# *Convex Optimization Theory*

# *Chapter 3*
# *Exercises and Solutions*

Dimitri P. Bertsekas

**Massachusetts Institute of Technology**

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

## CHAPTER 3: EXERCISES AND SOLUTIONS†

### 3.1 (Local Minima Along Lines)

(a) Consider a vector $x^*$ such that a given function $f : \Re^n \mapsto \Re$ is convex over a sphere centered at $x^*$. Show that $x^*$ is a local minimum of $f$ if and only if it is a local minimum of $f$ along every line passing through $x^*$ [i.e., for all $d \in \Re^n$, the function $g : \Re \mapsto \Re$, defined by $g(\alpha) = f(x^* + \alpha d)$, has $\alpha^* = 0$ as its local minimum].

(b) Consider the nonconvex function $f : \Re^2 \mapsto \Re$ given by

$$f(x_1, x_2) = (x_2 - px_1^2)(x_2 - qx_1^2),$$

where $p$ and $q$ are scalars with $0 < p < q$, and $x^* = (0, 0)$. Show that $f(y, my^2) < 0$ for $y \neq 0$ and $m$ satisfying $p < m < q$, so $x^*$ is not a local minimum of $f$ even though it is a local minimum along every line passing through $x^*$.

**Solution:** (a) If $x^*$ is a local minimum of $f$, evidently it is also a local minimum of $f$ along any line passing through $x^*$.

Conversely, let $x^*$ be a local minimum of $f$ along any line passing through $x^*$. Assume, to arrive at a contradiction, that $x^*$ is not a local minimum of $f$ and that we have $f(\overline{x}) < f(x^*)$ for some $\overline{x}$ in the sphere centered at $x^*$ within which $f$ is assumed convex. Then, by convexity of $f$, for all $\alpha \in (0, 1)$, we have

$$f\big(\alpha x^* + (1 - \alpha)\overline{x}\big) \leq \alpha f(x^*) + (1 - \alpha) f(\overline{x}) < f(x^*),$$

so $f$ decreases monotonically along the line segment connecting $x^*$ and $\overline{x}$. This contradicts the hypothesis that $x^*$ is a local minimum of $f$ along any line passing through $x^*$.

(b) Consider the function $f(x_1, x_2) = (x_2 - px_1^2)(x_2 - qx_1^2)$, where $0 < p < q$ and let $x^* = (0, 0)$.

We first show that $g(\alpha) = f(x^* + \alpha d)$ is minimized at $\alpha = 0$ for all $d \in \Re^2$. We have

$$g(\alpha) = f(x^* + \alpha d) = (\alpha d_2 - p\alpha^2 d_1^2)(\alpha d_2 - q\alpha^2 d_1^2) = \alpha^2 (d_2 - p\alpha d_1^2)(d_2 - q\alpha d_1^2).$$

† 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.

Also,

$$g'(\alpha) = 2\alpha(d_2 - p\alpha d_1^2)(d_2 - q\alpha d_1^2) + \alpha^2(-pd_1^2)(d_2 - q\alpha d_1^2) + \alpha^2(d_2 - p\alpha d_1^2)(-qd_1^2).$$

Thus $g'(0) = 0$. Furthermore,

$$\begin{aligned} g''(\alpha) = {} & 2(d_2 - p\alpha d_1^2)(d_2 - q\alpha d_1^2) + 2\alpha(-pd_1^2)(d_2 - q\alpha d_1^2) \\ & + 2\alpha(d_2 - p\alpha d_1^2)(-qd_1^2) + 2\alpha(-pd_1^2)(d_2 - q\alpha d_1^2) + \alpha^2(-pd_1^2)(-qd_1^2) \\ & + 2\alpha(d_2 - p\alpha d_1^2)(-qd_1^2) + \alpha^2(-pd_1^2)(-qd_1^2). \end{aligned}$$

Thus $g''(0) = 2d_2^2$, which is greater than 0 if $d_2 \neq 0$. If $d_2 = 0$, $g(\alpha) = pq\alpha^4 d_1^4$, which is clearly minimized at $\alpha = 0$. Therefore, $(0,0)$ is a local minimum of $f$ along every line that passes through $(0,0)$.

We now show that if $p < m < q$, $f(y, my^2) < 0$ if $y \neq 0$ and that $f(y, my^2) = 0$ otherwise. Consider a point of the form $(y, my^2)$. We have $f(y, my^2) = y^4(m-p)(m-q)$. Clearly, $f(y, my^2) < 0$ if and only if $p < m < q$ and $y \neq 0$. In any $\epsilon-$neighborhood of $(0,0)$, there exists a $y \neq 0$ such that for some $m \in (p,q)$, $(y, my^2)$ also belongs to the neighborhood. Since $f(0,0) = 0$, we see that $(0,0)$ is not a local minimum.

### 3.2 (Equality-Constrained Quadratic Programming)

(a) Consider the quadratic program

$$\begin{aligned} &\text{minimize} \quad \tfrac{1}{2}\|x\|^2 + c'x \\ &\text{subject to} \quad Ax = 0, \end{aligned} \tag{3.1}$$

where $c \in \Re^n$ and $A$ is an $m \times n$ matrix of rank $m$. Use the Projection Theorem to show that

$$x^* = -\big(I - A'(AA')^{-1}A\big)c \tag{3.2}$$

is the unique solution.

(b) Consider the more general quadratic program

$$\begin{aligned} &\text{minimize} \quad \tfrac{1}{2}(x - \overline{x})'Q(x - \overline{x}) + c'(x - \overline{x}) \\ &\text{subject to} \quad Ax = b, \end{aligned} \tag{3.3}$$

where $c$ and $A$ are as before, $Q$ is a symmetric positive definite matrix, $b \in \Re^m$, and $\overline{x}$ is a vector in $\Re^n$, which is feasible, i.e., satisfies $A\overline{x} = b$. Use the transformation $y = Q^{1/2}(x - \overline{x})$ to write this problem in the form of part (a) and show that the optimal solution is

$$x^* = \overline{x} - Q^{-1}(c - A'\lambda),$$

where $\lambda$ is given by
$$\lambda = \left(AQ^{-1}A'\right)^{-1} AQ^{-1}c.$$

(c) Apply the result of part (b) to the program

$$\begin{aligned} &\text{minimize} \quad \tfrac{1}{2}x'Qx + c'x \\ &\text{subject to} \quad Ax = b, \end{aligned}$$

and show that the optimal solution is

$$x^* = -Q^{-1}\left(c - A'\lambda - A'(AQ^{-1}A')^{-1}b\right).$$

**Solution:** (a) Consider the quadratic programming problem

$$\begin{aligned} &\text{minimize} \quad \tfrac{1}{2}\|x\|^2 + c'x \\ &\text{subject to} \quad Ax = 0, \end{aligned} \tag{3.4}$$

where $c$ is a given vector in $\Re^n$ and $A$ is an $m \times n$ matrix of rank $m$. By adding the constant term $\frac{1}{2}\|c\|^2$ to the cost function, we can equivalently write this problem as

$$\begin{aligned} &\text{minimize} \quad \tfrac{1}{2}\|c + x\|^2 \\ &\text{subject to} \quad Ax = 0, \end{aligned}$$

which is the problem of projecting the vector $-c$ on the subspace $X = \{x \mid Ax = 0\}$. By the optimality condition or projection, a vector $x^*$ such that $Ax^* = 0$ is the unique projection if and only if

$$(c + x^*)'x = 0, \qquad \forall\ x \text{ with } Ax = 0.$$

It can be seen that the vector

$$x^* = -\left(I - A'(AA')^{-1}A\right)c \tag{3.5}$$

satisfies this condition and is thus the unique solution of the quadratic programming problem (3.4). (The matrix $AA'$ is invertible because $A$ has rank $m$.)

(b) Consider now the more general quadratic program

$$\begin{aligned} &\text{minimize} \quad \tfrac{1}{2}(x - \overline{x})'Q(x - \overline{x}) + c'(x - \overline{x}) \\ &\text{subject to} \quad Ax = b, \end{aligned} \tag{3.6}$$

where $c$ and $A$ are as before, $Q$ is a symmetric positive definite matrix, $b$ is a given vector in $\Re^m$, and $\overline{x}$ is a given vector in $\Re^n$, which is feasible, that is, it satisfies $A\overline{x} = b$. By introducing the transformation $y = Q^{1/2}(x - \overline{x})$, we can write this problem as

$$\begin{aligned} &\text{minimize} \quad \tfrac{1}{2}\|y\|^2 + \left(Q^{-1/2}c\right)'y \\ &\text{subject to} \quad AQ^{-1/2}y = 0. \end{aligned}$$

Using Eq. (3.5) we see that the solution of this problem is

$$y^* = -\left(I - Q^{-1/2}A'\left(AQ^{-1}A'\right)^{-1}AQ^{-1/2}\right)Q^{-1/2}c$$

and by passing to the $x$-coordinate system through the inverse transformation $x^* - \overline{x} = Q^{-1/2}y^*$, we obtain the optimal solution

$$x^* = \overline{x} - Q^{-1}(c - A'\lambda), \tag{3.7}$$

where the vector $\lambda$ is given by

$$\lambda = \left(AQ^{-1}A'\right)^{-1}AQ^{-1}c. \tag{3.8}$$

The quadratic program (3.6) contains as a special case the program

$$\begin{aligned} &\text{minimize} && \tfrac{1}{2}x'Qx + c'x \\ &\text{subject to} && Ax = b. \end{aligned} \tag{3.9}$$

This special case is obtained when $\overline{x}$ is given by

$$\overline{x} = Q^{-1}A'(AQ^{-1}A')^{-1}b. \tag{3.10}$$

Indeed $\overline{x}$ as given above satisfies $A\overline{x} = b$ as required, and for all $x$ with $Ax = b$, we have

$$x'Q\overline{x} = x'A'(AQ^{-1}A')^{-1}b = b'(AQ^{-1}A')^{-1}b,$$

which implies that for all $x$ with $Ax = b$,

$$\tfrac{1}{2}(x-\overline{x})'Q(x-\overline{x}) + c'(x-\overline{x}) = \tfrac{1}{2}x'Qx + c'x + \left(\tfrac{1}{2}\overline{x}'Q\overline{x} - c'\overline{x} - b'(AQ^{-1}A')^{-1}b\right).$$

The last term in parentheses on the right-hand side above is constant, thus establishing that the programs (3.6) and (3.9) have the same optimal solution when $\overline{x}$ is given by Eq. (3.10). By combining Eqs. (3.7) and (3.10), we obtain the optimal solution of program (3.9):

$$x^* = -Q^{-1}\left(c - A'\lambda - A'(AQ^{-1}A')^{-1}b\right),$$

where $\lambda$ is given by Eq. (3.8).

### 3.3 (Approximate Minima of Convex Functions)

Let $X$ be a closed convex subset of $\Re^n$, and let $f : \Re^n \mapsto (-\infty, \infty]$ be a closed convex function such that $X \cap \text{dom}(f) \neq \varnothing$. Assume that $f$ and $X$ have no common nonzero direction of recession. Let $X^*$ be the set of minima of $f$ over $X$ (which is nonempty and compact by Prop. 3.2.2), and let $f^* = \inf_{x\in X} f(x)$. Show that:

(a) For every $\epsilon > 0$ there exists a $\delta > 0$ such that every vector $x \in X$ with $f(x) \leq f^* + \delta$ satisfies $\min_{x^* \in X^*} \|x - x^*\| \leq \epsilon$.

(b) If $f$ is real-valued, for every $\delta > 0$ there exists an $\epsilon > 0$ such that every vector $x \in X$ with $\min_{x^* \in X^*} \|x - x^*\| \leq \epsilon$ satisfies $f(x) \leq f^* + \delta$.

(c) Every sequence $\{x_k\} \subset X$ satisfying $f(x_k) \to f^*$ is bounded and all its limit points belong to $X^*$.

**Solution:** Let $X$ be a closed convex subset of $\Re^n$, and let $f : \Re^n \mapsto (-\infty, \infty]$ be a closed convex function such that $X \cap \text{dom}(f) \neq \emptyset$. Assume that $f$ and $X$ have no common nonzero direction of recession. Let $X^*$ be the set of minima of $f$ over $X$ (which is nonempty and compact by Prop. 3.2.2), and let $f^* = \inf_{x \in X} f(x)$. Show that:

(a) For every $\epsilon > 0$ there exists a $\delta > 0$ such that every vector $x \in X$ with $f(x) \leq f^* + \delta$ satisfies $\min_{x^* \in X^*} \|x - x^*\| \leq \epsilon$.

(b) If $f$ is real-valued, for every $\delta > 0$ there exists an $\epsilon > 0$ such that every vector $x \in X$ with $\min_{x^* \in X^*} \|x - x^*\| \leq \epsilon$ satisfies $f(x) \leq f^* + \delta$.

(c) Every sequence $\{x_k\} \subset X$ satisfying $f(x_k) \to f^*$ is bounded and all its limit points belong to $X^*$.

### 3.4 (Partial Minimization)

(a) Let $f : \Re^n \mapsto [-\infty, \infty]$ be a function and consider the subset of $\Re^{n+1}$ given by

$$E_f = \big\{(x, w) \mid f(x) < w\big\}.$$

Show that $E_f$ is related to the epigraph of $f$ as follows:

$$E_f \subset \text{epi}(f) \subset \text{cl}(E_f).$$

Show also that $f$ is convex if and only if $E_f$ is convex.

(b) Let $F : \Re^{m+n} \mapsto [-\infty, \infty]$ be a function and let

$$f(x) = \inf_{z \in \Re^m} F(x, z), \qquad x \in \Re^n.$$

Show that $E_f$ is the projection of the set $\big\{(x, z, w) \mid F(x, z) < w\big\}$ on the space of $(x, w)$.

(c) Use parts (a) and (b) to show that convexity of $F$ implies convexity of $f$ [this shows Prop. 3.3.1(a) by a somewhat different argument].

**Solution:** (a) We clearly have $E_f \subset \text{epi}(f)$. Let $(x, w) \in \text{epi}(f)$, so that $f(x) \leq w$. Let $\{w_k\}$ be a decreasing sequence such that $w_k \to w$, so that $f(x) < w_k$ and $(x, w_k) \in E_f$ for all $k$. Since $(x, w_k) \to (x, w)$, it follows that $(x, w) \in \text{cl}(E_f)$. Hence, $\text{epi}(f) \subset \text{cl}(E_f)$.

Let $f$ be convex, and let $(x, w), (y, v) \in E_f$. Then for any $\alpha \in [0, 1]$, we have

$$f\big(\alpha x + (1 - \alpha) y\big) \leq \alpha f(x) + (1 - \alpha) f(y) < \alpha w + (1 - \alpha) v.$$

It follows that $\big(\alpha x + (1-\alpha)y, \alpha w + (1-\alpha)v\big) \in E_f$, so that $E_f$ is convex.

Let $E_f$ be convex, and for any $x, y \in \text{dom}(f)$, let $\big\{(x, w_k)\big\}$, $\big\{(y, v_k)\big\}$ be sequences in $E_f$ such that $w_k \downarrow f(x)$ and $v_k \downarrow f(y)$, respectively [we allow the possibility that $f(x)$ and/or $f(y)$ are equal to $-\infty$]. Since $E_f$ is convex, we have

$$f\big(\alpha x + (1-\alpha)y\big) < \alpha w_k + (1-\alpha)v_k, \qquad \forall\ \alpha \in [0,1],\ k = 0, 1, \ldots.$$

Taking the limit as $k \to \infty$, we obtain

$$f\big(\alpha x + (1-\alpha)y\big) \leq \alpha f(x) + (1-\alpha)f(y), \qquad \forall\ \alpha \in [0,1],$$

so $f$ is convex.

(b) We have $f(x) < w$ if and only if there exists $z \in \Re^m$ such that $F(x,z) < w$:

$$E_f = \big\{(x,w) \mid F(x,z) < w \text{ for some } z \in \Re^m\big\}.$$

Thus $E_f$ is the projection of the set $\big\{(x,z,w) \mid F(x,z) < w\big\}$ on the space of $(x,w)$.

(c) If $F$ is convex, by part (a) the set $\big\{(x,z,w) \mid F(x,z) < w\big\}$ is convex, and by part (b), $E_f$ is the image of this set under a linear transformation. Therefore, $E_f$ is convex, so by part (a), $f$ is convex.

### 3.5 (Closures of Partially Minimized Functions)

Consider a function $F : \Re^{n+m} \mapsto (-\infty, \infty]$ and the function $f : \Re^n \mapsto [-\infty, \infty]$ defined by

$$f(x) = \inf_{z \in \Re^m} F(x,z).$$

(a) Show that

$$P\Big(\text{cl}\big(\text{epi}(F)\big)\Big) \subset \text{cl}\big(\text{epi}(f)\big),$$

where $P(\cdot)$ denotes projection on the space of $(x,w)$.

(b) Let $\overline{f}$ be defined by

$$\overline{f}(x) = \inf_{z \in \Re^m} (\text{cl}\, F)(x,z),$$

where $\text{cl}\, F$ is the closure of $F$. Show that the closures of $f$ and $\overline{f}$ coincide.

**Solution:** (a) We first note that from Prop. 3.3.1, we have

$$P\big(\text{epi}(F)\big) \subset \text{epi}(f) \subset \text{cl}\Big(P\big(\text{epi}(F)\big)\Big), \tag{3.11}$$

where $P(\cdot)$ denotes projection on the space of $(x,w)$. To show the equation

$$P\Big(\text{cl}\big(\text{epi}(F)\big)\Big) \subset \text{cl}\big(\text{epi}(f)\big), \tag{3.12}$$

let $(\overline{x}, \overline{w})$ belong to $P\big(\mathrm{cl}\big(\mathrm{epi}(F)\big)\big)$. Then there exists $\overline{z}$ such that $(\overline{x}, \overline{z}, \overline{w}) \in \mathrm{cl}\big(\mathrm{epi}(F)\big)$, and hence there is a sequence $(x_k, z_k, w_k) \in \mathrm{epi}(F)$ such that $x_k \to \overline{x}$, $z_k \to \overline{z}$, and $w_k \to \overline{w}$. Thus we have $f(x_k) \leq F(x_k, z_k) \leq w_k$, implying that $(x_k, w_k) \in \mathrm{epi}(f)$ for all $k$. It follows that $(\overline{x}, \overline{w}) \in \mathrm{cl}\big(\mathrm{epi}(f)\big)$.

(b) By taking closure in Eq. (3.11), we see that

$$\mathrm{cl}\big(\mathrm{epi}(f)\big) = \mathrm{cl}\Big(P\big(\mathrm{epi}(F)\big)\Big). \tag{3.13}$$

Denoting $\overline{F} = \mathrm{cl}\, F$ and replacing $F$ with $\overline{F}$, we also have

$$\mathrm{cl}\big(\mathrm{epi}(\overline{f})\big) = \mathrm{cl}\Big(P\big(\mathrm{epi}(\overline{F})\big)\Big). \tag{3.14}$$

On the other hand, by taking closure in Eq. (3.12), we have

$$\mathrm{cl}\Big(P\big(\mathrm{epi}(\overline{F})\big)\Big) \subset \mathrm{cl}\Big(P\big(\mathrm{epi}(F)\big)\Big),$$

which, in view of $\mathrm{epi}(\overline{F}) \supset \mathrm{epi}(F)$, implies that

$$\mathrm{cl}\Big(P\big(\mathrm{epi}(\overline{F})\big)\Big) = \mathrm{cl}\Big(P\big(\mathrm{epi}(F)\big)\Big). \tag{3.15}$$

By combining Eqs. (3.13)-(3.15), we see that

$$\mathrm{cl}\big(\mathrm{epi}(f)\big) = \mathrm{cl}\big(\mathrm{epi}(\overline{f})\big).$$

### 3.6 (Counterexample for Partially Minimized Functions)

Consider the function of $x \in \Re$ and $z = (z_1, z_2) \in \Re^2$ given by

$$F(x, z) = \begin{cases} e^{z_2} & \text{if } \|z\| \leq z_1 + x,\ x \geq 0, \\ \infty & \text{otherwise,} \end{cases}$$

which can be verified to be convex and closed. Show that the function $f(x) = \inf_{z \in \Re^2} F(x, z)$ is not closed.

**Solution:** We claim that $f$ is the nonclosed function

$$f(x) = \inf_{z \in \Re^2} F(x, z) = \begin{cases} 0 & \text{if } x > 0, \\ 1 & \text{if } x = 0, \\ \infty & \text{if } x < 0. \end{cases}$$

To see this note that when $x = 0$, the only $z$ that satisfies the constraint $\|z\| \leq z_1 + x$ is $z = 0$, but when $x > 0$, there exist $z = (z_1, z_2)$ with arbitrarily small $z_2$ such that $\|z\| \leq z_1 + x$.