i = F_ix + g_i,\ t_i \in \Re,\ i = 1, \ldots, p\},$$

$$C = \{(x, u, t) \mid x \in \Re^n,\ \|u_i\| \leq t_i, \quad i = 1, \ldots, p\}.$$

Then, by applying the result of Exercise 6.4.3 with $f(x, u, t) = \sum_{i=1}^{p} t_i$, and X, C defined above, we have

$$-C^\perp = \{(0, z, w) \mid \|z_i\| \leq w_i, \quad i = 1, \ldots, p\},$$

and

$$\begin{aligned}
g(0,z,w) &= \sup_{(x,u,t)\in X} \left\{ \sum_{i=1}^p z_i' u_i + \sum_{i=1}^p w_i t_i - \sum_{i=1}^p t_i \right\} \\
&= \sup_{x\in\Re^n,\, t\in\Re^p} \left\{ \sum_{i=1}^p z_i'(F_i x + g_i) + \sum_{i=1}^p (w_i - 1) t_i \right\} \\
&= \sup_{x\in\Re^n} \left\{ \left(\sum_{i=1}^p F_i' z_i \right)' x \right\} + \sup_{t\in\Re^p} \left\{ \sum_{i=1}^p (w_i - 1) t_i \right\} + \sum_{i=1}^p g_i' z_i \\
&= \begin{cases} \sum_{i=1}^p g_i' z_i & \text{if } \sum_{i=1}^p F_i' z_i = 0,\ w_i = 1,\ i = 1,\ldots,p \\ +\infty & \text{otherwise.} \end{cases}
\end{aligned}$$

Hence the dual problem is given by

$$\begin{aligned}
&\text{maximize } -\sum_{i=1}^p g_i' z_i \\
&\text{subject to } \sum_{i=1}^p F_i' z_i = 0, \quad \|z_i\| \le 1,\ i = 1,\ldots,p.
\end{aligned}$$

Now, consider the problem

$$\begin{aligned}
&\text{minimize } \max_{1\le i\le p} \|F_i x + g_i\| \\
&\text{subject to } x \in \Re^n.
\end{aligned}$$

By introducing a new variable x_{n+1}, we obtain

$$\begin{aligned}
&\text{minimize } x_{n+1} \\
&\text{subject to } \|F_i x + g_i\| \le x_{n+1},\ i = 1,\ldots,p,
\end{aligned}$$

or equivalently

$$\begin{aligned}
&\text{minimize } e_{n+1}' x \\
&\text{subject to } \|A_i x + g_i\| \le e_{n+1}' x,\ i = 1,\ldots,p,
\end{aligned}$$

where $x \in \Re^{n+1}$, $A_i = (F_i | 0)$, and $e_{n+1} = (0,\ldots,0,1)' \in \Re^{n+1}$. Evidently, this is a second-order cone programming problem. Its dual problem is given by

$$\begin{aligned}
&\text{maximize } -\sum_{i=1}^p g_i' z_i \\
&\text{subject to } \sum_{i=1}^p \left(\begin{pmatrix} F_i' \\ 0 \end{pmatrix} z_i + e_{n+1} w_i \right) = e_{n+1}, \quad \|z_i\| \le w_i, \quad i = 1,\ldots,p,
\end{aligned}$$

or equivalently

$$\begin{aligned}
&\text{maximize } -\sum_{i=1}^p g_i' z_i \\
&\text{subject to } \sum_{i=1}^p F_i' z_i = 0, \quad \sum_{i=1}^p w_i = 1, \quad \|z_i\| \le w_i, \quad i = 1,\ldots,p.
\end{aligned}$$

6.4.8

Let $f(x) = (1/2)x'Qx$. Since the problem has a unique optimal solution, we have that Q is positive definite on the nullspace of the matrix A, implying that $f(x) + (1/2c)\|x - x^k\|^2$ also has a unique minimum subject to $Ax = b$. Hence the algorithm

$$x^{k+1} = \arg\min_{Ax=b} \left\{ \frac{1}{2}x'Qx + \frac{1}{2c}\|x - x^k\|^2 \right\}$$

is well defined. This algorithm can also be written as

$$x^{k+1} = \arg\min_{Ax=b} \left\{ \frac{1}{2}x'Qx + \frac{1}{2c}\|x - x^k\|^2 + \frac{\gamma}{2}\|Ax - b\|^2 \right\}$$

for any scalar γ. If γ is sufficiently large, the quadratic function

$$\frac{1}{2}x'Qx + \frac{\gamma}{2}\|Ax - b\|^2$$

is positive definite by Lemma 3.2.1. For such γ, the above algorithm is equivalent to the proximal minimization algorithm and it inherits the corresponding convergence properties.

6.4.8 (Complex l_1 and l_∞ Approximation [LVB98])

For $v \in \mathcal{C}^p$ we have

$$||v||_1 = \sum_{i=1}^{p} |v_i| = \sum_{i=1}^{p} \left|\left| \begin{pmatrix} \mathcal{R}e(v_i) \\ \mathcal{I}m(v_i) \end{pmatrix} \right|\right|,$$

where $\mathcal{R}e(v_i)$ and $\mathcal{I}m(v_i)$ denote real and imaginary parts of v_i, respectively. Then the complex l_1 approximation problem is equivalent to

$$\begin{aligned} &\text{minimize } \sum_{i=1}^{p} \left|\left| \begin{pmatrix} \mathcal{R}e(a_i'x - b_i) \\ \mathcal{I}m(a_i'x - b_i) \end{pmatrix} \right|\right| \\ &\text{subject to } x \in \mathcal{C}^n, \end{aligned} \tag{1}$$

where a_i' is the i-th row of A (A is a $p \times n$ matrix). Note that

$$\begin{pmatrix} \mathcal{R}e(a_i'x - b_i) \\ \mathcal{I}m(a_i'x - b_i) \end{pmatrix} = \begin{pmatrix} \mathcal{R}e(a_i') & -\mathcal{I}m(a_i') \\ \mathcal{I}m(a_i') & \mathcal{R}e(a_i') \end{pmatrix} \begin{pmatrix} \mathcal{R}e(x) \\ \mathcal{I}m(x) \end{pmatrix} - \begin{pmatrix} \mathcal{R}e(b_i) \\ \mathcal{I}m(b_i). \end{pmatrix}$$

By introducing new variables $y = (\mathcal{R}e(x'), \mathcal{I}m(x'))'$, problem (1) can be rewritten as

$$\begin{aligned} &\text{minimize } \sum_{i=1}^{p} ||F_i y + g_i|| \\ &\text{subject to } y \in \Re^{2n}, \end{aligned}$$

where

$$F_i = \begin{pmatrix} \mathcal{R}e(a_i') & -\mathcal{I}m(a_i') \\ \mathcal{I}m(a_i') & \mathcal{R}e(a_i') \end{pmatrix}, \qquad g_i = -\begin{pmatrix} \mathcal{R}e(b_i) \\ \mathcal{I}m(b_i) \end{pmatrix}. \tag{2}$$

According to Exercise 6.4.19, the dual problem is given by

$$\begin{aligned} &\text{maximize } \sum_{i=1}^{p} \big(\, \mathcal{R}e(b_i), \mathcal{I}m(b_i) \,\big)\, z_i \\ &\text{subject to } \sum_{i=1}^{p} \begin{pmatrix} \mathcal{R}e(a_i') & \mathcal{I}m(a_i') \\ -\mathcal{I}m(a_i') & \mathcal{R}e(a_i') \end{pmatrix} z_i = 0, \qquad \|z_i\| \leq 1, \ i = 1, \ldots, p, \end{aligned}$$

where $z_i \in \Re^{2n}$ for all i.

For $v \in \mathcal{C}^p$ we have

$$\|v\|_\infty = \max_{1 \leq i \leq p} |v_i| = \max_{1 \leq i \leq p} \left\| \begin{pmatrix} \mathcal{R}e(v_i) \\ \mathcal{I}m(v_i) \end{pmatrix} \right\|.$$

Therefore the complex l_∞ approximation problem is equivalent to

$$\begin{aligned} &\text{minimize } \max_{1 \leq i \leq p} \left\| \begin{pmatrix} \mathcal{R}e(a_i'x - b_i) \\ \mathcal{I}m(a_i'x - b_i) \end{pmatrix} \right\| \\ &\text{subject to } x \in \mathcal{C}^n, \end{aligned}$$

By introducing new variables $y = (\mathcal{R}e(x'), \mathcal{I}m(x'))'$, this problem can be rewritten as

$$\begin{aligned} &\text{minimize } \max_{1 \leq i \leq p} \|F_i y + g_i\| \\ &\text{subject to } y \in \Re^{2n}, \end{aligned}$$

where F_i and g_i are given by Eq. (2). From Exercise 6.4.19, it follows that the dual problem is

$$\begin{aligned} &\text{maximize } \sum_{i=1}^{p} \big(\, \mathcal{R}e(b_i), \mathcal{I}m(b_i) \,\big)\, z_i \\ &\text{subject to } \sum_{i=1}^{p} \begin{pmatrix} \mathcal{R}e(a_i') & -\mathcal{I}m(a_i') \\ \mathcal{I}m(a_i') & \mathcal{R}e(a_i') \end{pmatrix} z_i = 0, \quad \sum_{i=1}^{p} w_i = 1, \quad \|z_i\| \leq w_i, \quad i = 1, \ldots, p, \end{aligned}$$

where $z_i \in \Re^2$ for all i.

6.4.11 (Strong Duality for One-Dimensional Problems)

For a function $h : \Re \mapsto [-\infty, \infty]$, the domain of h is the set

$$\text{dom}(h) = \big\{ x \mid -\infty < h(x) < \infty \big\}.$$

If h is lower semicontinuous over its domain, i.e., satisfies $h(x) \leq \liminf_{k \to \infty} h(x_k)$ for all $x \in \text{dom}(h)$ and all sequences $\{x_k\}$ with $x_k \to x$, it is called *domain lower semicontinuous* or DLSC for short. Note that a convex DLSC function need not be closed, i.e., need not have a closed epigraph.

Convex DLSC functions arise in the context of the constrained optimization problem

$$\begin{aligned} &\text{minimize} \quad f(x) \\ &\text{subject to} \quad g_j(x) \le 0, \quad j = 1, \ldots, r, \end{aligned} \tag{1}$$

where $f : \Re^n \mapsto (-\infty, \infty]$ and $g_j : \Re^n \mapsto (-\infty, \infty]$ are some proper extended real-valued functions. We denote by g the vector-valued function $g = (g_1, \ldots, g_r)$, and we denote compactly inequalities of the form $g_j(x) \le 0$, $j = 1, \ldots, r$, as $g(x) \le 0$.

The primal function of the problem, defined by

$$p(u) = \inf_{g(x) \le u} f(x),$$

determines whether there is a duality gap. In particular, assuming that p is convex and that $p(0) < \infty$ (i.e., that the problem is feasible), there is no duality gap if and only if p is lower semicontinuous at $u = 0$. More generally, assuming that p is convex, there is no duality gap for every feasible problem of the form

$$\begin{aligned} &\text{minimize} \quad f(x) \\ &\text{subject to} \quad g(x) \le u, \end{aligned}$$

if and only if p is a DLSC function.

The two most common approaches to ascertain that there is no duality gap in problem (1) are:

(a) To show that p is closed, so that it is also lower semicontinuous at any u, including $u = 0$.

(b) To show that p is subdifferentiable at $u = 0$, so that it is also lower semicontinuous at $u = 0$. This is guaranteed, in particular, under the assumption $0 \in \mathrm{ri}\big(\mathrm{dom}(p)\big)$ or under some other constraint qualification that guarantees the existence of a geometric multiplier for problem (1).

Note, however, that there are some important special cases that are not covered by one of the above two approaches. In these cases, p is a DLSC function but it is not necessarily closed or subdifferentiable at 0. As an example, consider the one-dimensional problem where

$$f(x) = \begin{cases} \frac{1}{x} & \text{if } x > 0, \\ \infty & \text{if } x \le 0, \end{cases}$$

and

$$g(x) = e^{-x}.$$

Then it can be seen that

$$p(u) = \begin{cases} 0 & \text{if } u > 0, \\ \infty & \text{if } u \le 0, \end{cases}$$

so p is a DLSC function but is not closed.

In the special case where the x is a scalar and the functions f and g_j are convex, proper, and DLSC, we can show that the function p is DLSC. This is consistent with the preceding example.

Proposition: If the functions f and g_j map $\Re$ into $(-\infty, \infty]$ and are convex, proper, and DLSC, then the primal function p is DLSC.

Proof: Without loss of generality, we assume that $0 \in \text{dom}(p)$. It will be sufficient to show that p is lower semicontinuous at $u = 0$. Let $q^* = \lim_{u \to 0^+} p(u)$. We will show that $p(0) = q^*$, which implies lower semicontinuity of p at 0, since p is monotonically nonincreasing. Let $\{x^k\}$ be a scalar sequence such that $f(x^k) \to q^*$ and $\max\big\{0, g_j(x^k)\big\} \to 0$ for all j. We consider three cases:

(a) $\{x^k\}$ has a subsequence that converges to a scalar $\overline{x}$. Without loss of generality, we assume that the entire sequence $\{x^k\}$ converges to $\overline{x}$. By the lower semicontinuity of f and g_j, we have $f(\overline{x}) \le q^*$ and $g_j(\overline{x}) \le 0$ for all j. Hence $\overline{x}$ is feasible for the problem corresponding to $u = 0$, and we have $p(0) \le f(\overline{x}) \le q^*$. Since p is monotonically noninceasing, we also have $q^* \le p(0)$ and we obtain $p(0) = q^*$.

(b) $\{x^k\}$ has a subsequence that tends to ∞. Without loss of generality, we assume that $x^k \to \infty$. Then the positive direction is a direction of recession of f and g_j for all j. This implies that $\inf_{x \in \Re} f(x) = q^*$, and also that $g(x^k) \le 0$ for all sufficiently large k [otherwise the problem corresponding to $u = 0$ would be infeasible, thereby violating the hypothesis that $0 \in \text{dom}(p)$]. Thus $p(0) = \inf_{x \in \Re} f(x) = q^*$.

(c) $\{x^k\}$ has a subsequence that tends to $-\infty$. Without loss of generality, we assume that $x^k \to -\infty$, and we proceed similar to case (b) above. **Q.E.D.**

The proposition implies that is no duality gap for the given problem, assuming that $-\infty < f^* < \infty$.

SECTION 6.5

6.5.3 (Separable Problems with Integer/Simplex Constraints)

Define $d_j(s) = f_j(s) - f_j(s-1)$ for $s = 1, \ldots m_j$ and $j = 1, \ldots, n$. By the convexity of f_j, we have

$$2f_j(s) \leq f_j(s-1) + f_j(s+1),$$

or equivalently

$$f_j(s) - f_j(s-1) \leq f_j(s+1) - f_j(s).$$

Therefore

$$d_j(1) \leq d_j(2) \leq \cdots \leq d_j(m_j) \quad j = 1, \ldots, n.$$

Consider the set $D = \{d_j(s) \mid s = 1, \ldots m_j,\ j = 1, \ldots, n\}$. At each iteration the algorithm chooses the smallest element in the set D as long as that smallest element is negative and the constraint is not violated. Let x^* be a solution generated by the algorithm and D^* be the set of elements of D that the algorithm chooses. Define

$$\lambda = \begin{cases} \max_{d_j(s) \in D^*} d_j(s) & \text{if } |D^*| = A \\ 0 & \text{if } |D^*| < A, \end{cases}$$

where $|\cdot|$ denotes the cardinality of a set. If $|D^*| = A$, then the algorithm chooses A smallest elements in D, which are all negative, so that $\lambda < 0$. If $|D^*| < A$, then either x^* has components $x_j^* = m_j$ or the set D has less then A negative elements. Consider the following function

$$\sum_{j=1}^n f_j(x_j) - \lambda \sum_{j=1}^n x_j. \tag{1}$$

We have

$$\sum_{j=1}^n (f_j(x_j) - \lambda x_j) = \sum_{j=1}^n \left(f_j(0) + (d_j(1) - \lambda) + \cdots + (d_j(x_j) - \lambda)\right).$$

By the definition of λ and D^*, we have

$$d_j(s) - \lambda \leq 0 \text{ if } d_j(s) \in D^*,$$

$$d_j(s) - \lambda \geq 0 \text{ if } d_j(s) \notin D^*.$$

Therefore the function given by Eq. (1) is minimized at $x = x^*$. Consequently, $-\lambda \geq 0$ is a Lagrange multiplier for the original problem and there is no duality gap. By Prop. 6.1.5, we have that x^* is an optimal solution.

6.5.4 (Monotone Discrete Problems [WuB01])

A detailed analysis appears in the paper

C. C. Wu and D. P. Bertsekas, "Distributed Power Control Algorithms for Wireless Networks," IEEE Trans. on Vehicular Technology, Vol. 50, pp. 504-514, 2001

which is available from the author's www site

http://web.mit.edu/dimitrib/www/home.html

6.5.5

Suppose that E is totally unimodular. Let J be a subset of $\{1,\ldots,n\}$. Define z by $z_j = 1$ if $j \in J$, and $z_j = 0$ otherwise. Also let $w = Ez$, $d_i = f_i = \frac{1}{2}w_i$ if w_i is even, and $d_i = \frac{1}{2}(w_i + 1)$, $f_i = \frac{1}{2}(w_i - 1)$ if w_i is odd. Since E is totally unimodular, the polyhedron

$$P = \{x \mid f \le Ex \le d,\ 0 \le x \le z\}$$

has integral extreme points and $z \notin P$. Note that $P \neq \emptyset$ because $\frac{1}{2}z \in P$. Therefore there is a vector $\hat{x} \in P$ such that $\hat{x}_j = 0$ for $j \notin J$, and $\hat{x}_j \in \{0,1\}$ for $j \in J$. We have $z_j - 2\hat{x}_j = \pm 1$ for $j \in J$. Define $J_1 = \{j \in J \mid z_j - 2\hat{x}_j = 1\}$ and $J_2 = \{j \in J \mid z_j - 2\hat{x}_j = -1\}$. We have

$$\sum_{j\in J_1} e_{ij} - \sum_{j\in J_2} e_{ij} = \sum_{j\in J_1} e_{ij}(z_j - 2\hat{x}_j) = Ez - 2E\hat{x} = \begin{cases} w_i - w_i = 0 & \text{if } w_i \text{ is even} \\ w_i - (w_i \pm 1) = \mp 1 & \text{if } w_i \text{ is odd.} \end{cases}$$

Thus

$$\left| \sum_{j\in J_1} e_{ij} - \sum_{j\in J_2} e_{ij} \right| \le 1, \quad \forall\ i = 1,\ldots,m. \tag{1}$$

Suppose that the matrix E is such that any $J \subset \{1,\ldots,n\}$ can be partitioned into two subsets such that Eq. (1) holds. For $J \subset \{1,\ldots,n\}$ with J consisting of a single element, we obtain from in Eq. (1) $e_{ij} \in \{-1,0,1\}$ for all i and j. The proof is by induction on the size of the nonsingular submatrices of E using the hypothesis that the determinant of every $(k-1)\times(k-1)$ submatrix of E equals $-1, 0$, or 1. Let B be a $k \times k$ nonsingular submatrix of E. Our objective is to prove that $|\det B| = 1$. By the induction hypothesis and Cramer's rule, we have $B^{-1} = \frac{B^*}{\det B}$, where $b^*_{ij} \in \{-1,0,1\}$. By the definition of B^*, we have $Bb^*_1 = (\det B)e_1$, where b^*_1 is the first column of B^* and $e_1 = (1,0,\ldots 0)'$.

Let $J = \{i \mid b^*_{i1} \neq 0\}$ and $J'_1 = \{i \in J \mid b^*_{i1} = 1\}$. Hence for $i = 2,\ldots,k$, we have

$$(Bb^*_1)_i = \sum_{j\in J'_1} b_{ij} - \sum_{j\in J\setminus J'_1} b_{ij} = 0.$$

Thus the cardinality of the set $\{i \in J \mid b^*ij \neq 0\}$ is even, so for any partition $(\tilde{J}_1, \tilde{J}_2)$ of J, it follows that $\sum_{j\in\tilde{J}_1} b_{ij} - \sum_{j\in\tilde{J}_2} b_{ij}$ is even for all $i = 2, \ldots, k$. Now by assumption, there is a partition (J_1, J_2) of J such that $\left|\sum_{j\in J_1} b_{ij} - \sum_{j\in J_2} b_{ij}\right| \leq 1$. Hence

$$\sum_{j\in J_1} b_{ij} - \sum_{j\in J_2} b_{ij} = 0, \quad \text{for } i = 2, \ldots, k.$$

Now consider the value $\alpha_1 = \left|\sum_{j\in J_1} b_{1j} - \sum_{j\in J_2} b_{1j}\right|$. If $\alpha_1 = 0$, define $y \in \Re^k$ by $y_i = 1$ for $i \in J_1$, $y_i = -1$ for $i \in J_2$, and $y_i = 0$ otherwise. Since $By = 0$ and B is nonsingular, we have $y = 0$, which contradicts $J \neq \emptyset$. Hence by hypothesis, we have $\alpha_1 = 1$ and $By = \pm e_1$. However, $Bb_1^* = (\det B)e_1$. Since y and b_1^* are $(0, \pm 1)$ vectors, it follows that $b_1^* = \pm y$ and $|\det B| = 1$. Therefore E is totally unimodular.

6.5.6

Note that E is totally unimodular if and only if its transpose E' is totally unimodular. Hence according to Exercise 6.5.5, an $m \times n$ matrix E is totally unimodular if and only if every $I \subset \{1, \ldots, m\}$ can be partitioned into two subsets I_1 and I_2 such that

$$\left|\sum_{i\in I_1} e_{ij} - \sum_{i\in I_2} e_{ij}\right| \leq 1, \quad \forall\, j = 1, \ldots, n.$$

Let E be an $m \times n$ matrix with entries $e_{ij} \in \{-1, 0, 1\}$, and such that each of its columns contains at most two nonzero entries. By assumption, the set $\{1, \ldots, m\}$ can be partitioned into two subsets M_1 and M_2 so that if a column has two nonzero entries, the following hold:

(1) If both nonzero entries have the same sign, then one is in a row contained in M_1 and the other is in a row contained in M_2.

(2) If the two nonzero entries have opposite sign, then both are in rows contained in the same subset.

It follows that

$$\left|\sum_{i\in M_1} e_{ij} - \sum_{i\in M_2} e_{ij}\right| \leq 1, \quad \forall\, j = 1, \ldots, n. \tag{1}$$

Let I be any subset of $\{1, \ldots, m\}$. Then $I_1 = I \cap M_1$ and $I_2 = I \cap M_2$ constitute a partition of I, which in view of Eq. (1) satisfies

$$\left|\sum_{i\in I_1} e_{ij} - \sum_{i\in I_2} e_{ij}\right| \leq 1, \quad \forall\, j = 1, \ldots, n.$$

Hence E is totally unimodular.

6.5.7

Since E is totally unimodular if and only if its transpose E' is totally unimodular, then according to Exercise 6.5.5, E is totally unimodular if and only if every $I \subset \{1, \ldots, m\}$ can be partitioned into two subsets I_1 and I_2 such that

$$\left| \sum_{i \in I_1} e_{ij} - \sum_{i \in I_2} e_{ij} \right| \leq 1, \quad \forall\, j = 1, \ldots, n.$$

Define $M_1 = \{i \mid i \text{ is odd}\}$ and $M_2 = \{i \mid i \text{ is even}\}$. Then

$$\left| \sum_{i \in M_1} e_{ij} - \sum_{i \in M_2} e_{ij} \right| \leq 1, \quad \forall\, j = 1, \ldots, n.$$

Let I be any subset of $\{1, \ldots, m\}$. Then $I_1 = I \cap M_1$ and $I_2 = I \cap M_2$ constitute a partition of I, which satisfies

$$\left| \sum_{i \in I_1} e_{ij} - \sum_{i \in I_2} e_{ij} \right| \leq 1, \quad \forall\, j = 1, \ldots, n,$$

and therefore E is totally unimodular.

6.5.10

(a) We have for every $\mu = (\mu_1, \ldots, \mu_r) \geq 0$

$$\begin{aligned}
\overline{q}(\mu_1, \ldots, \mu_r) &= \inf_{x \in \overline{X}} \left\{ f(x) + \sum_{j=1}^{\overline{r}} \mu_j g_j(x) \right\} \\
&= \inf_{\substack{x \in X \\ g_{\overline{r}+1}(x) \leq 0, \ldots, g_r(x) \leq 0}} \left\{ f(x) + \sum_{j=1}^{\overline{r}} \mu_j g_j(x) \right\} \\
&\geq \inf_{\substack{x \in X \\ g_{\overline{r}+1}(x) \leq 0, \ldots, g_r(x) \leq 0}} \left\{ f(x) + \sum_{j=1}^{r} \mu_j g_j(x) \right\} \\
&\geq \inf_{x \in X} \left\{ f(x) + \sum_{j=1}^{r} \mu_j g_j(x) \right\} \\
&= q(\mu_1, \ldots, \mu_r).
\end{aligned}$$

By taking the supremum of both sides over $\mu \geq 0$, we obtain $\overline{q}^* \geq q^*$. The inequality $\overline{q}^* \leq f^*$ holds by the Weak Duality Theorem.

(b) This is evident from the proof of part (a).

(c) Take any problem with two constraints that has a duality gap and has an optimal solution at which one of the two constraints is inactive. For example, consider the following problem, which is derived from Example 6.2.1:

$$\begin{aligned} &\text{minimize} \quad f(x) \\ &\text{subject to} \quad x_1 = 0, \quad x_1 \leq 1, \qquad x \in X = \{x \mid x \geq 0\}, \end{aligned}$$

where

$$f(x) = e^{-\sqrt{x_1 x_2}}, \qquad \forall\, x \in X,$$

and $f(x)$ is arbitrarily defined for $x \notin X$.

Consider the problem obtained by keeping the inactive constraint explicit (the constraint $x_1 \leq 1$ in the above example), and by lumping the other constraint together with X to form $\overline{X}$ ($\overline{X} = \{x \mid x \geq 0,\, x_1 = 0\}$ in the above example). Then, we have $q^* < \overline{q}^* = f^*$ ($q^* = 0$ and $\overline{q}^* = f^* = 1$ in the above example).