Convex Optimization Theory

Chapter 5

Exercises and Solutions: Extended Version

Dimitri P. Bertsekas

Massachusetts Institute of Technology

Athena Scientific, Belmont, Massachusetts
http://www.athenasc.com

CHAPTER 5: EXERCISES AND SOLUTIONS†

5.1 (Extended Representation)

Consider the convex programming problem

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & x \in X, \quad g(x) \le 0, \end{array} \tag{5.1}$$

of Section 5.3, and assume that the set X is described by equality and inequality constraints as

$$X = \big\{x \mid l_i(x) = 0,\, i = m+1, \ldots, \overline{m},\, g_j(x) \le 0,\, j = r+1, \ldots, \overline{r}\big\}.$$

Then the problem can alternatively be described without an abstract set constraint, in terms of all of the constraint functions

$$l_i(x) = 0, \quad i = 1, \ldots, \overline{m}, \qquad g_j(x) \le 0, \quad j = 1, \ldots, \overline{r}.$$

We call this the *extended representation* of (P). Show if there is no duality gap and there exists a dual optimal solution for the extended representation, the same is true for the original problem (5.1).

Solution: Assume that there exists a dual optimal solution in the extended representation. Thus there exist nonnegative scalars $\lambda_1^*, \ldots, \lambda_m^*, \lambda_{m+1}^*, \ldots, \lambda_{\overline{m}}^*$ and $\mu_1^*, \ldots, \mu_r^*, \mu_{r+1}^*, \ldots, \mu_{\overline{r}}^*$ such that

$$f^* = \inf_{x \in \Re^n} \left\{ f(x) + \sum_{i=1}^{\overline{m}} \lambda_i^* h_i(x) + \sum_{j=1}^{\overline{r}} \mu_j^* g_j(x) \right\},$$

from which we have

$$f^* \le f(x) + \sum_{i=1}^{\overline{m}} \lambda_i^* h_i(x) + \sum_{j=1}^{\overline{r}} \mu_j^* g_j(x), \qquad \forall\, x \in \Re^n.$$

† This set of exercises will be periodically updated as new exercises are added. Many of the exercises and solutions given here were developed as part of my earlier convex optimization book [BNO03] (coauthored with Angelia Nedić and Asuman Ozdaglar), and are posted on the internet of that book's web site. The contribution of my coauthors in the development of these exercises and their solutions is gratefully acknowledged. Since some of the exercises and/or their solutions have been modified and also new exercises have been added, all errors are my sole responsibility.

For any $x \in X$, we have $h_i(x) = 0$ for all $i = m+1, \ldots, \overline{m}$, and $g_j(x) \le 0$ for all $j = r+1, \ldots, \overline{r}$, so that $\mu_j^* g_j(x) \le 0$ for all $j = r+1, \ldots, \overline{r}$. Therefore, it follows from the preceding relation that

$$f^* \le f(x) + \sum_{i=1}^{m} \lambda_i^* h_i(x) + \sum_{j=1}^{r} \mu_j^* g_j(x), \qquad \forall\, x \in X.$$

Taking the infimum over all $x \in X$, it follows that

$$\begin{aligned}
f^* &\le \inf_{x \in X} \left\{ f(x) + \sum_{i=1}^{m} \lambda_i^* h_i(x) + \sum_{j=1}^{r} \mu_j^* g_j(x) \right\} \\
&\le \inf_{\substack{x \in X,\ h_i(x)=0,\ i=1,\ldots,m \\ g_j(x) \le 0,\ j=1,\ldots,r}} \left\{ f(x) + \sum_{i=1}^{m} \lambda_i^* h_i(x) + \sum_{j=1}^{r} \mu_j^* g_j(x) \right\} \\
&\le \inf_{\substack{x \in X,\ h_i(x)=0,\ i=1,\ldots,m \\ g_j(x) \le 0,\ j=1,\ldots,r}} f(x) \\
&= f^*.
\end{aligned}$$

Hence, equality holds throughout above, showing that the scalars $\lambda_1^*, \ldots, \lambda_m^*$, $\mu_1^*, \ldots, \mu_r^*$ constitute a dual optimal solution for the original representation.

5.2 (Dual Function for Linear Constraints)

Consider the problem

$$\begin{aligned}
&\text{minimize} && f(x) \\
&\text{subject to} && Ax \le b, \qquad Ex = d,
\end{aligned}$$

and the dual function

$$q(\lambda, \mu) = \inf_{x \in \Re^n} \big\{ f(x) + \lambda'(Ex - d) + \mu'(Ax - b) \big\}.$$

Show that

$$q(\lambda, \mu) = -d'\lambda - b'\mu - f^\star(-E'\lambda - A'\mu),$$

where $f^\star$ is the conjugate function of f.

Solution: Evident from the definition of a conjugate convex function

$$f^\star(y) = \sup_{x \in \Re^n} \big\{ y'x - f(x) \big\}.$$

5.3

Consider the two-dimensional problem

$$\begin{array}{ll} \text{minimize} & f(x) = e^{-x_1} \\ \text{subject to} & g(x) \le 0, \qquad x \in X, \end{array}$$

where

$$g(x) = \frac{x_1^2}{x_2}, \qquad X = \big\{(x_1, x_2) \mid x_2 > 0\big\}.$$

Show that g is convex over X, calculate the optimal primal and dual values, and show that there is a duality gap. Verify your finding by considering the corresponding MC/MC framework.

Solution: The Hessian of g for $x \in X$ is given by

$$\nabla^2 g(x) = \begin{pmatrix} 2x_2 & \frac{-2x_1}{x_2^2} \\ \frac{-2x_1}{x_2^2} & \frac{2x_1^2}{x_2^3} \end{pmatrix}.$$

Its determinant is nonnegative over X, so $\nabla^2 g(x)$ is positive semidefinite over X. Therefore, according to Prop. 1.1.10, g is convex over X.

For the MC/MC framework, we construct the set

$$M = \big\{(g(x), f(x)) \mid x \in X\big\}.$$

For $g(x) = 0$, we have $x_1 = 0, f(x) = 1$. For $g(x) = \frac{x_1^2}{x_2} = u > 0$, we have

$$x_1 = \pm\sqrt{ux_2}.$$

Therefore, x_1 changes from $(0, \infty)$, $f(x)$ can take any value in $(0, 1)$, and the set $\overline{M}$ in the MC/MC framework consists of the positive orthant and the halfline $\{(0, w) \mid w \ge 1\}$.

From the description of $\overline{M}$, we can see that $f^* = 1$, while from the description of $\overline{M}$, we can see that the dual function is

$$q(\mu) = \begin{cases} 0 & \text{if } \mu \ge 0, \\ -\infty & \text{if } \mu < 0. \end{cases}$$

Therefore, $q^* = 0$ and there is a duality gap of $f^* - q^* = 1$. Clearly, Slater's condition doesn't hold for this problem.

5.4 (Quadratic Programming Duality)

Consider the quadratic program

$$\begin{aligned} &\text{minimize} \quad c'x + \tfrac{1}{2}x'Qx \\ &\text{subject to} \quad x \in X, \qquad a_j'x \le b_j, \quad j = 1, \ldots, r, \end{aligned}$$

where X is a polyhedral set, Q is a symmetric positive semidefinite $n \times n$ matrix, $c, a_1, \ldots, a_r$ are vectors in $\Re^n$, and $b_1, \ldots, b_r$ are scalars, and assume that its optimal value is finite. Then there exist at least one optimal solution and at least one dual optimal solution. *Hint*: Use the extended representation of Exercise 5.1.

Solution: Consider the extended representation of the problem in which the linear inequalities that represent the polyhedral part are lumped with the remaining linear inequality constraints (cf. Exercise 5.1). From Prop. 1.4.12, finiteness of the optimal value implies that there exists a primal optimal solution and from the analysis of Section 5.3.1, there exists a dual optimal solution. From Exercise 5.1, it follows that there exists a dual optimal solution for the original representation of the problem.

5.5 (Sensitivity)

Consider the class of problems

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

where $u = (u_1, \ldots, u_r)$ is a vector parameterizing the right-hand side of the constraints. Given two distinct values $\bar{u}$ and $\tilde{u}$ of u, let $\overline{f}$ and $\tilde{f}$ be the corresponding optimal values, and assume that $\overline{f}$ and $\tilde{f}$ are finite. Assume further that $\overline{\mu}$ and $\tilde{\mu}$ are corresponding dual optimal solutions and that there is no duality gap. Show that

$$\tilde{\mu}'(\tilde{u} - \bar{u}) \le \overline{f} - \tilde{f} \le \overline{\mu}'(\tilde{u} - \overline{u}).$$

Solution: We have

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

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

Let $\overline{q}(\mu)$ denote the dual function of the problem corresponding to $\overline{u}$:

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

We have

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

where the last inequality holds because $\overline{\mu}$ maximizes $\overline{q}$.

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.

5.6 (Duality and Zero Sum Games)

Let A be an $n \times m$ matrix, and let X and Z be the unit simplices in $\Re^n$ and $\Re^m$, respectively:

$$
X = \left\{ x \;\middle|\; \sum_{i=1}^{n} x_i = 1,\ x_i \geq 0,\ i = 1, \ldots, n \right\},
$$

$$
Z = \left\{ z \;\middle|\; \sum_{j=1}^{m} z_j = 1,\ z_j \geq 0,\ j = 1, \ldots, m \right\}.
$$

Show that the minimax equality

$$
\max_{z \in Z} \min_{x \in X} x'Az = \min_{x \in X} \max_{z \in Z} x'Az
$$

is a special case of linear programming duality. *Hint*: For a fixed z, $\min_{x \in X} x'Az$ is equal to the minimum component of the vector Az, so

$$
\max_{z \in Z} \min_{x \in X} x'Az = \max_{z \in Z} \min \bigl\{ (Az)_1, \ldots, (Az)_n \bigr\} = \max_{\xi e \leq Az,\, z \in Z} \xi, \tag{5.2}
$$

where e is the unit vector in $\Re^n$ (all components are equal to 1). Similarly,

$$
\min_{x \in X} \max_{z \in Z} x'Az = \min_{\zeta e \geq A'x,\, x \in X} \zeta. \tag{5.3}
$$

Show that the linear programs in the right-hand sides of Eqs. (5.2) and (5.3) are dual to each other.

Solution: Consider the linear program

$$
\min_{\substack{\zeta e \geq A'x \\ \sum_{i=1}^n x_i = 1,\, x_i \geq 0}} \zeta,
$$

whose optimal value is equal to $\min_{x\in X}\max_{z\in Z} x'Az$. Introduce dual variables $z\in\Re^m$ and $\xi\in\Re$, corresponding to the constraints $A'x-\zeta e\le 0$ and $\sum_{i=1}^n x_i = 1$, respectively. The dual function is

$$
\begin{aligned}
q(z,\xi) &= \inf_{x_i\ge 0,\, i=1,\ldots,n}\left\{\zeta + z'(A'x-\zeta e) + \xi\left(1-\sum_{i=1}^n x_i\right)\right\}\\
&= \inf_{x_i\ge 0,\, i=1,\ldots,n}\left\{\zeta\left(1-\sum_{j=1}^m z_j\right) + x'(Az-\xi e)+\xi\right\}\\
&= \begin{cases} \xi & \text{if } \sum_{j=1}^m z_j = 1,\ \xi e - Az\le 0,\\ -\infty & \text{otherwise.}\end{cases}
\end{aligned}
$$

Thus the dual problem, which is to maximize $q(z,\xi)$ subject to $z\ge 0$ and $\xi\in\Re$, is equivalent to the linear program

$$
\max_{\xi e\le Az,\, z\in Z} \xi,
$$

whose optimal value is equal to $\max_{z\in Z}\min_{x\in X} x'Az$.

5.7 (Inconsistent Convex Systems of Inequalities)

Let $g_j : \Re^n \mapsto \Re$, $j=1,\ldots,r$, be convex functions over the nonempty convex subset of $\Re^n$. Show that the system

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

has no solution within X if and only if there exists a vector $\mu\in\Re^r$ such that

$$
\sum_{j=1}^r \mu_j = 1, \qquad \mu\ge 0,
$$

$$
\mu' g(x) \ge 0, \qquad \forall\ x\in X.
$$

Hint: Consider the convex program

$$
\begin{aligned}
&\text{minimize } y\\
&\text{subject to } x\in X,\quad y\in\Re,\qquad g_j(x)\le y,\quad j=1,\ldots,r.
\end{aligned}
$$

Solution: The dual function for the problem in the hint is

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

The problem in the hint satisfies the Slater condition, so 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.

5.8 (Finiteness of the Optimal Dual Value)

Consider the problem

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

where X is a convex set, and f and g_j are convex over X. Assume that the problem has at least one feasible solution. Show that the following are equivalent.

(i) The dual optimal value $q^* = \sup_{\mu \in \Re^r} q(\mu)$ is finite.

(ii) The primal function p is proper.

(iii) The set

$$M = \big\{(u, w) \in \Re^{r+1} \mid \text{there is an } x \in X \text{ such that } g(x) \leq u,\ f(x) \leq w\big\}$$

does not contain a vertical line.

Solution: We note that $-q$ is closed and convex, and that

$$q(\mu) = \inf_{u \in \Re^r} \big\{p(u) + \mu' u\big\}, \qquad \forall\ \mu \in \Re^r.$$

Since $q(\mu) \leq p(0)$ for all $\mu \in \Re^r$, given the feasibility of the problem [i.e., $p(0) < \infty$], we see that q^* is finite if and only if q is proper. Since q is the conjugate of $p(-u)$ and p is convex, by the Conjugacy Theorem [Prop. 1.6.1(b)], q is proper if and only if p is proper. Hence (i) is equivalent to (ii).

We note that the epigraph of p is the closure of M. Hence, given the feasibility of the problem, (ii) is equivalent to the closure of M not containing a vertical line. Since M is convex, its closure does not contain a line if and only if M does not contain a line (since the closure and the relative interior of M have the same recession cone). Hence (ii) is equivalent to (iii).

5.9

Show that for the function $f(x) = \|x\|$, we have

$$\partial f(x) = \begin{cases} \{x/\|x\|\} & \text{if } x \neq 0, \\ \{g \mid \|g\| \leq 1\} & \text{if } x = 0. \end{cases}$$

Solution: For $x \neq 0$, the function $f(x) = \|x\|$ is differentiable with $\nabla f(x) = x/\|x\|$, so that $\partial f(x) = \{\nabla f(x)\} = \{x/\|x\|\}$. Consider now the case $x = 0$. If a vector d is a subgradient of f at $x = 0$, then $f(z) \geq f(0) + d'z$ for all z, implying that

$$\|z\| \geq d'z, \qquad \forall\, z \in \Re^n.$$

By letting $z = d$ in this relation, we obtain $\|d\| \leq 1$, showing that $\partial f(0) \subset \{d \mid \|d\| \leq 1\}$.

On the other hand, for any $d \in \Re^n$ with $\|d\| \leq 1$, we have

$$d'z \leq \|d\| \cdot \|z\| \leq \|z\|, \qquad \forall\, z \in \Re^n,$$

which is equivalent to $f(0) + d'z \leq f(z)$ for all z, so that $d \in \partial f(0)$, and therefore $\{d \mid \|d\| \leq 1\} \subset \partial f(0)$.

5.10

Let $f : \Re^n \mapsto (-\infty, \infty]$ be a proper convex function, and let x and y be vectors in $\Re^n$. Show that if $g_x \in \partial f(x)$ and $g_y \in \partial f(y)$, then

$$(g_x - g_y)'(x - y) \geq 0.$$

Note: This is a generalization of a well-known inequality for gradients of a differentiable convex function. *Hint*: Write the subgradient inequalities for x, y, g_x and for x, y, g_y, and add.

Solution: Following the hint, we write

$$f(y) \geq f(x) + g_x'(y - x),$$

$$f(x) \geq f(y) + g_y'(x - y).$$

By adding, we obtain

$$(g_x - g_y)'(x - y) \geq 0.$$

5.11

Let $f : \Re^n \mapsto \Re$ be a convex function, and let x and y be given vectors in $\Re^n$. Consider the scalar function $\varphi : \Re \mapsto \Re$ defined by $\varphi(t) = f\big(tx + (1-t)y\big)$ for all $t \in \Re$, and show that

$$\partial\varphi(t) = \Big\{(x-y)'g \,\Big|\, g \in \partial f\big(tx + (1-t)y\big)\Big\}, \qquad \forall\, t \in \Re.$$

Hint: Apply the chain rule of Prop. 5.4.5.

Solution: We can view the function

$$\varphi(t) = f\big(tx + (1-t)y\big), \qquad t \in \Re$$

as the composition of the form

$$\varphi(t) = f\big(g(t)\big), \qquad t \in \Re,$$

where $g(t) : \Re \mapsto \Re^n$ is an affine function given by

$$g(t) = y + t(x-y), \qquad t \in \Re.$$

By using the Chain Rule (Prop. 5.4.5), where $A = (x-y)$, we obtain

$$\partial\varphi(t) = A'\,\partial f\big(g(t)\big), \qquad \forall\, t \in \Re,$$

or equivalently

$$\partial\varphi(t) = \Big\{(x-y)'d \mid d \in \partial f\big(tx + (1-t)y\big)\Big\}, \qquad \forall\, t \in \Re.$$

5.12 (Partial Differentiation)

Consider a proper convex function F of two vectors $x \in \Re^n$ and $y \in \Re^m$. For a fixed $(\overline{x}, \overline{y}) \in \mathrm{dom}(F)$, let $\partial_x F(\overline{x}, \overline{y})$ and $\partial_y F(\overline{x}, \overline{y})$ be the subdifferentials of the functions $F(\cdot, \overline{y})$ and $F(\overline{x}, \cdot)$ at $\overline{x}$ and $\overline{y}$, respectively.

(a) Show that

$$\partial F(\overline{x}, \overline{y}) \subset \partial_x F(\overline{x}, \overline{y}) \times \partial_y F(\overline{x}, \overline{y}),$$

and give an example showing that the inclusion may be strict in general.

(b) Assume that F has the form

$$F(x, y) = h_1(x) + h_2(y) + h(x, y),$$

where h_1 and h_2 are proper convex functions, and h is convex, real-valued, and differentiable. Show that the formula of part (a) holds with equality.

Solution: (a) We have $(g_x, g_y) \in \partial F(\overline{x}, \overline{y})$ if and only if

$$F(x, y) \geq F(\overline{x}, \overline{y}) + g_x'(x - \overline{x}) + g_y'(y - \overline{y}), \qquad \forall\ x \in \Re^n,\ y \in \Re^m.$$

By setting $y = \overline{y}$, we obtain that $g_x \in \partial_x F(\overline{x}, \overline{y})$, and by setting $x = \overline{x}$, we obtain that $g_y \in \partial_y F(\overline{x}, \overline{y})$, so that $(g_x, g_y) \in \partial_x F(\overline{x}, \overline{y}) \times \partial_y F(\overline{x}, \overline{y})$.

For an example where the inclusion is strict, consider any function whose subdifferential is not a Cartesian product at some point, such as $F(x, y) = |x+y|$ at points $(\overline{x}, \overline{y})$ with $\overline{x} + \overline{y} = 0$.

(b) Since F is the sum of functions of the given form, Prop. 5.4.6 applies and shows that

$$\partial F(\overline{x}, \overline{y}) = \big\{(g_x, 0) \mid g_x \in \partial h_1(\overline{x})\big\} + \big\{(0, g_y) \mid g_y \in \partial h_2(\overline{y})\big\} + \big\{\nabla h(\overline{x}, \overline{y})\big\}$$

[the relative interior condition of the proposition is clearly satisfied]. Since

$$\nabla h(\overline{x}, \overline{y}) = \big(\nabla_x h(\overline{x}, \overline{y}), \nabla_y h(\overline{x}, \overline{y})\big),$$

$$\partial_x F(\overline{x}, \overline{y}) = \partial h_1(\overline{x}) + \nabla_x h(\overline{x}, \overline{y}),$$

$$\partial_y F(\overline{x}, \overline{y}) = \partial h_2(\overline{y}) + \nabla_y h(\overline{x}, \overline{y}),$$

the result follows.

5.13 (Normal Cones of Level Sets)

Let $f : \Re^n \mapsto (-\infty, \infty]$ be a proper convex function, and let x be a vector that does not achieve a minimum of f, and is such that $\partial f(x) \neq \emptyset$.

(a) Show that the normal cone $N_V(x)$, where

$$V = \big\{z \mid f(z) \leq f(x)\big\},$$

is the closure of the convex cone generated by $\partial f(x)$.

(b) Show that if $x \in \text{int}\big(\text{dom}(f)\big)$, then $N_V(x)$ is the convex cone generated by $\partial f(x)$.

Solution: It is easy to show that $\text{cl}(\text{cone}(\partial f(x))) \subset N_v(x)$, so to prove that $N_v(x) = \text{cl}(\text{cone}(\partial f(x)))$, we show $N_v(x) \subset \text{cl}(\text{cone}(\partial f(x)))$. We will prove this by contradiction. Suppose the set $N_v(x)$ is strictly larger than $\text{cl}(\text{cone}(\partial f(x)))$. Since both sets are closed, the polar cone $(\text{cl}(\text{cone}(\partial f(x))))^*$ is strictly larger than the polar cone $(N_v(x))^*$. Thus there exists a pair of directions (y, d_1) such that

$$y' d_1 > 0, \quad y \in N_v(x) \setminus \text{clcone}(\partial f(x))), \quad d_1 \in (\text{cl}(\text{cone}(\partial f(x)))^*.$$

Furthermore, since x is not the minimum of f, we have $0 \notin \partial f(x)$, therefore the set

$$(\text{cl}(\text{cone}(\partial f(x))))^* \cap \text{aff}(\text{cone}(\partial f(x)))$$

contains points other than the origin. Hence we can choose a direction d_2 such that

$$d_2 \in (\mathrm{cl}(\mathrm{cone}(\partial f(x))))^* \cap \mathrm{aff}(\mathrm{cone}(\partial f(x))),$$

and

$$y'd > 0, \quad d \in (\mathrm{cl}(\mathrm{cone}(\partial f(x))))^*, \quad d \notin (\mathrm{aff}(\mathrm{cone}(\partial f(x))))^\perp.$$

where $d = d_1 + d_2$. Next, we show that $\sup_{z\in\partial f(x)} d'z < -\delta$ for some positive δ. Since $d \in (\mathrm{cl}(\mathrm{cone}(\partial f(x))))^*$, it can be seen that $\sup_{z\in\partial f(x)} d'z \leq 0$. To claim that strict inequality holds, since $0 \notin \partial f(x)$ and $\partial f(x)$ is a closed set, and furthermore, $d \notin (\mathrm{aff}(\mathrm{cone}(\partial f(x))))^\perp$, it follows that there exists some $\delta > 0$ such that

$$\sup_{z\in\partial f(x)} d'z < -\delta.$$

Finally, since the closure of the directional derivative function $f'(x;\cdot)$ is the support function of $\partial f(x)$, we have

$$(cl f')(x;d) = \sup_{z\in\partial f(x)} d'z < -\delta.$$

Therefore, by the definition of the closure of a function, we can choose a direction $\bar{d}$ sufficiently close to d such that

$$f'(x;\bar{d}) < 0, \quad y'\bar{d} > 0,$$

where the second inequality is possible because $y'd > 0$. The fact $f'(x;\bar{d}) < 0$ implies that along the direction $\bar{d}$ we can find a point z such that $f(z) < f(x)$. Since $y \in N_v$, we have $y'(z - x) \leq 0$, and this contradicts $y'\bar{d} > 0$.

5.14 (Functions of Legendre Type)

A convex function f on $\Re^n$ is said to be of *Legendre type* if it is real-valued, everywhere differentiable, and is such that for all $g \in \Re^n$, the minimum of $f(x) - g'x$ over $x \in \Re^n$ is attained at a unique point.

(a) Show that if f is real-valued, everywhere differentiable, and strictly convex, and $f(x) - g'x$ is coercive as a function of x for each g, it is of Legendre type. Furthermore, this is true if f is strongly convex [for example it is twice differentiable, and is such that for some $\alpha > 0$, $\nabla^2 f(x) - \alpha I$ is positive definite for all x; see the exercises of Chapter 1, Exercise 1.11].

(b) Show that for $1 < p < \infty$, the p-th power of the p-norm

$$f(x) = \sum_{i=1}^{n} |x_i|^p,$$

and all positive definite quadratic functions are functions of Legendre type.

(c) Give an example of a twice differentiable function, with everywhere positive definite Hessian, which is not of Legendre type.

(d) A convex function is of Legendre type if and only if its conjugate is of Legendre type.

Solution: (a) Consider the function $f_g(x) = f(x) - g'x$, which is real-valued and convex for any value of g. Since $f_g(x)$ is convex and coercive, it has an nonempty and compact set of minima. Also $f_g(x)$ is strictly convex by strict convexity of f, which implies that it has an unique global minimal point.

To sum up, $f(x)$ is real-valued, everywhere differentiable, and is such that for all $g \in \Re^n$, the minimum of $f(x) - g'x = f_g(x)$ over $x \in \Re^n$ is attained at a unique point. It follows that f is of Legendre type.

If f is strongly convex, the same is true for f_g, so by the result of Exercise 1.11, the minimum of f_g is attained at a unique point. In the special case where f is twice differentiable, and is such that for some $a > 0$, $\nabla^2 f(x) - aI$ is positive definite for all $x \in \Re^n$. A more direct proof is possible. Then the Hessian matrix of $f_g(x)$ also satisfies $\nabla^2 f_g(x) - aI > 0$ for all x. It follows that $\lim_{\|x\| \to \infty} f_g(x) \geq \lim_{\|x\| \to \infty} (f_g(0) + x'(aI)x) = \infty$, where $f_g(0) \neq -\infty$ since f is real-valued. Thus f_g is coercive and it follows that f is of Legendre type.

(b) Consider the p-th power of the p-norm, $f(x) = \Sigma_{i=1}^n \|x_i\|^p$, $x = (x_1, x_2, ..., x_n)'$, $1 < p < \infty$. Note that f is real-valued, everywhere differentiable, and we have

$$\partial f / \partial x_i = \begin{cases} 0, & \text{if } x_i = 0 \\ p x_i^{p-1}, & \text{if } x_i > 0 \\ -p(-x_i)^{p-1}, & \text{if } x_i < 0 \end{cases} .$$

The function $f(x) - g'x$ is coercive, because

$$\lim_{x \leftarrow \infty} \Sigma_{i=1}^n \|x_i\|^n - g'x \geq \Sigma_{i=1}^n \|x_i\|^p - \|g'x\| = \infty.$$

Using part (a) it follows that f is of Legendre type.

Consider an arbitrary positive definite quadratic function

$$f(x) = x'Hx + Ax + b.$$

The Hessian of f is the positive definite matrix H, so f is strongly convex. Using part (a) it follows that f is of Legendre type.

(c) Consider the exponential function $f(x) = e^{-x}$. Then f is twice differentiable, with everywhere positive definite Hessian. However, f is not of Legendre type, because the function $f(x) - g'x$ with $g > 0$ does not have a minimum.

(d) Let $f : \Re^n \mapsto \Re$ be of Legendre type, and let h be its conjugate. Then, for every $g \in \Re^n$, the minimum of $f(x) - g'x$ over $x \in \Re^n$ is attained at a unique point, denoted $x(g)$. It follows that $h(g)$ is the real number $g'x(g) - f\big(x(g)\big)$, and by the equivalence of (i) and (iii) in Prop. 5.4.3, the unique subgradient of h at g is $x(g)$, so h is everywhere differentiable. By the equivalence of (i) and (ii) in Prop. 5.4.3, for every x, $\nabla f(x)$ is the unique point at which $h(g) - x'g$ attains its minimum over $g \in \Re^n$. Hence h is of Legendre type.

By interchanging the roles of f and h in the preceding argument, and by using the Conjugacy Theorem [Prop. 1.6.1(c)], it follows that if h is of Legendre type, then f is also of Legendre type.

5.15 (Duality Gap and Nondifferentiabilities)

This exercise shows how a duality gap results in nondifferentiability of the dual function. Consider the problem

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

and assume that for all $\mu \ge 0$, the infimum of the Lagrangian $L(x, \mu)$ over X is attained by at least one $x_\mu \in X$. Show that if there is a duality gap, then the dual function $q(\mu) = \inf_{x \in X} L(x, \mu)$ is nondifferentiable at every dual optimal solution. *Hint*: If q is differentiable at a dual optimal solution μ^*, by the theory of Section 5.3, we must have $\partial q(\mu^*)/\partial \mu_j \le 0$ and $\mu_j^* \partial q(\mu^*)/\partial \mu_j = 0$ for all j. Use optimality conditions for μ^*, together with any vector x_{μ^*} that minimizes $L(x, \mu^*)$ over X, to show that there is no duality gap.

Solution: To obtain a contradiction, assume that q is differentiable at some dual optimal solution $\mu^* \in M$, where $M = \{\mu \in \Re^r \mid \mu \ge 0\}$. Then

$$
\nabla q(\mu^*)(\mu^* - \mu) \ge 0, \qquad \forall\ \mu \ge 0.
$$

If $\mu_j^* = 0$, then by letting $\mu = \mu^* + \gamma e_j$ for a scalar $\gamma \ge 0$, and the vector e_j whose jth component is 1 and the other components are 0, from the preceding relation we obtain $\partial q(\mu^*)/\partial \mu_j \le 0$. Similarly, if $\mu_j^* > 0$, then by letting $\mu = \mu^* + \gamma e_j$ for a sufficiently small scalar γ (small enough so that $\mu^* + \gamma e_j \in M$), from the preceding relation we obtain $\partial q(\mu^*)/\partial \mu_j = 0$. Hence

$$
\partial q(\mu^*)/\partial \mu_j \le 0, \qquad \forall\ j = 1, \ldots, r,
$$

$$
\mu_j^* \partial q(\mu^*)/\partial \mu_j = 0, \qquad \forall\ j = 1, \ldots, r.
$$

Since q is differentiable at μ^*, we have that

$$
\nabla q(\mu^*) = g(x^*),
$$

for some vector $x^* \in X$ such that $q(\mu^*) = L(x^*, \mu^*)$. This and the preceding two relations imply that x^* and μ^* satisfy the necessary and sufficient optimality conditions for an optimal primal and dual optimal solution pair. It follows that there is no duality gap, a contradiction.

5.16 (Saddle Points for Quadratic Functions)

Consider a quadratic function $\phi : X \times Z \mapsto \Re$ of the form

$$
\phi(x, z) = x'Qx + c'x + z'Mx - z'Rz - d'z,
$$

where Q and R are symmetric positive semidefinite $n \times n$ and $m \times m$ matrices, respectively, M is an $n \times m$ matrix, $c \in \Re^n$, $d \in \Re^m$, and X and Z are subsets of $\Re^n$ and $\Re^m$, respectively. Derive conditions based on Prop. 5.5.6 for ϕ to have at least one saddle point.

Solution: We assume that X and Z are closed and convex sets. We will calculate the recession cone and the lineality space of the functions

$$t(x) = \begin{cases} \sup_{z\in Z} \phi(x,z) & \text{if } x \in X, \\ \infty & \text{if } x \notin X, \end{cases}$$

and

$$r(z) = \begin{cases} -\inf_{x\in X} \phi(x,z) & \text{if } z \in Z, \\ \infty & \text{if } z \notin Z, \end{cases}$$

and we will then use Prop. 5.5.6.

For each $z \in Z$, the function $t_z : \Re^n \mapsto (-\infty, \infty]$ defined by

$$t_z(x) = \begin{cases} \phi(x,z) & \text{if } x \in X, \\ \infty & \text{if } x \notin X, \end{cases}$$

is closed and convex in view of the assumption that Q is positive semidefinite symmetric. Similarly, for each $x \in X$, the function $r_x : \Re^m \mapsto (-\infty, \infty]$ defined by

$$r_x(z) = \begin{cases} -\phi(x,z) & \text{if } z \in Z, \\ \infty & \text{if } z \notin Z, \end{cases}$$

is closed and convex in view of the assumption that R is a positive semidefinite symmetric. Hence, Assumption 3.3.1 is satisfied. By the positive semidefiniteness of Q and the calculations of Example 2.1.1, it can be seen that for each $z \in Z$, the recession cone of the function t_z is given by

$$R_{t_z} = R_X \cap N(Q) \cap \{y \mid y'(Mz + c) \le 0\},$$

where R_X is the recession cone of the convex set X and $N(Q)$ is the null space of the matrix Q. Similarly, for each $z \in Z$, the constancy space of the function t_z is given by

$$L_{t_z} = L_X \cap N(Q) \cap \{y \mid y'(Mz + c) = 0\},$$

where L_X is the lineality space of the set X. By the positive semidefiniteness of R, for each $x \in X$, it can be seen that the recession cone of the function r_x is given by

$$R_{r_x} = R_Z \cap N(R) \cap \{y \mid y'(Mx - d) \ge 0\},$$

where R_Z is the recession cone of the convex set Z and $N(R)$ is the null space of the matrix R. Similarly, for each $x \in X$, the constancy space of the function r_x is given by

$$L_{r_x} = L_Z \cap N(R) \cap \{y \mid y'(Mx - d) = 0\},$$

where L_Z is the lineality space of the set Z.

Since $t(x) = \sup_{z\in Z} t_z(x)$, the recession cone of t is

$$R_t = \cap_{z\in Z} R_{t_z} = R_X \cap N(Q) \cap \left(\cap_{z\in Z}\{y \mid y'(Mz + c) \le 0\}\right),$$

or

$$R_t = R_X \cap N(Q) \cap (MZ + c)^*,$$

where

$$(MZ + c)^* = \{y \mid y'w \le 0, \forall\, w \in MZ + c\}.$$

[Note that $(MZ + c)^*$ is the polar cone of the set $MZ + c$, as defined in Chapter 4.] Similarly, the lineality space of t is

$$L_t = L_X \cap N(Q) \cap (MZ)^\perp,$$

where

$$(MZ + c)^\perp = \{y \mid y'w = 0, \forall\, w \in MZ + c\}.$$

By the same calculations, we also have

$$R_r = R_Z \cap N(R) \cap (-MX + d)^*, \qquad L_r = L_Z \cap N(R) \cap (-MX + d)^\perp,$$

where

$$(-MX + d)^* = \{y \mid y'w \le 0, \forall\, w \in -MX + d\},$$
$$(-MX + d)^\perp = \{y \mid y'w = 0, \forall\, w \in -MX + d\}.$$

If

$$R_t = R_r = \{0\},$$

then it follows from Prop. 5.5.6(a) that the set of saddle points of ϕ is nonempty and compact. (In particular, this condition holds if Q and R are positive definite matrices, or if X and Z are compact.)

Similarly, if

$$R_t = L_t, \qquad R_r = L_r,$$

then it follows from Prop. 5.5.6(b) that the set of saddle points of ϕ is nonempty.

5.17 (Goldman-Tucker Complementarity Theorem [GoT56])

Consider the linear programming problem

$$\begin{aligned} &\text{minimize} \quad c'x \\ &\text{subject to} \quad Ax = b, \qquad x \ge 0, \end{aligned} \tag{LP}$$

where A is an $m \times n$ matrix, c is a vector in $\Re^n$, and b is a vector in $\Re^m$. Consider also the dual problem

$$\begin{aligned} &\text{maximize} \quad b'\lambda \\ &\text{subject to} \quad A'\lambda \le c. \end{aligned} \tag{DLP}$$

Assume that the sets of optimal solutions of LP and DLP, denoted X^* and Λ^*, respectively, are nonempty. Show that the index set $\{1, \ldots, n\}$ can be partitioned into two disjoint subsets I and $\overline{I}$ with the following two properties:

(1) For all $x^* \in X^*$ and $\lambda^* \in \Lambda^*$, we have

$$x_i^* = 0, \qquad \forall\, i \in \overline{I}, \qquad (A'\lambda^*)_i = c_i, \qquad \forall\, i \in I,$$

where x_i^* and $(A'\lambda^*)_i$ are the ith components of x^* and $A'\lambda^*$, respectively.

(2) There exist vectors $x^* \in X^*$ and $\lambda^* \in \Lambda^*$ such that

$$x_i^* > 0, \quad \forall\, i \in I, \qquad x_i^* = 0, \quad \forall\, i \in \overline{I},$$

$$(A'\lambda^*)_i = c_i, \quad \forall\, i \in I, \qquad (A'\lambda^*)_i < c_i, \quad \forall\, i \in \overline{I}.$$

Hint: Apply the Tucker Complementarity Theorem (Exercises of Chapter 2).

Solution: Consider the subspace

$$S = \big\{(x, w) \mid bw - Ax = 0,\, c'x = wv,\, x \in \Re^n,\, w \in \Re\big\},$$

where v is the optimal value of (LP). Its orthogonal complement is the range of the matrix

$$\begin{pmatrix} -A' & c \\ b & -v \end{pmatrix},$$

so it has the form

$$S^\perp = \big\{(c\zeta - A'\lambda, b'\lambda - v\zeta) \mid \lambda \in \Re^m,\, \zeta \in \Re\big\}.$$

Applying the Tucker Complementarity Theorem (Exercises of Chapter 2) for this choice of S, we obtain a partition of the index set $\{1, \ldots, n+1\}$ in two subsets. There are two possible cases: (1) the index $n+1$ belongs to the first subset, or (2) the index $n+1$ belongs to the second subset. Since the vectors $(x, 1)$ such that $x \in X^*$ satisfy $Ax - bw = 0$ and $c'x = wv$, we see that case (1) holds, i.e., the index $n+1$ belongs to the first index subset. In particular, we have that there exist disjoint index sets I and $\overline{I}$ such that $I \cup \overline{I} = \{1, \ldots, n\}$ and the following properties hold:

(a) There exist vectors $(x, w) \in S$ and $(\lambda, \zeta) \in \Re^{m+1}$ with the property

$$x_i > 0, \quad \forall\, i \in I, \qquad x_i = 0, \quad \forall\, i \in \overline{I}, \qquad w > 0, \tag{5.4}$$

$$c_i\zeta - (A'\lambda)_i = 0, \quad \forall\, i \in I, \quad c_i\zeta - (A'\lambda)_i > 0, \quad \forall\, i \in \overline{I}, \qquad b'\lambda = v\zeta. \tag{5.5}$$

(b) For all $(x, w) \in S$ with $x \geq 0$, and $(\lambda, \zeta) \in \Re^{m+1}$ with $c\zeta - A'\lambda \geq 0$, $v\zeta - b'\lambda \geq 0$, we have

$$x_i = 0, \qquad \forall\, i \in \overline{I},$$

$$c_i\zeta - (A'\lambda)_i = 0, \qquad \forall\, i \in I, \qquad b'\lambda = v\zeta.$$

By dividing (x, w) by w, we obtain [cf. Eq. (5.4)] an optimal primal solution $x^* = x/w$ such that

$$x_i^* > 0, \quad \forall\, i \in I, \qquad x_i^* = 0, \quad \forall\, i \in \overline{I}.$$

Similarly, if the scalar ζ in Eq. (5.5) is positive, by dividing with ζ in Eq. (5.5), we obtain an optimal dual solution $\lambda^* = \lambda/\zeta$, which satisfies the desired property

$$c_i - (A'\lambda^*)_i = 0, \quad \forall\, i \in I, \qquad c_i - (A'\lambda^*)_i > 0, \quad \forall\, i \in \overline{I}.$$

If the scalar ζ in Eq. (5.5) is nonpositive, we choose any optimal dual solution λ^*, and we note, using also property (b), that we have

$$c_i - (A'\lambda^*)_i = 0, \qquad \forall\, i \in I, \qquad c_i - (A'\lambda^*)_i \geq 0, \quad \forall\, i \in \overline{I}, \qquad b'\lambda^* = v. \tag{5.6}$$

Consider the vector

$$\tilde{\lambda} = (1 - \zeta)\lambda^* + \lambda.$$

By multiplying Eq. (5.6) with the positive number $1 - \zeta$, and by combining it with Eq. (5.5), we see that

$$c_i - (A'\tilde{\lambda})_i = 0, \qquad \forall\, i \in I, \qquad c_i - (A'\tilde{\lambda})_i > 0, \quad \forall\, i \in \overline{I}, \qquad b'\tilde{\lambda} = v.$$

Thus, $\tilde{\lambda}$ is an optimal dual solution that satisfies the desired property.

REFERENCES

[GoT56] Goldman, A. J., and Tucker, A. W., 1956. "Theory of Linear Programming," in Linear Inequalities and Related Systems, H. W. Kuhn and A. W. Tucker, eds., Princeton University Press, Princeton, N.J., pp. 53-97.