Chapter 14 Unconstrained and Constrained Optimization Problems

Shuguang Cui[‡], Man-Cho Anthony So[♯], and Rui Zhang[♮]
[‡]Texas A&M University, College Station, USA
[♯]Chinese University of Hong Kong, Hong Kong, China
[♮]National University of Singapore, Singapore

In the first section of this chapter, we will give an overview of the basic mathematical tools that are useful for analyzing both unconstrained and constrained optimization problems. In order to allow the readers to focus on the applications of these tools and not to be burdened with too many technical details, we shall state most of the results without proof. However, the readers are strongly encouraged to refer to the texts [1–4] for expositions of these results and other further developments. In the second section, we provide three application examples to illustrate how we could apply the optimization techniques to solve real-world problems, with a focus on communications, networking, and signal processing. In the last section, several exercise questions are given to help the audience gain a deeper understanding of the material.

14.1 Basics of Convex Analysis

The notion of convexity plays a very important role in both the theoretical and algorithmic aspects of optimization. Before we discuss the relevance of convexity in optimization, let us first introduce the notions of convex sets and convex functions and state some of their properties.

Definition 14.1.1. Let $S \subset \mathbb{R}^n$ be a set. We say that

1. S is **affine** if $\alpha\mathbf{x} + (1 - \alpha)\mathbf{y} \in S$ whenever $\mathbf{x}, \mathbf{y} \in S$ and $\alpha \in \mathbb{R}$;

2. S is **convex** if $\alpha\mathbf{x} + (1 - \alpha)\mathbf{y} \in S$ whenever $\mathbf{x}, \mathbf{y} \in S$ and $\alpha \in [0, 1]$.

Given $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$, the vector $\mathbf{z} = \alpha\mathbf{x} + (1-\alpha)\mathbf{y}$ is called an **affine combination** of $\mathbf{x}$ and $\mathbf{y}$. If $\alpha \in [0,1]$, then $\mathbf{z}$ is called a **convex combination** of $\mathbf{x}$ and $\mathbf{y}$.

Geometrically, when $\mathbf{x}$ and $\mathbf{y}$ are distinct points in $\mathbb{R}^n$, the set

$$L = \{\mathbf{z} \in \mathbb{R}^n : \mathbf{z} = \alpha\mathbf{x} + (1-\alpha)\mathbf{y}, \, \alpha \in \mathbb{R}\}$$

of all affine combinations of $\mathbf{x}$ and $\mathbf{y}$ is simply the *line* determined by $\mathbf{x}$ and $\mathbf{y}$; and the set

$$S = \{\mathbf{z} \in \mathbb{R}^n : \mathbf{z} = \alpha\mathbf{x} + (1-\alpha)\mathbf{y}, \, \alpha \in [0,1]\}$$

is the *line segment* between $\mathbf{x}$ and $\mathbf{y}$. By convention, the empty set $\emptyset$ *is* convex.

It is clear that one can generalize the notion of affine (resp. convex) combination of two points to any finite number of points. In particular, an affine combination of the points $\mathbf{x_1}, \ldots, \mathbf{x_k} \in \mathbb{R}^n$ is a point $\mathbf{z} = \sum_{i=1}^k \alpha_i \mathbf{x_i}$, where $\sum_{i=1}^k \alpha_i = 1$. Similarly, a convex combination of the points $\mathbf{x_1}, \ldots, \mathbf{x_k} \in \mathbb{R}^n$ is a point $\mathbf{z} = \sum_{i=1}^k \alpha_i \mathbf{x_i}$, where $\sum_{i=1}^k \alpha_i = 1$ *and* $\alpha_1, \ldots, \alpha_k \geq 0$.

Here are some sets in Euclidean space whose convexity can be easily established by first principles:

Example 14.1.1. (Some Examples of Convex Sets)

1. **Non-Negative Orthant**: $\mathbb{R}_+^n = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \geq \mathbf{0}\}$.

2. **Hyperplane**: $H(\mathbf{s}, c) = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{s}^T\mathbf{x} = c\}$.

3. **Halfspaces**: $H^+(\mathbf{s}, c) = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{s}^T\mathbf{x} \leq c\}$, $H^-(\mathbf{s}, c) = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{s}^T\mathbf{x} \geq c\}$.

4. **Euclidean Ball**: $B(\bar{\mathbf{x}}, r) = \{\mathbf{x} \in \mathbb{R}^n : l|\mathbf{x} - \bar{\mathbf{x}}|l_2 \leq \gamma\}$.

5. **Ellipsoid**: $E(\bar{\mathbf{x}}, \mathbf{Q}, r) = \{\mathbf{x} \in \mathbb{R}^n : (\mathbf{x} - \bar{\mathbf{x}})^T Q(\mathbf{x} - \bar{\mathbf{x}}) \leq r^2\}$, where $\mathbf{Q}$ is an $n \times n$ symmetric, positive definite matrix (i.e., $\mathbf{x}^T\mathbf{Q}\mathbf{x} > 0$ for all $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$), and is denoted by $\mathbf{Q} \succ \mathbf{0}$.

6. **Simplex**: $\Delta = \left\{\sum_{pi=1}^n d : x_1 \sum_{i=0}^n \alpha := 1, \alpha :\geq 0 \text{for} i = 0, 1, \ldots, n\right\}$, where $\mathbf{x_0}, \mathbf{x_1}, \ldots, \mathbf{x_n}$ are vectors in $\mathbb{R}^n$ such that the vectors $\mathbf{x_1} - \mathbf{x_0}, \mathbf{x_2} - \mathbf{x_0}, \ldots, \mathbf{x_n} - \mathbf{x_0}$ are linearly independent (equivalently, the vectors $\mathbf{x_0}, \mathbf{x_1}, \ldots, \mathbf{x_n}$ are *affinely* independent).

7. **Positive Semidefinite Cone**:
   $\mathcal{S}_+^n = \{\mathbf{A} \in \mathbb{R}^{n \times n} : \mathbf{A} \text{ is symmetric and } \mathbf{x}^T\mathbf{A}\mathbf{x} \geq 0 \text{ for all } \mathbf{x} \in \mathbb{R}^n\}$ (a symmetric matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ is said to be positive semidefinite if $\mathbf{x}^T\mathbf{A}\mathbf{x} \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n$, and is denoted by $\mathbf{A} \succeq \mathbf{0}$).

$\square$

Let us now turn to the notion of a convex function.

482CHAPTER 14. UNCONSTRAINED & CONSTRAINED OPTIMIZATION PROBLEMS

Definition 14.1.2. Let $S \subset \mathbb{R}^n$ be a nonempty convex set, and let $f : S \to \mathbb{R}$ be a real-valued function.

1. We say that f is **convex** on S if

$$f(\alpha \mathbf{x_1} + (1 - \alpha)\mathbf{x_2}) \leq \alpha f(\mathbf{x_1}) + (1 - \alpha)f(\mathbf{x_2}) \qquad (14.1)$$

   for all $\mathbf{x_1}, \mathbf{x_2} \in S$ and $\alpha \in [0, 1]$. We say that f is **concave** if $-f$ is convex.

2. We say that f is **strictly convex** on S if

$$f(\alpha \mathbf{x_1} + (1 - \alpha)\mathbf{x_2}) < \alpha f(\mathbf{x_1}) + (1 - \alpha)f(\mathbf{x_2})$$

   for all $\mathbf{x_1}, \mathbf{x_2} \in S$ and $\alpha \in (0, 1)$.

3. The **epigraph** of f is the set $\mathrm{epi}(f) = \{(\mathbf{x}, r) \in S \times \mathbb{R} : f(\mathbf{x}) \leq r\}$.

The relationship between convex sets and convex functions can be summarized as follows:

Proposition 14.1.1. Let f be as in Definition 14.1.2. Then, f is convex (as a function) iff $\mathrm{epi}(f)$ is convex (as a set in $S \times \mathbb{R}$).

Let $r \in \mathbb{R}$ be arbitrary. A set closely related to the epigraph is the so-called r–**level set** of f, which is defined as $L(r) = \{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) \leq r\}$. It is clear that if f is convex, then $L(r)$ is convex for all $r \in \mathbb{R}$. However, the converse is *not* true, as illustrated by the function $x \mapsto x^3$. A function $f : S \to \mathbb{R}$ whose domain is convex and whose r–level sets are convex for all $r \in \mathbb{R}$ is called **quasi-convex**.

One of the most desirable features of convexity is the following:

Proposition 14.1.2. Consider the optimization problem:

$$\begin{array}{ll} \text{minimize} & f(\mathbf{x}) \\ \text{subject to} & \mathbf{x} \in S, \end{array}$$

where $S \subset \mathbb{R}^n$ is a convex set and $f : S \to \mathbb{R}$ is convex. Then, any local minimum of f is also a global minimum[1].

Now, let $S \subset \mathbb{R}^n$ be an open convex set, and let $f : S \to \mathbb{R}$ be an arbitrary function. When f has suitable degree of differentiability, we can characterize its convexity by its gradient or Hessian. Specifically, we have the following.

Theorem 14.1.1. *Let $S \subset \mathbb{R}^n$ be an open convex set, and let $f : S \to \mathbb{R}$ be a differentiable function on S. Then, f is convex on S iff*

$$f(\mathbf{x_1}) \geq f(\mathbf{x_2}) + (\nabla f(\mathbf{x_2}))^T (\mathbf{x_1} - \mathbf{x_2})$$

for all $\mathbf{x_1}, \mathbf{x_2} \in S$. Furthermore, if f is twice continuously differentiable function on S, then f is convex on S iff $\nabla^2 f(\mathbf{x})$ is positive semidefinite for all $\mathbf{x} \in S$.

---

[1] Recall that for a generic optimization problem $\min_{\mathbf{x} \in S \subset \mathbb{R}^n} f(\mathbf{x})$, a point $\mathbf{x}^* \in S$ is called a **global minimum** if $f(\mathbf{x}^*) \leq f(\mathbf{x})$ for all $\mathbf{x} \in S$. On the other hand, if there exists an $\epsilon > 0$ such that the point $\mathbf{x}^* \in S$ satisfies $f(\mathbf{x}^*) \leq f(\mathbf{x})$ for all $\mathbf{x} \in S \cap B^\circ(\mathbf{x}^*, \epsilon)$, then it is called a **local minimum**. Here, $B^\circ(\bar{\mathbf{x}}, \epsilon) =$ denotes the *open* ball centered at $\bar{\mathbf{x}} \in \mathbb{R}^n$ of radius $\epsilon > 0$.

Sometimes it may be difficult to verify directly from the definition whether a given function is convex or not. However, a function can often be obtained as a composition of several, more elementary functions. When each of those elementary functions is convex, it is natural to ask whether their composition is also convex. In general, the answer is no. On the other hand, here are some transformations that preserve convexity.

Theorem 14.1.2. *Let $S \subset \mathbb{R}^n$ be a nonempty convex set. Then, the following hold:*

1. **(Non-negative Combinations)** *Let $f_1, \ldots, f_m : S \to \mathbb{R}$ be convex functions, and let $\alpha_1, \ldots, \alpha_m \geq 0$. Then, the function $\sum_{i=1}^{m} \alpha_i f_i$ is convex on S.*

2. **(Pointwise Supremum)** *Let $\{f_i\}_{i \in I}$ be an arbitrary family of convex functions on S. Then, the pointwise supremum $f = \sup_{i \in I} f_i$ is convex on S.*

3. **(Affine Composition)** *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function and $A : \mathbb{R}^m \to \mathbb{R}^n$ be an affine mapping[2]. Then, the function $f \circ A : \mathbb{R}^m \to \mathbb{R}^n$ given by $(f \circ A)(\mathbf{x}) = f(A(\mathbf{x}))$ is convex on $\mathbb{R}^m$.*

4. **(Composition with an Increasing Convex Function)** *Let $f : S \to \mathbb{R}$ be a convex function, and let $g : \mathbb{R} \to \mathbb{R}$ be an increasing convex function. Then, the function $g \circ f : S \to \mathbb{R}$ defined by $(g \circ f)(\mathbf{x}) = g(f(\mathbf{x}))$ is convex on S.*

5. **(Restriction on Lines)** *Let $f : S \to \mathbb{R}$ be a function. Given $\mathbf{x_0} \in S$ and $\mathbf{h} \in \mathbb{R}^n$, define the function $\tilde{f}_{\mathbf{x_0},\mathbf{h}} : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ by*

$$\tilde{f}_{\mathbf{x_0},\mathbf{h}}(t) = \begin{cases} f(\mathbf{x_0} + t\mathbf{h}) & \text{if } \mathbf{x_0} + t\mathbf{h} \in S, \\ +\infty & \text{otherwise.} \end{cases}$$

Then, f is convex on S iff $\tilde{f}_{\mathbf{x_0},\mathbf{h}}$ is convex on $\mathbb{R}$ for any $\mathbf{x_0} \in S$ and $\mathbf{h} \in \mathbb{R}^n$.

Let us now illustrate an application of Theorem 14.1.2.

Example 14.1.2. Let $f : \mathbb{R}^{m \times n} \to \mathbb{R}_+$ be given by $f(\mathbf{X}) = \|\mathbf{X}\|_2$, where $\|\cdot\|_2$ denotes the **spectral norm** or **largest singular value** of the $m \times n$ matrix $\mathbf{X}$. By the Courant–Fischer theorem (see, e.g., [5]), we have

$$f(\mathbf{X}) = \sup. \tag{14.2}$$

Now, for each $\mathbf{u} \in \mathbb{R}^m$ and $\mathbf{v} \in \mathbb{R}^n$ with $\|\mathbf{u}\|_2 = \|\mathbf{v}\|_2 = 1$, define the function $f_{\mathbf{u},\mathbf{v}} : \mathbb{R}^{m \times n} \to \mathbb{R}$ by

$$f_{\mathbf{u},\mathbf{v}}(\mathbf{X}) = \mathbf{u}^T \mathbf{X} \mathbf{v}.$$

---

[2]A map $A : \mathbb{R}^m \to \mathbb{R}^n$ is said to be **affine** if there exists an $n \times m$ matrix $\mathbf{B}$ and a vector $\mathbf{d} \in \mathbb{R}^n$ such that $A(\mathbf{x}) = \mathbf{B}\mathbf{x} + \mathbf{d}$ for all $\mathbf{x} \in \mathbb{R}^m$.

Note that $f_{\mathbf{u},\mathbf{v}}$ is a convex (in fact, linear) function of $\mathbf{X}$ for each $\mathbf{u}, \mathbf{v}$. Hence, it follows from (14.2) that f is a pointwise supremum of a family of linear functions of $\mathbf{X}$. By Theorem 14.1.2, this implies that f is convex. $\qquad\square$

14.2   Unconstrained vs.   Constrained Optimization

14.2.1   Optimality Conditions for Unconstrained Optimization

One of the most fundamental problems in optimization is to derive conditions for identifying potential optimal solutions to an optimization problem. Typically, such conditions, which are known as **optimality conditions**, would enable us to reduce the original optimization problem to that of checking the validity of certain geometric conditions, or to that of checking the consistency of certain system of inequalities. As an illustration and to motivate our discussion, let us first consider a univariate, twice continuously differentiable function $f : \mathbb{R} \to \mathbb{R}$. Recall from basic calculus that if $\bar{x} \in \mathbb{R}$ is a local minimum of f, then we must have

$$\frac{df(x)}{dx}\Big|_{x=\bar{x}} = 0. \tag{14.3}$$

In other words, condition (14.3) is a *necessary condition* for $\bar{x}$ to be a local minimum. However, it is *not* a sufficient condition, as an $\bar{x} \in \mathbb{R}$ that satisfies (14.3) can be a local maximum or just a stationary point. In order to certify that $\bar{x}$ is indeed a local minimum, one could check, in addition to (14.3), whether

$$\frac{d^2 f(x)}{dx^2}\Big|_{x=\bar{x}} > 0. \tag{14.4}$$

In particular, condition (14.4) is a *sufficient condition* for $\bar{x}$ to be a local minimum.

In the above discussion, conditions (14.3) and (14.4) together yield a system of inequalities whose solutions are local minima of the function f. Alternatively, they can be viewed as stating the geometric fact that there is no descent direction in a neighborhood of a local minimum. In particular, the former is an algebraic interpretation of local optimality, while the latter is a geometric interpretation. It is worth noting that each interpretation has its own advantage. Indeed, the geometric interpretation can often help us gain intuitions about the problem at hand, and the algebraic interpretation would help to make those intuitions precise. Thus, it is good to keep both interpretations in mind.

To derive optimality conditions for the local minima of a *multivariate* twice continuously differentiable function $f : \mathbb{R}^n \to \mathbb{R}$, we first recall that $\nabla f(\mathbf{x})$, the gradient of f at $\mathbf{x} \in \mathbb{R}^n$, is the direction of steepest ascent at $\mathbf{x}$. Thus, if $\nabla f(\mathbf{x}) \neq \mathbf{0}$, then starting at $\mathbf{x}$, we can proceed in the direction $-\nabla f(\mathbf{x})$ and achieve a smaller function value. More specifically, we have the following

Proposition 14.2.1. Suppose that $f : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable at $\bar{\mathbf{x}} \in \mathbb{R}^n$. If there exists a $\mathbf{d} \in \mathbb{R}^n$ such that $\nabla \left(f(\bar{\mathbf{x}}) \right)^T \mathbf{d} < 0$, then there exists an $\alpha_0 > 0$ such that $f(\bar{\mathbf{x}} + \alpha \mathbf{d}) < f(\bar{\mathbf{x}})$ for all $\alpha \in (0, \alpha_0)$. In other words, $\mathbf{d}$ is a **descent direction** of f at $\bar{\mathbf{x}}$.

Using Proposition 14.2.1, we can establish the following.

Corollary 14.2.1. (First Order Necessary Condition for Unconstrained Optimization) Suppose that $f : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable at $\bar{\mathbf{x}} \in \mathbb{R}^n$. If $\bar{\mathbf{x}}$ is a local minimum, then we have $\nabla f(\bar{\mathbf{x}}) = \mathbf{0}$. In particular, we have $= \emptyset$.

Similar to the univariate case, even if $\bar{\mathbf{x}} \in \mathbb{R}^n$ satisfies $\nabla f(\bar{\mathbf{x}}) = \mathbf{0}$, we cannot conclude that $\bar{\mathbf{x}}$ is a local minimum. For instance, consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ given by $f(x_1, x_2) = -x_1^2 - (x_1 - x_2)^2$. Then, we have

$$\nabla f(\mathbf{x}) = -2(2x_1 - x_2, x_2 - x_1).$$

In particular, the (unique) solution to $\nabla f(\mathbf{x}) = \mathbf{0}$ is $\bar{x}_1 = \bar{x}_2 = 0$. However, as can be easily verified, the point $(\bar{x}_1, \bar{x}_2) = (0, 0)$ is a global maximum of f.

The above example shows that some extra conditions are needed in order to guarantee that a solution to the equation $\nabla f(\mathbf{x}) = \mathbf{0}$ is a local minimum of f. For instance, we have the following proposition, which states that if f is convex at $\bar{\mathbf{x}}$, then the necessary condition in Corollary 14.2.1 is also sufficient[3]:

Proposition 14.2.2. Suppose that $f : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable and convex at $\bar{\mathbf{x}}$. Then, $\bar{\mathbf{x}}$ is a global minimum iff $\nabla f(\bar{\mathbf{x}}) = \mathbf{0}$.

Alternatively, if $\nabla f(\bar{\mathbf{x}}) = \mathbf{0}$ and $\nabla^2 f(\bar{\mathbf{x}})$, the Hessian of f at $\bar{\mathbf{x}}$, is positive definite, then $\bar{\mathbf{x}}$ is a local minimum. Specifically, we have the following proposition, which generalizes the corresponding result for the univariate case (cf. (14.3) and (14.4)).

Proposition 14.2.3. (Second Order Sufficient Condition for Unconstrained Optimization) Suppose that $f : \mathbb{R}^n \to \mathbb{R}$ is twice continuously differentiable at $\bar{\mathbf{x}} \in \mathbb{R}^n$. If $\nabla f(\bar{\mathbf{x}}) = \mathbf{0}$ and $\nabla^2 f(\bar{\mathbf{x}})$ is positive definite, then $\bar{\mathbf{x}}$ is a local minimum.

Let us now illustrate the above results with an example.

Example 14.2.1. Let $f : \mathbb{R}^n \to \mathbb{R}$ be defined by $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T \mathbf{Q} \mathbf{x} + \mathbf{c}^T \mathbf{x}$, where $\mathbf{Q} \in \mathbf{S}^n$ and $\mathbf{c} \in \mathbb{R}^n$ are given. Then, f is continuously differentiable, and we have $\nabla f(\mathbf{x}) = \mathbf{Q}\mathbf{x} + \mathbf{c}$ and $\nabla^2 f(\mathbf{x}) = \mathbf{Q}$. Now, if f is convex, or equivalently, if $\mathbf{Q} \succeq \mathbf{0}$, then by Proposition 14.2.2, any $\bar{\mathbf{x}} \in \mathbb{R}^n$ that satisfies $\mathbf{Q}\bar{\mathbf{x}} + \mathbf{c} = \mathbf{0}$ will be a global minimum of f. Note that in this case, we cannot even conclude

---

[3]Let S be a nonempty convex subset of $\mathbb{R}^n$. We say that $f : S \to \mathbb{R}$ is **convex at** $\bar{\mathbf{x}} \in S$ if $f(\alpha\bar{\mathbf{x}} + (1 - \alpha)\mathbf{x}) \leq \alpha f(\bar{\mathbf{x}}) + (1 - \alpha)f(\mathbf{x})$ for all $\alpha \in (0, 1)$ and $\mathbf{x} \in S$. Note that a function $f : S \to \mathbb{R}$ can be convex at a particular point $\bar{\mathbf{x}} \in S$ without being convex on S.

from Proposition 14.2.3 that $\bar{\mathbf{x}}$ is a local minimum of f, since we only have $\mathbf{Q} \succeq \mathbf{0}$. On the other hand, suppose that $\mathbf{Q} \succ \mathbf{0}$. Then, $\mathbf{Q}$ is invertible, and by Proposition 14.2.3, the point $\bar{\mathbf{x}} = -\mathbf{Q}^{-1}\mathbf{c}$ is a local minimum of f. However, since f is convex, Proposition 14.2.2 allows us to draw a stronger conclusion, namely, the point $\bar{\mathbf{x}} = -\mathbf{Q}^{-1}\mathbf{c}$ is in fact the unique global minimum. $\qquad\square$

14.2.2 Optimality Conditions for Constrained Optimization

After deriving optimality conditions for unconstrained optimization problems, let us turn our attention to *constrained* optimization problems of the form

$$\min_{\mathbf{x} \in S} f(\mathbf{x}), \tag{14.5}$$

where S is a nonempty subset of $\mathbb{R}^n$. Note that due to the constraint $\mathbf{x} \in S$, even if $\bar{\mathbf{x}} \in \mathbb{R}^n$ satisfies $\nabla f(\bar{\mathbf{x}}) = \mathbf{0}$ and $\nabla^2 f(\bar{\mathbf{x}}) \succ \mathbf{0}$, it may not be a solution to (14.5), since $\bar{\mathbf{x}}$ need not lie in S. Similarly, a local minimum $\bar{\mathbf{x}}$ of f over S need not satisfy $\nabla f(\bar{\mathbf{x}}) = \mathbf{0}$, since all the descent directions of f at $\bar{\mathbf{x}}$ may lead to points that do not lie in S. Thus, in order to derive optimality conditions for (14.5), we need to consider not only the **set of descent directions** at $\bar{\mathbf{x}}$, i.e.,

$$D = \left\{ \mathbf{d} \in \mathbb{R}^n : \nabla f(\bar{\mathbf{x}})^T \mathbf{d} < 0 \right\}, \tag{14.6}$$

but also the **set of feasible directions** at $\bar{\mathbf{x}}$, i.e.,

$$F = \{\mathbf{d} \in \mathbb{R}^n \setminus \{\mathbf{0}\} : \text{there exists an } \alpha_0 > 0 \text{ such that } \bar{\mathbf{x}} + \alpha\mathbf{d} \in S \text{ for all } \alpha \in (0, \alpha_0)\}. \tag{14.7}$$

We emphasize that in order for $\mathbf{d} \in F$, the *entire* open line segment $\{\bar{\mathbf{x}} + \alpha\mathbf{d} : \alpha \in (0, \alpha_0)\}$ must belong to S. This is to ensure that whenever $\mathbf{d} \in D$, one can find a feasible solution $\bar{\mathbf{x}}' \in S$ with $f(\bar{\mathbf{x}}') < f(\bar{\mathbf{x}})$ by proceeding from $\bar{\mathbf{x}}$ in the direction $\mathbf{d}$. Indeed, by Proposition 14.2.1, if $\mathbf{d} \in D$, then there exists an $\alpha_1 > 0$ such that $f(\bar{\mathbf{x}} + \alpha\mathbf{d}) < f(\bar{\mathbf{x}})$ for all $\alpha \in (0, \alpha_1)$. However, if $\bar{\mathbf{x}} + \alpha\mathbf{d} \notin S$ for any $\alpha \in (0, \alpha_1)$, then we cannot rule out the local minimality of $\bar{\mathbf{x}}$, even if $\bar{\mathbf{x}} + \alpha\mathbf{d} \in S$ for some $\alpha > \alpha_1$.

As the following proposition shows, the sets D and F provide a necessary, and under some additional assumptions, even sufficient condition for optimality.

Proposition 14.2.4. Consider Problem (14.5). Suppose that $f : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable at $\bar{\mathbf{x}} \in S$. If $\bar{\mathbf{x}}$ is a local minimum, then we have $D \cap F = \emptyset$. Conversely, suppose that (i) $D \cap F = \emptyset$, (ii) f is convex at $\bar{\mathbf{x}}$, and (iii) there exists an $\epsilon > 0$ such that $\mathbf{d} = \mathbf{x} - \bar{\mathbf{x}} \in F$ for any $\mathbf{x} \in S \cap B^\circ(\bar{\mathbf{x}}, \epsilon)$. Then, $\bar{\mathbf{x}}$ is a local minimum of f over S.

REMARKS: Condition (iii) is to ensure that the entire line segment $\{\bar{\mathbf{x}} + \alpha(\mathbf{x} - \bar{\mathbf{x}}) : \alpha \in [0, 1]\}$ lies in S for any $\mathbf{x} \in S \cap B^\circ(\bar{\mathbf{x}}, \epsilon)$, so that $\mathbf{d} = \mathbf{x} - \bar{\mathbf{x}} \in F$; see the remark after (14.7).

So far we have only discussed optimality conditions for a very general class of optimization problems, i.e., problems of the form (14.5). In particular, we derived a necessary condition for local optimality in terms of the sets D and F, namely that $D \cap F = \emptyset$. However, such a condition is largely geometric, and it is not as easy to manipulate as algebraic conditions (e.g., a system of inequalities). On the other hand, as we will show below, if the feasible region has more structure, then one can circumvent such difficulty and derive algebraic optimality conditions. To begin, let us consider the following class of optimization problems:

$$
\begin{aligned}
\text{minimize} \quad & f(\mathbf{x}) \\
\text{subject to} \quad & g_i(\mathbf{x}) \le 0 \quad \text{for } i = 1, \ldots, m, \\
& \mathbf{x} \in X,
\end{aligned} \tag{14.8}
$$

where $f : \mathbb{R}^n \to \mathbb{R}$ and $g_i : \mathbb{R}^n \to \mathbb{R}$ are continuously differentiable functions, and X is a nonempty open subset of $\mathbb{R}^n$ (usually we take $X = \mathbb{R}^n$). We then have the following.

Proposition 14.2.5. Let $S = \{\mathbf{x} \in X : g_i(\mathbf{x}) \le 0 \text{ for } i = 1, \ldots, m\}$ be the feasible region of problem (14.8), and let $\bar{\mathbf{x}} \in S$. Define

$$
I = \{i \in \{1, \ldots, m\} : g_i(\bar{\mathbf{x}}) = 0\}
$$

to be the index set for the **active** or **binding** constraints. Furthermore, define

$$
\begin{aligned}
G &= \left\{ \mathbf{d} \in \mathbb{R}^n : \nabla g_i(\bar{\mathbf{x}})^T \mathbf{d} < 0 \text{ for } i \in I \right\}, \tag{14.9} \\
\overline{G} &= \left\{ \mathbf{d} \in \mathbb{R}^n \setminus \{\mathbf{0}\} : \nabla g_i(\bar{\mathbf{x}})^T \mathbf{d} \le 0 \text{ for } i \in I \right\}.
\end{aligned}
$$

Then, we have $G \subset F \subset \overline{G}$, where F is defined in (14.7). Moreover, if the functions g_i, where $i \in I$, are strictly convex (resp. concave) at $\bar{\mathbf{x}}$, then $F = G$ (resp. $F = \overline{G}$).

Using Proposition 14.2.4 and Proposition 14.2.5, we can establish the following geometric optimality condition for (14.8):

Corollary 14.2.2. Let S be the feasible region of problem (14.8). Let $\bar{\mathbf{x}} \in S$, and define $I = \{i \in \{1, \ldots, m\} : g_i(\bar{\mathbf{x}}) = 0\}$. If $\bar{\mathbf{x}}$ is a local minimum, then $D \cap G = \emptyset$, where D is defined in (14.6) and G is defined in (14.9).

The intuition behind Corollary 14.2.2 is quite straightforward. Indeed, suppose that $\mathbf{d} \in D \cap G$. Then, by Proposition 14.2.1, there exists an $\alpha_0 > 0$ such that $f(\bar{\mathbf{x}} + \alpha \mathbf{d}) < f(\bar{\mathbf{x}})$ and $g_i(\bar{\mathbf{x}} + \alpha \mathbf{d}) < g_i(\bar{\mathbf{x}}) = 0$ for all $i \in I$ and $\alpha \in (0, \alpha_0)$. Moreover, by the continuity of the functions $g_1, \ldots, g_m$, for sufficiently small $\alpha > 0$, we have $g_i(\bar{\mathbf{x}} + \alpha \mathbf{d}) < 0$ for all $i \notin I$. It follows that there exists an $\alpha_1 > 0$ such that $\bar{\mathbf{x}} + \alpha \mathbf{d} \in S$ and $f(\bar{\mathbf{x}} + \alpha \mathbf{d}) < f(\bar{\mathbf{x}})$ for all $\alpha \in (0, \alpha_1)$. In other words, $\bar{\mathbf{x}}$ is not a local minimum.

The upshot of Corollary 14.2.2 is that it allows us to derive optimality conditions for (14.8) that is more algebraic in nature. Specifically, Corollary 14.2.2, together with Farkas' lemma, yields the following.

488 *CHAPTER 14. UNCONSTRAINED & CONSTRAINED OPTIMIZATION PROBLEMS*

Theorem 14.2.1. (Karush–Kuhn–Tucker Necessary Conditions) *Let $\bar{\mathbf{x}} \in S$ be a local minimum of problem* (14.8)*, and let $I = \{i \in \{1,\ldots,m\} : g_i(\bar{\mathbf{x}}) = 0\}$ be the index set for the active constraints. Suppose that the family $\{\nabla g_i(\bar{\mathbf{x}})\}_{i \in I}$ of vectors is linearly independent. Then, there exist $\bar{u}_1,\ldots,\bar{u}_m \in \mathbb{R}$ such that*

$$
\begin{aligned}
\nabla f(\bar{\mathbf{x}}) + \sum_{i=1}^{m} \bar{u}_i \nabla g_i(\bar{\mathbf{x}}) &= \mathbf{0}, \\
\bar{u}_i g_i(\bar{\mathbf{x}}) &= 0 \qquad for\ i = 1,\ldots,m, \\
\bar{u}_i &\geq 0 \qquad for\ i = 1,\ldots,m.
\end{aligned}
\tag{14.10}
$$

We say that $\bar{\mathbf{x}} \in \mathbb{R}^n$ is a **KKT point** if (i) $\bar{\mathbf{x}} \in S$ and (ii) there exist **Lagrange multipliers** $\bar{u}_1,\ldots,\bar{u}_m$ such that $(\bar{\mathbf{x}},\bar{u}_1,\ldots,\bar{u}_m)$ satisfies the system (14.10).

Note that if the gradient vectors of the active constraints are not linearly independent, then the KKT conditions are *not* necessary for local optimality, *even when the optimization problem is convex*. This is demonstrated in the following example.

Example 14.2.2. Consider the following optimization problem:

$$
\begin{aligned}
\text{minimize} \quad & x_1 \\
\text{subject to} \quad & (x_1 - 1)^2 + (x_2 - 1)^2 \leq 1, \\
& (x_1 - 1)^2 + (x_2 + 1)^2 \leq 1.
\end{aligned}
\tag{14.11}
$$

Since there is only one feasible solution (i.e., $(x_1, x_2) = (1,0)$), it is naturally optimal. Besides the primal feasibility condition, the KKT conditions of (14.11) are given by

$$
\begin{aligned}
\begin{bmatrix} 1 \\ 0 \end{bmatrix} + 2u_1 \begin{bmatrix} x_1 - 1 \\ x_2 - 1 \end{bmatrix} + 2u_2 \begin{bmatrix} x_1 - 1 \\ x_2 + 1 \end{bmatrix} &= \mathbf{0}, \\
u_1 \left((x_1 - 1)^2 + (x_2 - 1)^2 - 1\right) &= 0, \\
u_2 \left((x_1 - 1)^2 + (x_2 + 1)^2 - 1\right) &= 0. \\
u_1, u_2 &\geq 0
\end{aligned}
$$

However, it is clear that there is no solution $(u_1, u_2) \geq \mathbf{0}$ to the above system when $(x_1, x_2) = (1,0)$. $\qquad\square$

Let us now illustrate Theorem 14.2.1 with an example.

Example 14.2.3. (Optimization of a Matrix Function) Let $\mathbf{A} \succ \mathbf{0}$ and $b > 0$ be given. Consider the following problem:

$$
\begin{aligned}
\text{minimize} \quad & -\log \det(\mathbf{Z}) \\
\text{subject to} \quad & \mathrm{tr}(\mathbf{A}\mathbf{Z}) \leq b, \\
& \mathbf{Z} \succ \mathbf{0}.
\end{aligned}
\tag{14.12}
$$

Note that (14.12) is of the form (14.8), since we may write (14.12) as

$$\begin{aligned}
\text{minimize} \quad & -\log\det(\mathbf{Z}) \\
\text{subject to} \quad & \operatorname{tr}(\mathbf{AZ}) \le b, \\
& \mathbf{Z} \in \mathbf{S}_{++}^n,
\end{aligned}$$

and $\mathbf{S}_{++}^n \subset \mathbb{R}^{n(n+1)/2}$ is an open set. Now, it is known that for $\mathbf{X} \succ \mathbf{0}$,

$$\nabla \log\det(\mathbf{X}) = \mathbf{X}^{-1}, \quad \nabla \operatorname{tr}(\mathbf{AX}) = \mathbf{A};$$

see, e.g., [6]. Hence, the KKT conditions associated with (14.12) are given by

$$\begin{aligned}
\operatorname{tr}(\mathbf{AZ}) &\le & b, \quad \mathbf{Z} &\succ& \mathbf{0}, & \quad (a) \\
-\mathbf{Z}^{-1} + u\mathbf{A} &=& \mathbf{0}, \quad u &\ge& 0, & \quad (b) \\
u(\operatorname{tr}(\mathbf{AZ}) - b) &=& 0. & & & \quad (c)
\end{aligned}$$

Condition (a) is simply primal feasibility. Condition (c) is known as **complementarity**. As we shall see later, condition (b) can be interpreted as feasibility with respect to a certain dual of (14.12). □

Note that Theorem 14.2.1 applies only to inequality-constrained optimization problems of the form (14.8). However, by extending the geometric arguments used to prove Corollary 14.2.2, one can establish similar necessary optimality conditions for optimization problems of the form

$$\begin{aligned}
\text{minimize} \quad & f(\mathbf{x}) \\
\text{subject to} \quad & g_i(\mathbf{x}) \le 0 \quad \text{for } i = 1, \ldots, m_1, \\
& h_j(\mathbf{x}) = 0 \quad \text{for } j = 1, \ldots, m_2, \\
& \mathbf{x} \in X,
\end{aligned} \tag{14.13}$$

where $f, g_1, \ldots, g_{m_1}, h_1, \ldots, h_{m_2} : \mathbb{R}^n \to \mathbb{R}$ are continuously differentiable functions, and X is a nonempty open subset of $\mathbb{R}^n$. Specifically, we have the following.

Theorem 14.2.2. (Karush–Kuhn–Tucker Necessary Conditions) *Let S be the feasible region of Problem* (14.13). *Suppose that $\bar{\mathbf{x}} \in S$ is a local minimum of problem* (14.13), *with $I = \{i \in \{1, \ldots, m_1\} : g_i(\bar{\mathbf{x}}) = 0\}$ being the index set for the active constraints. Furthermore, suppose that $\bar{\mathbf{x}}$ is* regular, *i.e., the family $\{\nabla g_i(\bar{\mathbf{x}})\}_{i \in I} \cup \{\nabla h_j(\bar{\mathbf{x}})\}_{j=1}^{m_2}$ of vectors is linearly independent. Then, there exist $\bar{v}_1, \ldots, \bar{v}_{m_1} \in \mathbb{R}$ and $\bar{w}_1, \ldots, \bar{w}_{m_2} \in \mathbb{R}$ such that*

$$\begin{aligned}
\nabla f(\bar{\mathbf{x}}) + \sum_{i=1}^{m_1} \bar{v}_i \nabla g_i(\bar{\mathbf{x}}) + \sum_{j=1}^{m_2} \bar{w}_j \nabla h_j(\bar{\mathbf{x}}) &= \mathbf{0}, \\
\bar{v}_i g_i(\bar{\mathbf{x}}) &= 0 \quad \text{for } i = 1, \ldots, m_1, \\
\bar{v}_i &\ge 0 \quad \text{for } i = 1, \ldots, m_1.
\end{aligned} \tag{14.14}$$

490 *CHAPTER 14. UNCONSTRAINED & CONSTRAINED OPTIMIZATION PROBLEMS*

As demonstrated in Exercise 14.2.2, the linear independence of the gradient vectors of the active constraints is generally needed to guarantee the existence of Lagrange multipliers. However, such a regularity condition is not always easy to check. As it turns out, there are other forms of regularity conditions, a more well-known of which is the following:

Theorem 14.2.3. *Suppose that in Problem* (14.13)*, the functions* $g_1, \ldots, g_{m_1}$ *are convex and* $h_1, \ldots, h_{m_2}$ *are linear. Let* $\bar{\mathbf{x}} \in S$ *be a local minimum, and let* $I = \{i \in \{1, \ldots, m_1\} : g_i(\bar{\mathbf{x}}) = 0\}$. *If the* Slater condition *is satisfied, i.e., if there exists an* $\mathbf{x}' \in S$ *such that* $g_i(\mathbf{x}') < 0$ *for all* $i \in I$*, then* $\bar{\mathbf{x}}$ *satisfies the KKT conditions* (14.14)*.*

Another setting in which the existence of Lagrange multipliers is guaranteed is the following:

Theorem 14.2.4. *Suppose that in Problem* (14.13)*, the functions* $g_1, \ldots, g_{m_1}$ *are concave and* $h_1, \ldots, h_{m_2}$ *are linear. Let* $\bar{\mathbf{x}} \in S$ *be a local minimum. Then,* $\bar{\mathbf{x}}$ *satisfies the KKT conditions* (14.14)*.*

In particular, Theorem 14.2.4 implies that when all the constraints in problem (14.13) are linear, one can always find Lagrange multipliers for any local minimum of problem (14.13).

So far we have only discussed necessary optimality conditions for constrained optimization problems. Let us now turn our attention to sufficient conditions. The following theorem can be viewed as an extension of the first order sufficient condition in Proposition 14.2.2 to the constrained setting.

Theorem 14.2.5. *Suppose that in Problem* (14.13)*, the functions* $f, g_1, \ldots, g_{m_1}$ *are convex,* $h_1, \ldots, h_{m_2}$ *are linear, and* $X = \mathbb{R}^n$*. Let* $\bar{\mathbf{x}} \in \mathbb{R}^n$ *be feasible for* (14.13)*. If there exist vectors* $\bar{\mathbf{v}} \in \mathbb{R}^{m_1}$ *and* $\bar{\mathbf{w}} \in \mathbb{R}^{m_2}$ *such that* $(\bar{\mathbf{x}}, \bar{\mathbf{v}}, \bar{\mathbf{w}})$ *satisfies the KKT conditions* (14.14)*, then* $\bar{\mathbf{x}}$ *is a global minimum.*

To demonstrate the usage of the above results, let us consider the following example:

Example 14.2.4. (Linear Programming) Consider the standard form linear programming (LP):

$$
\begin{aligned}
\text{minimize} \quad & f(\mathbf{x}) \equiv \mathbf{c}^T \mathbf{x} \\
\text{subject to} \quad & h_j(\mathbf{x}) \equiv \mathbf{a_j}^T \mathbf{x} - b_j = 0 \quad \text{for } j = 1, \ldots, m, \\
& g_i(\mathbf{x}) \equiv -x_i \leq 0 \qquad \text{for } i = 1, \ldots, n,
\end{aligned}
\tag{14.15}
$$

where $\mathbf{a_1}, \ldots, \mathbf{a_m}, \mathbf{c} \in \mathbb{R}^n$ and $b_1, \ldots, b_m \in \mathbb{R}$. Since

$$
\begin{aligned}
\nabla f(\mathbf{x}) &= \mathbf{c}, \\
\nabla g_i(\mathbf{x}) &= -\mathbf{e_i} \quad \text{for } i = 1, \ldots, n, \\
\nabla h_j(\mathbf{x}) &= \mathbf{a_j} \quad \text{for } j = 1, \ldots, m,
\end{aligned}
$$

the KKT conditions associated with (14.15) are given by

$$
\mathbf{c} - \sum_{i=1}^{n} v_i \mathbf{e_i} + \sum_{j=1}^{m} w_j \mathbf{a_j} = \mathbf{0},
$$

$$
\begin{aligned}
x_i v_i &= 0 \quad \text{for } i = 1, \ldots, n, \\
v_i &\geq 0 \quad \text{for } i = 1, \ldots, n, \\
\mathbf{a_j}^T \mathbf{x} &= b \quad \text{for } j = 1, \ldots, m, \\
x_i &\geq 0 \quad \text{for } i = 1, \ldots, n.
\end{aligned}
$$

The above system can be written more compactly as follows:

$$
\begin{aligned}
\mathbf{A}\mathbf{x} &= \mathbf{b}, \quad \mathbf{x} \geq \mathbf{0}, &\quad (a) \\
\mathbf{A}^T \mathbf{w} + \mathbf{c} &= \mathbf{v}, \quad \mathbf{v} \geq \mathbf{0}, &\quad (b) \\
\mathbf{x}^T \mathbf{v} &= 0, &\quad (c)
\end{aligned}
$$

where $\mathbf{A}$ is an $m \times n$ matrix whose j–th row is $\mathbf{a_j}$, where $j = 1, \ldots, m$. Readers who are familiar with the theory of linear programming will immediately recognize that (a) is primal feasibility, (b) is dual feasibility, and (c) is complementarity. In particular, when we apply Theorem 14.2.4 to Problem (14.15), we obtain the strong duality theorem of linear programming.   □

14.2.3   Lagrangian Duality

Given an optimization problem $\mathcal{P}$ (the primal problem), we can associate with it a dual problem whose properties are closely related to those of $\mathcal{P}$. To begin our investigation, consider the following primal problem:

$$
(P) \quad
\begin{aligned}
v_p^* = \quad &\inf \quad f(\mathbf{x}) \\
&\text{subject to} \quad g_i(\mathbf{x}) \leq 0 \quad \text{for } i = 1, \ldots, m_1, \\
&\qquad\qquad\; h_j(\mathbf{x}) = 0 \quad \text{for } j = 1, \ldots, m_2, \\
&\qquad\qquad\; \mathbf{x} \in X.
\end{aligned}
$$

Here, $f, g_1, \ldots, g_{m_1}, h_1, \ldots, h_{m_2} : \mathbb{R}^n \to \mathbb{R}$ are *arbitrary* functions, and X is an *arbitrary* nonempty subset of $\mathbb{R}^n$. For the sake of brevity, we shall write the first two sets of constraints in (P) as $\mathbf{g}(\mathbf{x}) \leq \mathbf{0}$ and $\mathbf{h}(\mathbf{x}) = \mathbf{0}$, where $\mathbf{g} : \mathbb{R}^n \to \mathbb{R}^{m_1}$ is given by $\mathbf{g}(\mathbf{x}) = (g_1(\mathbf{x}), \ldots, g_{m_1}(\mathbf{x}))$ and $\mathbf{h} : \mathbb{R}^n \to \mathbb{R}^{m_2}$ is given by $\mathbf{h}(\mathbf{x}) = (h_1(\mathbf{x}), \ldots, h_{m_2}(\mathbf{x}))$.

Now, the **Lagrangian dual problem** associated with (P) is the following problem:

$$
(D) \quad
\begin{aligned}
v_d^* = \quad &\sup \quad \theta(\mathbf{u}, \mathbf{v}) \equiv \inf_{\mathbf{x} \in X} L(\mathbf{x}, \mathbf{u}, \mathbf{v}) \\
&\text{subject to} \quad \mathbf{u} \geq \mathbf{0}.
\end{aligned}
$$

492*CHAPTER 14. UNCONSTRAINED & CONSTRAINED OPTIMIZATION PROBLEMS*

Here, $L : \mathbb{R}^n \times \mathbb{R}^{m_1} \times \mathbb{R}^{m_2} \to \mathbb{R}$ is the Lagrangian function given by

$$L(\mathbf{x}, \mathbf{u}, \mathbf{v}) = f(\mathbf{x}) + \sum_{i=1}^{m_1} u_i g_i(\mathbf{x}) + \sum_{j=1}^{m_2} v_j h_j(\mathbf{x}) = f(\mathbf{x}) + \mathbf{u}^T \mathbf{g}(\mathbf{x}) + \mathbf{v}^T \mathbf{h}(\mathbf{x}). \quad (14.16)$$

Observe that the above formulation can be viewed as a penalty function approach, in the sense that we incorporate the primal constraints $\mathbf{g}(\mathbf{x}) \leq \mathbf{0}$ and $\mathbf{h}(\mathbf{x}) = \mathbf{0}$ into the objective function of (D) using the Lagrange multipliers $\mathbf{u}$ and $\mathbf{v}$. Also, since the set X is arbitrary, there can be many different Lagrangian dual problems for the same primal problem, depending on which constraints are handled as $\mathbf{g}(\mathbf{x}) \leq \mathbf{0}$ and $\mathbf{h}(\mathbf{x}) = \mathbf{0}$, and which constraints are treated by X. However, different choices of the Lagrangian dual problem will in general lead to different outcomes, both in terms of the dual optimal value as well as the computational efforts required to solve the dual problem.

Let us now investigate the relationship between (P) and (D). For any $\bar{\mathbf{x}} \in X$ and $(\bar{\mathbf{u}}, \bar{\mathbf{v}}) \in \mathbb{R}_+^{m_1} \times \mathbb{R}^{m_2}$, we have

$$\inf_{\mathbf{x} \in X} L(\mathbf{x}, \bar{\mathbf{u}}, \bar{\mathbf{v}}) \leq f(\bar{\mathbf{x}}) + Trns\bar{\mathbf{u}}\mathbf{g}(\bar{\mathbf{x}}) + \bar{\mathbf{v}}^T \mathbf{h}(\bar{\mathbf{x}}) \leq \sup_{\mathbf{u} \geq \mathbf{0}} L(\bar{\mathbf{x}}, \mathbf{u}, \mathbf{v}).$$

This implies that

$$\sup_{\mathbf{u} \geq \mathbf{0}} \inf_{\mathbf{x} \in X} L(\mathbf{x}, \mathbf{u}, \mathbf{v}) \leq \inf_{\mathbf{x} \in X} \sup_{\mathbf{u} \geq \mathbf{0}} L(\mathbf{x}, \mathbf{u}, \mathbf{v}). \quad (14.17)$$

In particular, we have the following weak duality theorem, which asserts that the dual objective value is always a lower bound on the primal objective value:

Theorem 14.2.6. (Weak Duality) *Let $\bar{\mathbf{x}}$ be feasible for (P) and $(\bar{\mathbf{u}}, \bar{\mathbf{v}})$ be feasible for (D). Then, we have $\theta(\bar{\mathbf{u}}, \bar{\mathbf{v}}) \leq f(\bar{\mathbf{x}})$. In particular, if $v_d^* = +\infty$, then (P) has no feasible solution.*

Given the primal–dual pair of problems (P) and (D), the **duality gap** between them is defined as $\Delta = v_p^* - v_d^*$. By Theorem 14.2.6, we always have $\Delta \geq 0$. It would be nice to have $\Delta = 0$ (i.e., zero duality gap). However, as the following example shows, this is not true in general.

Example 14.2.5. Consider the following problem from [1, Example 6.2.2]:

$$\begin{aligned} \text{minimize} \quad & f(\mathbf{x}) \equiv -2x_1 + x_2 \\ \text{subject to} \quad & h(\mathbf{x}) \equiv x_1 + x_2 - 3 = 0, \\ & \mathbf{x} \in X, \end{aligned} \quad (14.18)$$

where $X \subset \mathbb{R}^2$ is the following discrete set:

$$X = \{(0,0), (0,4), (4,4), (4,0), (1,2), (2,1)\}.$$

By enumeration, we see that the optimal value of (14.18) is -3, attained at the point $(x_1, x_2) = (2, 1)$. Now, one can verify that the Lagrangian function is given by

$$
\begin{aligned}
\theta(v) &= \min_{x \in X} \left\{ -2x_1 + x_2 + v(x_1 + x_2 - 3) \right\} \\
&= \begin{cases}
-4 + 5v & \text{for } v \leq -1, \\
-8 + v & \text{for } -1 \leq v \leq 2, \\
-3v & \text{for } v \geq 2.
\end{cases}
\end{aligned}
$$

It follows that $\max_v \theta(v) = -6$, which is attained at $v = 2$. Note that the duality gap in this example is $\Delta = -3 - (-6) = 3 > 0$. □

The above example raises the important question of when the duality gap is zero. It turns out that there is a relatively simple answer to this question. Before we proceed, let us introduce the following definition:

Definition 14.2.1. We say that $(\bar{\mathbf{x}}, \bar{\mathbf{u}}, \bar{\mathbf{v}})$ is a **saddle point** of the Lagrangian function L defined in (14.16) if the following conditions are satisfied:

1. $\bar{\mathbf{x}} \in X$,

2. $\bar{\mathbf{u}} \geq \mathbf{0}$, and

3. for all $\mathbf{x} \in X$ and $(\mathbf{u}, \mathbf{v}) \in \mathbb{R}^{m_1} \times \mathbb{R}^{m_2}$ with $\mathbf{u} \geq \mathbf{0}$, we have

$$
L(\bar{\mathbf{x}}, \mathbf{u}, \mathbf{v}) \leq L(\bar{\mathbf{x}}, \bar{\mathbf{u}}, \bar{\mathbf{v}}) \leq L(\mathbf{x}, \bar{\mathbf{u}}, \bar{\mathbf{v}}).
$$

In particular, observe that $(\bar{\mathbf{x}}, \bar{\mathbf{u}}, \bar{\mathbf{v}})$ is a saddle point of L if $\bar{\mathbf{x}}$ *minimizes* L over X when $(\mathbf{u}, \mathbf{v})$ is fixed at $(\bar{\mathbf{u}}, \bar{\mathbf{v}})$, and that $(\bar{\mathbf{u}}, \bar{\mathbf{v}})$ *maximizes* L over all $(\mathbf{u}, \mathbf{v}) \in \mathbb{R}^{m_1} \times \mathbb{R}^{m_2}$ with $\mathbf{u} \geq \mathbf{0}$ when $\mathbf{x}$ is fixed at $\bar{\mathbf{x}}$.

We are now ready to state the following theorem:

Theorem 14.2.7. (Saddle Point Optimality Conditions) *The point* $(\bar{\mathbf{x}}, \bar{\mathbf{u}}, \bar{\mathbf{v}})$ *with* $\bar{\mathbf{x}} \in X$ *and* $\bar{\mathbf{u}} \geq \mathbf{0}$ *is a saddle point of* L *iff*

1. $L(\bar{\mathbf{x}}, \bar{\mathbf{u}}, \bar{\mathbf{v}}) = \min_{\mathbf{x} \in X} L(\mathbf{x}, \bar{\mathbf{u}}, \bar{\mathbf{v}})$,

2. $\mathbf{g}(\bar{\mathbf{x}}) \leq \mathbf{0}$ *and* $\mathbf{h}(\bar{\mathbf{x}}) = \mathbf{0}$, *and*

3. $\bar{\mathbf{u}}^T \mathbf{g}(\bar{\mathbf{x}}) = 0$.

Moreover, the point $(\bar{\mathbf{x}}, \bar{\mathbf{u}}, \bar{\mathbf{v}})$ *is a saddle point of* L *iff* $\bar{\mathbf{x}}$ *and* $(\bar{\mathbf{u}}, \bar{\mathbf{v}})$ *are the optimal solutions to (P) and (D), respectively, with* $f(\bar{\mathbf{x}}) = \theta(\bar{\mathbf{u}}, \bar{\mathbf{v}})$, *i.e., there is no duality gap.*

494 *CHAPTER 14. UNCONSTRAINED & CONSTRAINED OPTIMIZATION PROBLEMS*

In other words, the existence of a saddle point $(\bar{\mathbf{x}}, \bar{\mathbf{u}}, \bar{\mathbf{v}})$ of L implies that

$$\inf_{\mathbf{x} \in X} L(\mathbf{x}, \bar{\mathbf{u}}, \bar{\mathbf{v}}) = L(\bar{\mathbf{x}}, \bar{\mathbf{u}}, \bar{\mathbf{v}}) = \sup_{\mathbf{u} \geq \mathbf{0}} L(\bar{\mathbf{x}}, \mathbf{u}, \mathbf{v}),$$

which in turn implies that

$$\sup_{\mathbf{u} \geq \mathbf{0}} \inf_{\mathbf{x} \in X} L(\mathbf{x}, \mathbf{u}, \mathbf{v}) = \inf_{\mathbf{x} \in X} \sup_{\mathbf{u} \geq \mathbf{0}} L(\mathbf{x}, \mathbf{u}, \mathbf{v}),$$

i.e., inequality (14.17) holds with equality, and $v_p^* = v_d^*$.

Now, if we want to apply Theorem 14.2.7 to certify that the duality gap between (P) and (D) is zero, we need to produce a saddle point of the Lagrangian function L, which is not always an easy task. The following theorem, which is an application of Sion's minimax theorem [7] (see [8] for an elementary proof), provides an easy-to-check sufficient condition for certifying zero duality gap.

Theorem 14.2.8. *Let L be the Lagrangian function defined in* (14.16). *Suppose that*

1. *X is a compact convex subset of $\mathbb{R}^n$,*

2. *$(\mathbf{u}, \mathbf{v}) \mapsto L(\mathbf{x}, \mathbf{u}, \mathbf{v})$ is continuous and concave on $\mathbb{R}_+^{m_1} \times \mathbb{R}^{m_2}$ for each $\mathbf{x} \in X$, and*

3. *$\mathbf{x} \mapsto L(\mathbf{x}, \mathbf{u}, \mathbf{v})$ is continuous and convex on X for each $(\mathbf{u}, \mathbf{v}) \in \mathbb{R}_+^{m_1} \times \mathbb{R}^{m_2}$.*

Then, we have

$$\sup_{\mathbf{u} \geq \mathbf{0}} \inf_{\mathbf{x} \in X} L(\mathbf{x}, \mathbf{u}, \mathbf{v}) = \inf_{\mathbf{x} \in X} \sup_{\mathbf{u} \geq \mathbf{0}} L(\mathbf{x}, \mathbf{u}, \mathbf{v}).$$

Let us now illustrate some of the above results with an example.

Example 14.2.6. (Semidefinite Programming) Consider the following standard form semidefinite programming (SDP):

$$\begin{aligned} \inf \quad & f(\mathbf{Z}) \equiv \operatorname{tr}(\mathbf{CZ}), \\ \text{subject to} \quad & h_j(\mathbf{Z}) \equiv b_j - \operatorname{tr}(\mathbf{a_j Z}) = 0 \quad \text{for } j = 1, \ldots, m, \\ & \mathbf{Z} \in X \equiv \mathbf{S}_+^n, \end{aligned} \tag{14.19}$$

where $\mathbf{c}, \mathbf{a_1}, \ldots, \mathbf{a_m} \in \mathbb{R}^{n \times n}$ are symmetric matrices, $b_1, \ldots, b_m \in \mathbb{R}$ and $\mathbf{S}_+^n$ is the set of $n \times n$ symmetric positive semidefinite matrices. The Lagrangian dual associated with (14.19) is given by

$$\sup \quad \theta(\mathbf{v}) \equiv \inf_{\mathbf{Z} \in \mathbf{S}_+^n} \left\{ \operatorname{tr}(\mathbf{CZ}) + \sum_{j=1}^m v_j (b_j - \operatorname{tr}(\mathbf{a_j Z})) \right\}. \tag{14.20}$$

Now, for any fixed $\mathbf{v} \in \mathbb{R}^m$, we have

$$\theta(\mathbf{v}) = \begin{cases} \mathbf{b}^T\mathbf{v} & \text{if } \mathbf{C} - \sum_{j=1}^{m} v_j\mathbf{a_j} \in \mathbf{S}_+^n, \\ -\infty & \text{otherwise.} \end{cases} \tag{14.21}$$

To see this, let $\mathbf{U\Lambda U}^T$ be the spectral decomposition of $\mathbf{C} - \sum_{j=1}^{m} v_j\mathbf{a_j}$, and suppose that $\Lambda_{ii} < 0$ for some $i = 1, \ldots, n$. Consider the matrix $\mathbf{Z}(\alpha) = \alpha\mathbf{Ue_ie_i}^T\mathbf{U}$. Clearly, we have $\mathbf{Z}(\alpha) \in \mathbf{S}_+^n$ for all $\alpha > 0$. Moreover, as $\alpha \to \infty$, we have

$$\text{tr}\left(\left(\mathbf{C} - \sum_{j=1}^{m} v_j\mathbf{a_j}\right)\mathbf{Z}(\alpha)\right) = \alpha\cdot\text{tr}\left((\mathbf{U\Lambda U}^T)(\mathbf{Ue_ie_i}^T\mathbf{U}^T)\right) = \alpha\cdot\text{tr}\left(\mathbf{\Lambda e_ie_i}^T\right) = \alpha\Lambda_{ii} \to -\infty,$$

whence

$$\theta(\mathbf{v}) = \mathbf{b}^T\mathbf{v} + \inf_{\mathbf{Z}\in\mathbf{S}_+^n} \text{tr}\left(\left(\mathbf{C} - \sum_{j=1}^{m} v_j\mathbf{a_j}\right)\mathbf{Z}\right) = -\infty.$$

On the other hand, if $\mathbf{C} - \sum_{j=1}^{m} v_j\mathbf{a_j} \in \mathbf{S}_+^n$, then we have $\text{tr}\left(\left(\mathbf{C} - \sum_{j=1}^{m} v_j\mathbf{a_j}\right)\mathbf{Z}\right) \geq 0$ for any $\mathbf{Z} \in \mathbf{S}_+^n$. It follows that $\theta(\mathbf{v}) = \mathbf{b}^T\mathbf{v}$ in this case (by taking, say, $\mathbf{Z} = \mathbf{0}$).

Now, using (14.21), we see that (14.20) is equivalent to

$$\begin{aligned} \sup \quad & \mathbf{b}^T\mathbf{v} \\ \text{subject to} \quad & \mathbf{C} - \sum_{j=1}^{m} v_j\mathbf{a_j} \in \mathbf{S}_+^n, \end{aligned} \tag{14.22}$$

which is known as a dual standard form SDP. $\qquad\square$

14.3 Application Examples

In the past decade optimization techniques, especially convex optimization techniques, have been widely used in various engineering fields such as industrial engineering, mechanical engineering, and electrical engineering. For electrical engineering in particular, optimization techniques have been applied to solve problems in communications [9–14], networking [15–19], signal processing [20–22], and even circuit design [23]. In this section, we briefly go through several examples in communications, networking, and signal processing to illustrate how we could apply the results introduced in the previous section to solve real-world problems.

Example 14.3.1. (Power Allocation Optimization in Parallel AWGN Channels) Consider the transmission over n parallel AWGN channels. The ith channel, $i \in \{1, \ldots, n\}$, is characterized by the channel power gain, $h_i \geq 0$, and

the additive Gaussian noise power, $\sigma_i > 0$. Let the transmit power allocated to the ith channel be denoted by $p_i \geq 0$. The maximum information rate that can be reliably transmitted over the ith channel is given by [24]

$$r_i = \log\left(1 + \frac{h_i p_i}{\sigma_i}\right). \tag{14.23}$$

Given a constraint P on the total transmit power over n channels, i.e., $\sum_{i=1}^{n} p_i \leq P$, we want to optimize the allocated power $p_1, \ldots, p_n$ such that the sum rate of n channels, $\sum_{i=1}^{n} r_i$, is maximized. This problem is thus formulated as

$$\begin{aligned} \text{maximize} \quad & \sum_{i=1}^{n} \log\left(1 + \frac{h_i p_i}{\sigma_i}\right) \\ \text{subject to} \quad & \sum_{i=1}^{n} p_i \leq P, \\ & p_i \geq 0 \ \text{ for } i = 1, \ldots, n. \end{aligned} \tag{14.24}$$

For convenience, we rewrite the above problem equivalently as

$$\begin{aligned} \text{minimize} \quad & f(\mathbf{p}) \equiv -\sum_{i=1}^{n} \log\left(1 + \frac{h_i p_i}{\sigma_i}\right) \\ \text{subject to} \quad & h(\mathbf{p}) \equiv \sum_{i=1}^{n} p_i - P \leq 0, \\ & g_i(\mathbf{p}) \equiv -p_i \leq 0 \ \text{ for } i = 1, \ldots, n, \\ & \mathbf{p} \in \mathbb{R}^n, \end{aligned} \tag{14.25}$$

where $\mathbf{p} = [p_1, \ldots, p_n]^T$. It is easy to verify that $f(\mathbf{p})$ is convex, and $h(\mathbf{p})$, $g_1(\mathbf{p}), \ldots, g_n(\mathbf{p})$ are all affine and thus convex. According to Theorem 14.2.5, if we can find a set of feasible solutions $\bar{\mathbf{p}} = [\bar{p}_1, \ldots, \bar{p}_n]^T \in \mathbb{R}^n$ for the above constrained minimization problem as well as a set of $u \geq 0$ and $v_i \geq 0, i = 1, \ldots, n$ such that the following KKT conditions are satisfied,

$$\begin{aligned} \nabla f(\bar{\mathbf{p}}) + u\nabla h(\bar{\mathbf{p}}) + \sum_{i=1}^{n} v_i \nabla g_i(\bar{\mathbf{p}}) &= \mathbf{0}, & \text{(a)} \\ u h(\bar{\mathbf{p}}) &= 0, & \text{(b)} \\ v_i g_i(\bar{\mathbf{p}}) &= 0 \quad \text{ for } i = 1, \ldots, n, & \text{(c)} \end{aligned} \tag{14.26}$$

then we can claim that $\bar{\mathbf{p}}$ is a global minimum for this problem. Suppose that $u > 0$. From (b), it follows that $h(\bar{\mathbf{p}}) = 0$, i.e., $\sum_{i=1}^{n} \bar{p}_i = P$. From (a), it follows that

$$\bar{p}_i = \frac{1}{u - v_i} - \frac{\sigma_i}{h_i} \quad \text{for } i = 1, \ldots, n. \tag{14.27}$$

Suppose that $\bar{p}_i > 0$. From (c), it follows that $v_i = 0$. Then from (14.27), it follows that $\bar{p}_i = \frac{1}{u} - \frac{\sigma_i}{h_i} > 0$. Clearly, if this inequality holds, the corresponding $\bar{p}_i$ will satisfy both (a) and (c). Otherwise, the preassumption of $\bar{p}_i > 0$ cannot be true and the only feasible value for $\bar{p}_i$ is $\bar{p}_i = 0$. In this case, since $\frac{1}{u} - \frac{\sigma_i}{h_i} \leq 0$, we

can always find a $v_i \geq 0$ such that $\bar{p}_i = 0$ holds in (14.27). To summarize, for any $u > 0$, the set of feasible values for $\bar{p}_i$ that satisfy both (a) and (c) are given by

$$\bar{p}_i = \left(\frac{1}{u} - \frac{\sigma_i}{h_i} \right)^+ \quad \text{for } i = 1, \ldots, n, \tag{14.28}$$

where $(x)^+ = \max(0, x)$ for $x \in \mathbb{R}$. Furthermore, recall that this set of $\bar{p}_i$'s need to satisfy $\sum_{i=1}^n \bar{p}_i = P$, i.e.,

$$\sum_{i=1}^n \left(\frac{1}{u} - \frac{\sigma_i}{h_i} \right)^+ = P. \tag{14.29}$$

Note that for any $P > 0$, in the above equation there exists a unique positive root of u (which can be found numerically by a simple bisection search over the interval $0 < u < \max_i(h_i/\sigma_i)$). With the root of u, the corresponding $\bar{p}_i$'s given in (14.28) satisfy all the KKT conditions in (a), (b), and (c), and are thus the global optimal solutions for Problem (14.25). It is worth noting that the structure for the optimal power allocation in (14.28) is known as the "water-filling" solution [24].   □

Example 14.3.2. (Transmit Optimization for MIMO AWGN Channels with Per-Antenna Power Constraints) Consider the transmission over a MIMO AWGN channel with n transmitting antennas and m receiving antennas. The propagation channel from the transmitter to the receiver is represented by a real matrix, $\mathbf{H} \in \mathbb{R}^{m \times n}$, in which all the columns are assumed to be "non-empty", i.e., there is at least one element in each column being non-zero. The additive noises at m receiving antennas are assumed to be i.i.d. Gaussian random variables with zero mean and unit variance. The transmit signals from the ith antenna, $i \in \{1, \ldots, n\}$, are denoted by $x_i(t) \in \mathbb{R}, t = 0, 1, \ldots$, which are subject to a per-antenna average power constraint P_i, i.e., $E\left\{ (x_i(t))^2 \right\} \leq P_i$, where $E\{\cdot\}$ denotes the expectation. Let $\mathbf{Z} \in \mathbf{S}_+^n$ denote the transmit covariance matrix, i.e., $\mathbf{Z} = E\left\{ \mathbf{x}(t) (\mathbf{x}(t))^T \right\}$, where $\mathbf{x}(t) = [x_1(t), \ldots, x_n(t)]^T$. The set of per-antenna transmit power constraints can then be expressed as

$$\text{tr}(\mathbf{A}_i \mathbf{Z}) \leq P_i \quad \text{for } i = 1, \ldots, n, \tag{14.30}$$

where $\mathbf{A}_i \in \mathbb{R}^{n \times n}$ is a matrix with all zero elements expect for the ith diagonal element being one.

For any transmit covariance matrix $\mathbf{Z} \in \mathbf{S}_+^n$, the maximum transmit rate over the MIMO AWGN channel is given by [25]

$$r = \log \det \left(\mathbf{I} + \mathbf{H}\mathbf{Z}\mathbf{H}^T \right), \tag{14.31}$$

where $\mathbf{I}$ denotes an identity matrix. The problem of our interest here is to maximize the rate r over $\mathbf{Z} \in \mathbf{S}_+^n$ subject to the set of per-antenna transmit power constraints,

where $\mathbf{U} \in \mathbb{R}^{m \times m}$ and $\mathbf{V} \in \mathbb{R}^{n \times n}$ are unitary matrices, and $\mathbf{\Lambda} \in \mathbb{R}^{m \times n}$ is a diagonal matrix with the diagonal elements being denoted by $\lambda_1, \ldots, \lambda_k, k = \min(m, n)$, and $\lambda_i \geq 0, i = 1, \ldots, k$. Substituting (14.37) into (14.36) and using the fact that $\log \det(\mathbf{I} + \mathbf{AB}) = \log \det(\mathbf{I} + \mathbf{BA})$ yield

$$\min_{\bar{\mathbf{Z}} \in \mathbf{S}_+^n} -\log \det\left(\mathbf{I} + \mathbf{\Lambda} \mathbf{V}^T \bar{\mathbf{Z}} \mathbf{V} \mathbf{\Lambda}^T\right) + \operatorname{tr}\left(\bar{\mathbf{Z}}\right). \tag{14.38}$$

By letting $\hat{\mathbf{Z}} = \mathbf{V}^T \bar{\mathbf{Z}} \mathbf{V}$ and using the fact that $\operatorname{tr}\left(\hat{\mathbf{Z}}\right) = \operatorname{tr}\left(\bar{\mathbf{Z}}\right)$, we obtain an equivalent problem of (14.38) as

$$\min_{\hat{\mathbf{Z}} \in \mathbf{S}_+^n} -\log \det\left(\mathbf{I} + \mathbf{\Lambda} \hat{\mathbf{Z}} \mathbf{\Lambda}^T\right) + \operatorname{tr}\left(\hat{\mathbf{Z}}\right). \tag{14.39}$$

Recall the Hadamard's inequality [24], which states that for any $\mathbf{X} \in \mathbf{S}_+^m$, $\det(\mathbf{X}) \leq \prod_{i=1}^{m} \mathbf{X}_{ii}$, iff $\mathbf{X}$ is a diagonal matrix, where $\mathbf{X}_{ii}$ denotes the ith diagonal element of $\mathbf{X}$, $i = 1, \ldots, m$. Applying this result to Problem (14.39), it follows that the minimum value for this problem is attained iff $\hat{\mathbf{Z}}$ is a diagonal matrix. Let the diagonal elements of $\hat{\mathbf{Z}}$ be denoted by $p_1, \ldots, p_n$. Since $\hat{\mathbf{Z}} \in \mathbf{S}_+^n$, problem (14.39) can be simplified as

$$\begin{aligned} &\text{minimize} && -\sum_{i=1}^{n} \log(1 + \lambda_i^2 p_i) + \sum_{i=1}^{n} p_i \\ &\text{subject to} && p_i \geq 0 \;\; \text{for } i = 1, \ldots, n. \end{aligned} \tag{14.40}$$

Note that in the above problem, for convenience we have assumed that $\lambda_i = 0$, for $i = k + 1, \ldots, n$. Similar to Exercise 14.3.1, the global minimum for the above problem can be shown to be the following water-filling solution:

$$p_i = \left(1 - \frac{1}{\lambda_i^2}\right)^+ \quad \text{for } i = 1, \ldots, n. \tag{14.41}$$

To summarize, for any given $\mathbf{u} > \mathbf{0}$, the optimal solution for the minimization problem in (14.35) is given by

$$\mathbf{Z}_u = \mathbf{A}_u^{-1/2} \mathbf{V} \hat{\mathbf{Z}} \mathbf{V}^T \mathbf{A}_u^{-1/2}, \tag{14.42}$$

where $\hat{\mathbf{Z}}$ is a diagonal matrix with the diagonal elements given in (14.41). Moreover, the dual function $\theta(\mathbf{u})$ in (14.35) can be simplified to be

$$\theta(\mathbf{u}) = \begin{cases} -\sum_{i=1}^{k} \left(\log(\lambda_i^2)\right)^+ + \sum_{i=1}^{k} \left(1 - \frac{1}{\lambda_i^2}\right)^+ - \sum_{i=1}^{n} u_i P_i & \text{if } \mathbf{u} > \mathbf{0} \\ -\infty & \text{otherwise,} \end{cases} \tag{14.43}$$

where $\lambda_1, \ldots, \lambda_k$ are related to $\mathbf{u}$ via (14.37).

Next, we solve the dual problem (14.34) by maximizing the dual function $\theta(\mathbf{u})$ in (14.43) over $\mathbf{u} \geq \mathbf{0}$. The corresponding dual optimal solution of $\mathbf{u}$ then leads to the optimal solution of $\mathbf{Z}_u$ in (14.42) for the primal problem (14.32). Since $v_d^* = v_p^* \geq 0$, in fact we only need to consider the maximization of $\theta(\mathbf{u})$ over $\mathbf{u} > \mathbf{0}$ in (14.43). However, due to the coupled structure of λ_i's and u_i's shown in (14.37), it is not evident whether $\theta(\mathbf{u})$ in (14.43) is differentiable over u_i's for $\mathbf{u} > \mathbf{0}$. As a result, conventional decent methods to find the global minimum for differentiable convex functions such as Newton's method are ineffective for our problem at hand. Thus, we resort to an alternative method, known as *subgradient* based method, to handle the non-differentiable function $\theta(\mathbf{u})$. First, we introduce the definition of subgradient for an arbitrary real-valued function $z(\mathbf{x})$ defined over a nonempty convex set $S \subset \mathbb{R}^n$. We assume that $z(\mathbf{x})$ has a finite maximum. However, $z(\mathbf{x})$ need not be continuously differentiable nor have an analytical expression for its differential. In this case, a vector $\mathbf{v} \in \mathbb{R}^n$ is called the subgradient of $z(\mathbf{x})$ at point $\mathbf{x} = \mathbf{x}_0$ if for any $\mathbf{x} \in S$, the following inequality holds:

$$z(\mathbf{x}) \leq z(\mathbf{x}_0) + \mathbf{v}^T(\mathbf{x} - \mathbf{x}_0). \tag{14.44}$$

If at any point $\mathbf{x} \in S$ a corresponding subgradient $\mathbf{v}$ for $z(\mathbf{x})$ is attainable, then the maximum of $z(\mathbf{x})$ can be found via an iterative search over $\mathbf{x} \in S$ based on $\mathbf{v}$ (see, e.g., the ellipsoid method [26]). Since $\theta(\mathbf{u})$ is defined over a convex set $\mathbf{u} > \mathbf{0}$ and has a finite maximum, the dual problem (14.34) can thus be solved by a subgradient based method. Next, we show that the subgradient of $\theta(\mathbf{u})$ at any point $\mathbf{u} > \mathbf{0}$ is given by $[\mathrm{tr}\,(\mathbf{A}_1\mathbf{Z}_u) - P_1, \ldots, \mathrm{tr}\,(\mathbf{A}_n\mathbf{Z}_u) - P_n]^T$, where $\mathbf{Z}_u$ is given in (14.42). Suppose that at any two points $\mathbf{u} > \mathbf{0}$ and $\mathbf{u}' > \mathbf{0}$, $\theta(\mathbf{u})$ and $\theta(\mathbf{u}')$ are attained by $\mathbf{Z} = \mathbf{Z}_u$ and $\mathbf{Z} = \mathbf{Z}_u'$, respectively. Then, we have the following inequalities:

$$\begin{aligned}
\theta(\mathbf{u}') &= L(\mathbf{Z}_u', \mathbf{u}') \\
&= \min_{\mathbf{Z} \in \mathbf{S}_+^n} L(\mathbf{Z}, \mathbf{u}') \\
&\leq L(\mathbf{Z}_u, \mathbf{u}') \\
&= -\log\det\left(\mathbf{I} + \mathbf{H}\mathbf{Z}_u\mathbf{H}^T\right) + [\mathrm{tr}\,(\mathbf{A}_1\mathbf{Z}_u) - P_1, \ldots, \mathrm{tr}\,(\mathbf{A}_n\mathbf{Z}_u) - P_n]\mathbf{u}' \\
&= -\log\det\left(\mathbf{I} + \mathbf{H}\mathbf{Z}_u\mathbf{H}^T\right) + [\mathrm{tr}\,(\mathbf{A}_1\mathbf{Z}_u) - P_1, \ldots, \mathrm{tr}\,(\mathbf{A}_n\mathbf{Z}_u) - P_n]\mathbf{u} \\
&\quad + [\mathrm{tr}\,(\mathbf{A}_1\mathbf{Z}_u) - P_1, \ldots, \mathrm{tr}\,(\mathbf{A}_n\mathbf{Z}_u) - P_n](\mathbf{u}' - \mathbf{u}) \\
&= L(\mathbf{Z}_u, \mathbf{u}) + [\mathrm{tr}\,(\mathbf{A}_1\mathbf{Z}_u) - P_1, \ldots, \mathrm{tr}\,(\mathbf{A}_n\mathbf{Z}_u) - P_n](\mathbf{u}' - \mathbf{u}) \\
&= \theta(\mathbf{u}) + [\mathrm{tr}\,(\mathbf{A}_1\mathbf{Z}_u) - P_1, \ldots, \mathrm{tr}\,(\mathbf{A}_n\mathbf{Z}_u) - P_n](\mathbf{u}' - \mathbf{u}),
\end{aligned}$$

from which the subgradient of $\theta(\mathbf{u})$ follows.

Last, we can verify that the optimal primal and dual solutions, $\mathbf{Z}_u$ given in

(14.42) and the corresponding $\mathbf{u} > 0$ satisfy (a) of the following KKT conditions:

$$\nabla f(\mathbf{Z}_u) + \sum_{i=1}^{n} u_i \nabla g_i(\mathbf{Z}_u) \;=\; \mathbf{0}, \qquad\qquad \text{(a)}$$
$$u_i g_i(\mathbf{Z}_u) \;=\; 0 \quad \text{ for } i = 1, \ldots, n, \qquad \text{(b)}$$

(14.45)

while since $\mathbf{u} > 0$, from (b) it follows that $g_i(\mathbf{Z}_u) = 0$, i.e., $\text{tr}(\mathbf{A}_i \mathbf{Z}_u) = P_i$ must hold for $i = 1, \ldots, n$. Thus, all transmit antennas should transmit with their maximum power levels with the optimal transmit covariance matrix $\mathbf{Z}_u$, which is consistent with the observation that the subgradient of the dual function $\theta(\mathbf{u})$ at the optimal dual solution of $\mathbf{u}$ should vanish to $\mathbf{0}$. $\qquad\square$

Example 14.3.3. (Power Efficient Beamforming in Two-Way Relay Network via SDP Relaxation) In this example, we illustrate how an originally nonconvex problem can be solved via convex techniques. As shown in Figure 14.1, we consider a two-way relay channel (TWRC) consisting of two source nodes, S1 and S2, each with a single antenna and a relay node, R, equipped with M antennas, $M \geq 2$. It is assumed that the transmission protocol of TWRC uses two consecutive equal-duration timeslots for one round of information exchange between S1 and S2 via R. During the first timeslot, both S1 and S2 transmit concurrently to R, which linearly processes the received signal and then broadcasts the resulting signal to S1 and S2 during the second timeslot. It is also assumed that perfect synchronization has been established among S1, S2, and R prior to data transmission. The received baseband signal at R in the first timeslot is expressed as

$$\mathbf{y}_R(n) = \mathbf{h}_1 \sqrt{p_1} s_1(n) + \mathbf{h}_2 \sqrt{p_2} s_2(n) + \mathbf{z}_R(n) \tag{14.46}$$

where $\mathbf{y}_R(n) \in \mathbb{C}^M$ is the received signal vector at symbol index n, $n = 1, \ldots, N$, with N denoting the total number of transmitted symbols during one timeslot; $\mathbf{h}_1 \in \mathbb{C}^M$ and $\mathbf{h}_2 \in \mathbb{C}^M$ represent the channel vectors from S1 to R and from S2 to R, respectively, which are assumed to be constant during the two timeslots; $s_1(n)$ and $s_2(n)$ are the transmitted symbols from S1 and S2, respectively, with $E\{|s_1(n)|\} = 1$, $E\{|s_2(n)|\} = 1$, and $|\cdot|$ denoting the absolute value for a complex number; p_1 and p_2 denote the transmit powers of S1 and S2, respectively; and $\mathbf{z}_R(n) \in \mathbb{C}^M$ is the receiver noise vector, independent over n, and without loss of generality, it is assumed that $\mathbf{z}_R(n)$ has a circular symmetric complex Gaussian (CSCG) distribution with zero mean and identity covariance matrix, denoted by $\mathbf{z}_R(n) \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}), \forall n$. Upon receiving the mixed signal from S1 and S2, R processes it with amplify-and-forward (AF) relay operation, also known as *linear analogue relaying*, and then broadcasts the processed signal to S1 and S2 during the second timeslot. Mathematically, the linear processing (beamforming) operation at the relay can be concisely represented as

$$\mathbf{x}_R(n) = \mathbf{A} \mathbf{y}_R(n), \ n = 1, \ldots, N \tag{14.47}$$
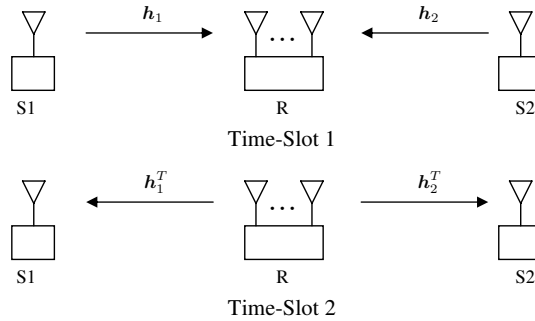
Figure 14.1: The two-way multiantenna relay channel.

where $\mathbf{x}_R(n) \in \mathbb{C}^M$ is the transmitted signal at R, and $\mathbf{A} \in \mathbb{C}^{M \times M}$ is the relay processing matrix.

Note that the transmit power of R can be shown equal to

$$
\begin{aligned}
p_R(\mathbf{A}) &= \mathrm{E}\left[\mathrm{tr}\left(\mathbf{x}_R(n)\mathbf{x}_R^H(n)\right)\right] \\
&= \|\mathbf{A}\mathbf{h}_1\|_2^2\, p_1 + \|\mathbf{A}\mathbf{h}_2\|_2^2\, p_2 + \mathrm{tr}(\mathbf{A}\mathbf{A}^H).
\end{aligned} \tag{14.48}
$$

We can assume w.l.o.g. that channel reciprocity holds for TWRC during uplink and downlink transmissions, i.e., the channels from R to S1 and S2 during the second timeslot are given as $\mathbf{h}_1^T$ and $\mathbf{h}_2^T$, respectively. Thus, the received signals at S1 can be written as

$$
\begin{aligned}
y_1(n) &= \mathbf{h}_1^T \mathbf{x}_R(n) + z_1(n) \\
&= \mathbf{h}_1^T \mathbf{A}\mathbf{h}_1 \sqrt{p_1}\, s_1(n) + \mathbf{h}_1^T \mathbf{A}\mathbf{h}_2 \sqrt{p_2}\, s_2(n) + \mathbf{h}_1^T \mathbf{A}\mathbf{z}_R(n) + z_1(n) \quad (14.49)
\end{aligned}
$$

for $n = 1, \ldots, N$, where $z_1(n)$'s are the independent receiver noise samples at S1, and it is assumed that $z_1(n) \sim \mathcal{CN}(0,1), \forall n$. Note that on the right-hand side of (14.49), the first term is the self-interference of S1, while the second term contains the desired message from S2. Assuming that both $\mathbf{h}_1^T \mathbf{A}\mathbf{h}_1$ and $\mathbf{h}_1^T \mathbf{A}\mathbf{h}_2$ are perfectly known at S1 via training-based channel estimation prior to data transmission, S1 can first subtract its self-interference from $y_1(n)$ and then coherently demodulate $s_2(n)$. The above practice is known as analogue network coding (ANC). From (14.49), subtracting the self-interference from $y_1(n)$ yields

$$
\tilde{y}_1(n) = \tilde{h}_{21}\sqrt{p_2}\, s_2(n) + \tilde{z}_1(n), \ n = 1, \ldots, N \tag{14.50}
$$

where $\tilde{h}_{21} = \mathbf{h}_1^T \mathbf{A}\mathbf{h}_2$, and $\tilde{z}_1(n) \sim \mathcal{CN}(0, \left\|\mathbf{A}^H\mathbf{h}_1^*\right\|_2^2 + 1)$, where $*$ denotes the complex conjugate. From (14.50), for a given $\mathbf{A}$, the maximum achievable SNR

for the end-to-end link from S2 to S1 via R, denoted by γ_{21}, is given as

$$\gamma_{21} = \frac{\left| \mathbf{h}_1^T \mathbf{A} \mathbf{h}_2 \right|^2 p_2}{\left\| \mathbf{A}^H \mathbf{h}_1^* \right\|_2^2 + 1} \tag{14.51}$$

Similarly, it can be shown that the maximum SNR γ_{12} for the link from S1 to S2 via R is given as

$$\gamma_{12} = \frac{\left| \mathbf{h}_2^T \mathbf{A} \mathbf{h}_1 \right|^2 p_1}{\left\| \mathbf{A}^H \mathbf{h}_2^* \right\|_2^2 + 1}. \tag{14.52}$$

Now we minimize the relay transmission power given in (14.48), under the constraints that the achievable SNRs γ_{21} and γ_{12} over the two directions are above two target values, $\bar{\gamma}_1$ and $\bar{\gamma}_2$. As such, the optimization can be formulated as

$$\begin{aligned}
\text{minimize}_{\mathbf{A}} \quad & p_R := \|\mathbf{A}\mathbf{h}_1\|_2^2 p_1 + \|\mathbf{A}\mathbf{h}_2\|_2^2 p_2 + \text{tr}(\mathbf{A}\mathbf{A}^H) \\
\text{subject to} \quad & \left| \mathbf{h}_1^T \mathbf{A} \mathbf{h}_2 \right|^2 \geq \frac{\bar{\gamma}_1}{p_2} \left\| \mathbf{A}^H \mathbf{h}_1^* \right\|_2^2 + \frac{\bar{\gamma}_1}{p_2} \\
& \left| \mathbf{h}_2^T \mathbf{A} \mathbf{h}_1 \right|^2 \geq \frac{\bar{\gamma}_2}{p_1} \left\| \mathbf{A}^H \mathbf{h}_2^* \right\|_2^2 + \frac{\bar{\gamma}_2}{p_1},
\end{aligned} \tag{14.53}$$

For the convenience of analysis, we further modify the above problem as follows. First, let $\text{Vec}(\mathbf{Q})$ be a $K^2 \times 1$ vector associated with a $K \times K$ square matrix $\mathbf{Q} = [\mathbf{q}_1, \ldots, \mathbf{q}_K]^T$, where $\mathbf{q}_k \in \mathbb{C}^K, k = 1, \ldots, K$, by the rule of $\text{Vec}(\mathbf{Q}) = \left[\mathbf{q}_1^T, \ldots, \mathbf{q}_K^T \right]^T$. Next, with $\mathbf{b} = \text{Vec}(\mathbf{A})$ and $\boldsymbol{\Theta} = p_1 \mathbf{h}_1 \mathbf{h}_1^H + p_2 \mathbf{h}_2 \mathbf{h}_2^H + \mathbf{I}$, we can express p_R in the objective function of (14.53) as $p_R = \text{tr} \left(\mathbf{A} \boldsymbol{\Theta} \mathbf{A}^H \right) = \|\boldsymbol{\Phi} \mathbf{b}\|_2^2$, with $\boldsymbol{\Phi} = (\text{diag}(\boldsymbol{\Theta}^T, \boldsymbol{\Theta}^T))^{\frac{1}{2}}$, where $\text{diag}(\mathbf{A}, \mathbf{B})$ denotes a block-diagonal matrix with $\mathbf{A}$ and $\mathbf{B}$ as the diagonal square matrices. Similarly, let $\mathbf{f}_1 = \text{Vec} \left(\mathbf{h}_1 \mathbf{h}_2^T \right)$ and $\mathbf{f}_2 = \text{Vec} \left(\mathbf{h}_2 \mathbf{h}_1^T \right)$. Then, from (14.53) it follows that $\left| \mathbf{h}_1^T \mathbf{A} \mathbf{h}_2 \right|^2 = \left| \mathbf{f}_1^T \mathbf{b} \right|^2$ and $\left| \mathbf{h}_2^T \mathbf{A} \mathbf{h}_1 \right|^2 = \left| \mathbf{f}_2^T \mathbf{b} \right|^2$. Furthermore, by defining

$$\mathbf{h}_i = \left[\begin{array}{cccc} \mathbf{h}_i(1,1) & 0 & \mathbf{h}_i(2,1) & 0 \\ 0 & \mathbf{h}_i(1,1) & 0 & \mathbf{h}_i(2,1) \end{array} \right], \; i = 1, 2,$$

we have $\left\| \mathbf{A}^H \mathbf{h}_i^* \right\|_2^2 = \|\mathbf{h}_i \mathbf{b}\|_2^2, i = 1, 2$. Using the above transformations, (14.53) can be rewritten as

$$\begin{aligned}
\text{minimize}_{\mathbf{b}} \quad & p_R := \|\boldsymbol{\Phi} \mathbf{b}\|_2^2 \\
\text{subject to} \quad & \left| \mathbf{f}_1^T \mathbf{b} \right|^2 \geq \frac{\bar{\gamma}_1}{p_2} \|\mathbf{h}_1 \mathbf{b}\|_2^2 + \frac{\bar{\gamma}_1}{p_2} \\
& \left| \mathbf{f}_2^T \mathbf{b} \right|^2 \geq \frac{\bar{\gamma}_2}{p_1} \|\mathbf{h}_2 \mathbf{b}\|_2^2 + \frac{\bar{\gamma}_2}{p_1}.
\end{aligned} \tag{14.54}$$

504 *CHAPTER 14. UNCONSTRAINED & CONSTRAINED OPTIMIZATION PROBLEMS*

The above problem can be shown to be still nonconvex. However, in the following, we show that the exact optimal solution could be obtained via a relaxed semidefinite programming (SDP) problem.

We first define $\mathbf{E}_0 = \mathbf{\Phi}^H\mathbf{\Phi}$, $\mathbf{E}_1 = \frac{p_2}{\gamma_1}\mathbf{f}_1^*\mathbf{f}_1^T - \mathbf{h}_1^H\mathbf{h}_1$, and $\mathbf{E}_2 = \frac{p_1}{\gamma_2}\mathbf{f}_2^*\mathbf{f}_2^T - \mathbf{h}_2^H\mathbf{h}_2$. Since standard SDP formulations only involve real variables and constants, we introduce a new real matrix variable as $\mathbf{X} = [\mathbf{b}_R; \mathbf{b}_I] \times [\mathbf{b}_R; \mathbf{b}_I]^T$, where $\mathbf{b}_R = Re(\mathbf{b})$ and $\mathbf{b}_I = Im(\mathbf{b})$ are the real and imaginary parts of $\mathbf{b}$, respectively. To rewrite the norm representations at (14.54) in terms of $\mathbf{X}$, we need to rewrite $\mathbf{E}_0$, $\mathbf{E}_1$, and $\mathbf{E}_2$, as expanded matrices $\mathbf{F}_0$, $\mathbf{F}_1$, and $\mathbf{F}_2$, respectively, in terms of their real and imaginary parts. Specifically, to write out $\mathbf{F}_0$, we first define the short notations $\mathbf{\Phi}_R = Re(\mathbf{\Phi})$ and $\mathbf{\Phi}_I = Im(\mathbf{\Phi})$; then we have

$$\mathbf{F}_0 \;\;=\;\; \left[\begin{array}{cc} \mathbf{\Phi}_R^T\mathbf{\Phi}_R + \mathbf{\Phi}_I^T\mathbf{\Phi}_I & \mathbf{\Phi}_I^T\mathbf{\Phi}_R - \mathbf{\Phi}_R^T\mathbf{\Phi}_I \\ \mathbf{\Phi}_R^T\mathbf{\Phi}_I - \mathbf{\Phi}_I^T\mathbf{\Phi}_R & \mathbf{\Phi}_R^T\mathbf{\Phi}_R + \mathbf{\Phi}_I^T\mathbf{\Phi}_I \end{array} \right].$$

The expanded matrices $\mathbf{F}_1$ and $\mathbf{F}_2$ can be generated from $\mathbf{E}_1$ and $\mathbf{E}_2$ in a similar way, where the two terms in $\mathbf{E}_1$ or $\mathbf{E}_2$ could first be expanded separately then summed together.

As such, problem (14.54) can be equivalently rewritten as

$$\begin{aligned} &\text{minimize}_{\mathbf{X}} \;\; p_R := \text{tr}\,(\mathbf{F}_0\mathbf{X}) \\ &\text{subject to} \;\; \text{tr}\,(\mathbf{F}_1\mathbf{X}) \geq 1, \;\; \text{tr}\,(\mathbf{F}_2\mathbf{X}) \geq 1, \;\; \mathbf{X} \succeq 0, \\ &\hspace{4.2em} \text{rank}(\mathbf{X}) = 1. \end{aligned} \tag{14.55}$$

The above problem is still not convex given the last rank-one constraint. However, if we remove such a constraint, this problem is relaxed into a convex SDP problem as shown below.

$$\begin{aligned} &\text{minimize}_{\mathbf{X}} \;\; p_R := \text{tr}\,(\mathbf{F}_0\mathbf{X}) \\ &\text{subject to} \;\; \text{tr}\,(\mathbf{F}_1\mathbf{X}) \geq 1, \;\; \text{tr}\,(\mathbf{F}_2\mathbf{X}) \geq 1, \;\; \mathbf{X} \succeq 0. \end{aligned} \tag{14.56}$$

Given the convexity of the above SDP problem, the optimal solution could be efficiently found by various convex optimization methods. Note that SDP relaxation usually leads to an optimal $\mathbf{X}$ for problem (14.56) that is of rank r with $r \geq 1$, which makes it impossible to reconstruct the exact optimal solution for Problem (14.54) when $r > 1$. A commonly adopted method in the literature to obtain a feasible rank-one (but in general suboptimal) solution from the solution of SDP relaxation is via "randomization" (see, e.g., [27] and references therein). Fortunately, we show in the following that with the special structure in Problem (14.56), we could efficiently reconstruct an optimal rank-one solution from its optimal solution that could be of rank r with $r > 1$, based on some elegant results derived for SDP relaxation in [28]. In other words, we could obtain the exact optimal solution for the nonconvex problem in (14.55) without losing any optimality, and as efficiently as solving a convex problem.

Theorem 14.3.1. *Assume that an optimal solution* $\mathbf{X}^\star$ *of rank* $r > 1$ *has been found for Problem* (14.56)*, we could efficiently construct another feasible optimal solution* $\mathbf{X}^{\star\star}$ *of rank one, i.e.,* $\mathbf{X}^{\star\star}$ *is the optimal solution for both* (14.55) *and* (14.56)*.*

Proof: Please refer to [21]. $\qquad\square$

Note that the above proof is self-constructive, based on which we could easily obtain a routine to obtain an optimal rank-one solution for Problem (14.55) from $\mathbf{X}^\star$. Then we could map the solution back to obtain an optimal solution for the problem in (14.53). $\qquad\square$

14.4 Exercises

Exercise 14.4.1. Please indicate whether the following sets are convex or not.

1. $\left\{\mathbf{x} : \frac{\mathbf{a}^T\mathbf{x}-\mathbf{b}}{\mathbf{c}^T\mathbf{x}+\mathbf{d}} \leq 1; \mathbf{c}^T\mathbf{x} + \mathbf{d} < 0\right\}$;

2. $\{\mathbf{x} : \mathbf{A}\mathbf{x} = \mathbf{b}, \|\mathbf{x}\|_2 = 1\}$;

3. $\{\mathbf{X} : X_{11}\mathbf{a_0} + X_{22}\mathbf{a_1} \succeq 0, \mathbf{a_0} \in \mathbf{S}^n, \mathbf{a_1} \in \mathbf{S}^n\}$; ($X_{ij}$ stands for the ijth element in matrix $\mathbf{X}$)

4. $\left\{\mathbf{X} : \mathbf{a}^T\mathbf{X}\mathbf{a} = 1\right\}$

Exercise 14.4.2. Please indicate whether the following functions are convex or concave or neither.

1. $f(\mathbf{x}) = \sup_w \left\{\log \sum_{i=1}^n e^{\frac{x_i}{w}}\right\}$;

2. $f(\mathbf{x}) = -(x_1 x_2 x_3)^{1/3}$, $\mathbf{x} > \mathbf{0}$;

3. $f(\mathbf{X}) = \log \det\left(\mathbf{A}^T\mathbf{X}\mathbf{A}\right)$; $\mathbf{X} \succ \mathbf{0}$;

4. $f(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x} + 2\mathbf{x} - 5$, $\mathbf{A} = \left[\begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array}\right]$.

Exercise 14.4.3. With the following problem formulation, answer the followup questions.

$$\begin{array}{ll} \text{minimize}_{\mathbf{x}} & -(x_1 + x_2) \\ \text{subject to} & \|\mathbf{a_1}\mathbf{x}\|_2 \leq 1, \\ & \|\mathbf{a_2}\mathbf{x} - \mathbf{b_2}\|_2 \leq 1, \end{array}$$

where $\mathbf{x} = [x_1, x_2]^T$, $\mathbf{a_1} = \mathbf{a_2} = \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right]$, and $\mathbf{b_2} = [1, 0]^T$.

506*CHAPTER 14. UNCONSTRAINED & CONSTRAINED OPTIMIZATION PROBLEMS*

1. Is this problem convex?

2. Does Slater's constraint condition hold?

3. What is the optimal solution for this problem? (Hint: Try to solve this problem graphically if the KKT conditions are hard to solve.)

4. What is the optimal objective value for the dual problem?

5. What is the optimal value for the dual variable associated with the second constraint?

Exercise 14.4.4. Given the optimization problem shown in Exercise 14.4.3, please reformulate it as a semidefinite programming (SDP) problem, then derive the dual problem of the resulting SDP problem.

Exercise 14.4.5. With the following optimization problem, answer the followup questions.

$$\text{maximize}_{\mathbf{P}} \qquad \sum_{i=1}^{n} \log(1 + \frac{P_i}{\delta_i})$$

$$\text{subject to} \qquad \sum_{i=1}^{n} P_i = P_{total},$$

$$\mathbf{P} \geq \mathbf{0},$$

where $\mathbf{P} = [P_1, \ldots, P_n]^T$, and $\delta_i > 0, i = 1, \ldots, n$.

1. Is KKT sufficient for us to get the optimal solution for the above problem?

2. Is KKT necessary for the optimal solution?

3. Please write out the KKT conditions for this problem.

4. Please solve the general form of optimal P_i's.

5. If $n = 3$, $\delta_1 = 2$, $\delta_2 = 10$, $\delta_3 = 5$, and $P_{total} = 10$, what are the optimal P_i values?

Exercise 14.4.6. Let $1 \leq m \leq n$ be integers, and let $\mathbf{A}$ be an $m \times n$ matrix with full row rank. Furthermore, let $\mathbf{c} \in \mathbb{R}^n$ and $\mathbf{Q} \in \mathbf{S}_{++}^n$ be given. Consider the following optimization problem:

$$\begin{aligned} \text{minimize} \quad & \frac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x} + \mathbf{c}^T\mathbf{x} \\ \text{subject to} \quad & \mathbf{A}\mathbf{x} = \mathbf{0}. \end{aligned} \qquad (14.57)$$

1. Explain why the KKT conditions are necessary and sufficient for (14.57).

2. Write down the KKT conditions associated with (14.57). Hence, express the optimal solution to (14.57) in closed form.

Exercise 14.4.7. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a differentiable convex function. Consider the following problem:

$$
\begin{aligned}
\text{minimize} \quad & f(\mathbf{x}) \\
\text{subject to} \quad & \mathbf{x} \geq \mathbf{0}.
\end{aligned}
\tag{14.58}
$$

Show that $\bar{\mathbf{x}} \in \mathbb{R}^n$ is an optimal solution to (14.58) iff it satisfies the following system:

$$
\begin{aligned}
\nabla f(\bar{\mathbf{x}}) &\geq \mathbf{0}, \\
\bar{\mathbf{x}} &\geq \mathbf{0}, \\
\bar{\mathbf{x}}^T \nabla f(\bar{\mathbf{x}}) &= 0.
\end{aligned}
$$

Exercise 14.4.8. This problem is concerned with finding the minimum-volume enclosing ellipsoid of a set of vectors.

1. Let $\mathbf{u} \in \mathbb{R}^n$ be fixed, and define the function $g : \mathbf{S}^n \to \mathbb{R}_+$ by $g(\mathbf{X}) = \|\mathbf{X}\mathbf{u}\|_2^2$. Find $\nabla g(\mathbf{X})$.

2. Let $V = \{\mathbf{v^1}, \ldots, \mathbf{v^m}\} \subset \mathbb{R}^n$ be a set of vectors that span $\mathbb{R}^n$. Consider the following problem:

$$
\begin{aligned}
\inf \quad & -\log \det(\mathbf{X}) \\
\text{subject to} \quad & \|\mathbf{X}\mathbf{v^i}\|_2^2 \leq 1 \quad i = 1, \ldots, m, \\
& \mathbf{X} \in \mathbf{S}^n_{++}.
\end{aligned}
\tag{14.59}
$$

Let $\bar{\mathbf{X}}$ be an optimal solution to (14.59) (it can be shown that such an $\bar{\mathbf{X}}$ exists). Write down the KKT conditions that $\bar{\mathbf{X}}$ must satisfy.

3. Suppose that $m = n$ and $\mathbf{v^i} = \mathbf{e_i}$ for $i = 1, \ldots, n$, where $\mathbf{e_i}$ is the i–th standard basis vector. Using the above result, determine the optimal solution to (14.59) and find the corresponding Lagrange multipliers.

Exercise 14.4.9. Let $\mathbf{a} \in \mathbb{R}^n$, $b \in \mathbb{R}$ and $\mathbf{c} \in \mathbb{R}^n$ be such that $\mathbf{a}, \mathbf{c} > \mathbf{0}$ and $b > 0$. Consider the following problem:

$$
\begin{aligned}
\text{minimize} \quad & \sum_{i=1}^{n} \frac{c_i}{x_i} \\
\text{subject to} \quad & \sum_{i=1}^{n} a_i x_i = b, \\
& \mathbf{x} \geq \mathbf{0}.
\end{aligned}
\tag{14.60}
$$

1. Let $u_1 \in \mathbb{R}$ and $\mathbf{u_2} \in \mathbb{R}^n$ be the Lagrange multipliers associated with the equality and inequality constraints, respectively. Write down the KKT conditions associated with (14.60).

2. Give explicit expressions for $\bar{\mathbf{x}} \in \mathbb{R}^n$, $\bar{u}_1 \in \mathbb{R}$ and $\bar{\mathbf{u}}_2 \in \mathbb{R}^n$ such that $(\bar{\mathbf{x}}, \bar{u}_1, \bar{\mathbf{u}}_2)$ satisfies the KKT conditions above.

3. Is the solution $\bar{\mathbf{x}} \in \mathbb{R}^n$ found above an optimal solution to (14.60)? Explain.

References

[1] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear Programming: Theory and Algorithms*, 2nd ed., ser. Wiley-Interscience Series in Discrete Mathematics and Optimization.   New York: John Wiley & Sons, Inc., 1993.

[2] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed.   Belmont, Massachusetts: Athena Scientific, 1999.

[3] S. Boyd and L. Vandenberghe, *Convex Optimization*.   Cambridge: Cambridge University Press, 2004, available online at http://www.stanford.edu/~boyd/cvxbook/.

[4] D. G. Luenberger and Y. Ye, *Linear and Nonlinear Programming*, 3rd ed., ser. International Series in Operations Research and Management Science.   New York: Springer Science+Business Media, LLC, 2008, vol. 116.

[5] R. A. Horn and C. R. Johnson, *Matrix Analysis*.   Cambridge: Cambridge University Press, 1985.

[6] M. Brookes, "The Matrix Reference Manual," 2005, available online at http://www.ee.ic.ac.uk/hp/staff/dmb/matrix/intro.html.

[7] M. Sion, "On General Minimax Theorems," *Pacific Journal of Mathematics*, vol. 8, no. 1, pp. 171–176, 1958.

[8] H. Komiya, "Elementary Proof for Sion's Minimax Theorem," *Kodai Mathematical Journal*, vol. 11, no. 1, pp. 5–7, 1988.

[9] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-constrained modulation optimization," *IEEE Transactions on Wireless Communications*, vol. 4, no. 5, pp. 2349–2360, September 2005.

[10] S. Cui, M. Kisialiou, Z.-Q. Luo, and Z. Ding, "Robust blind multiuser detection against signature waveform mismatch based on second order cone programming," *IEEE Transactions on Wireless Communications*, vol. 4, no. 4, pp. 1285–1291, July 2005.

[11] R. Zhang and Y.-C. Liang, "Exploiting multi-antennas for opportunistic spectrum sharing in cognitive radio networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, no. 1, pp. 88–102, Feb. 2008.

[12] W. Yu and R. Lui, "Dual methods for nonconvex spectrum optimization of multicarrier systems," *IEEE Transactions on Communications*, vol. 54, no. 7, pp. 1310–1322, July 2006.

[13] L. Zhang, R. Zhang, Y.-C. Liang, Y. Xin, and S. Cui, "On the relationship between the multi-antenna secrecy communications and cognitive radio communications," *IEEE Transactions on Communications*, vol. 58, no. 6, pp. 1877–1886, June 2010.

[14] R. Zhang, S. Cui, and Y.-C. Liang, "On ergodic sum capacity of fading cognitive multiple-access and broadcast channels," *IEEE Transactions on Information Theory*, vol. 55, no. 11, pp. 5161–5178, November 2009.

[15] M. Chiang, "Balancing transport and physical layers in wireless multihop networks: Jointly optimal congestion control and power control," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 1, pp. 104–116, January 2005.

[16] J. Xiao, S. Cui, Z. Q. Luo, and A. J. Goldsmith, "Joint estimation in sensor networks under energy constraint," *IEEE Transactions on Signal Processing*, vol. 54, no. 2, pp. 413–422, February 2005.

[17] A. So and Y. Ye, "Theory of semidefinite programming for sensor network localization," *Mathematical Programming*, vol. 109, pp. 367–384, 2007.

[18] S. Cui and A. J. Goldsmith, "Cross-layer design in energy-constrained networks using cooperative MIMO techniques," *EURASIP's Signal Processing Journal, Special Issue on Advances in Signal Processing-based Cross-layer Designs*, vol. 86, pp. 1804–1814, August 2006.

[19] R. Madan, S. Cui, S. Lall, and A. Goldsmith, "Modeling and optimization of transmission schemes in energy-constrained wireless sensor networks," *IEEE/ACM Transactions on Networking*, vol. 15, no. 6, pp. 1359–1372, December 2007.

[20] Z. Quan, S. Cui, H. V. Poor, and A. Sayed, "Collaborative wideband sensing for cognitive radios," *IEEE Signal Processing Magazine special issue on cognitive radios*, vol. 25, no. 6, pp. 60–73, January 2009.

[21] R. Zhang, Y.-C. Liang, C.-C. Chai, and S. Cui, "Optimal beamforming for two-way multi-antenna relay channel with analogue network coding," *IEEE Journal on Selected Areas of Communications*, vol. 27, no. 6, pp. 699–712, June 2009.

[22] R. Zhang and S. Cui, "Cooperative interference management with miso beam-forming," *IEEE Transactions on Signal Processing*, vol. 58, no. 10, pp. 5450 –5458, October 2010.

[23] S. P. Boyd, S.-J. Kim, D. D. Patil, and M. A. Horowitz, "Digital Circuit Optimization via Geometric Programming," *Operations Research*, vol. 53, no. 6, pp. 899–932, November 2005.

[24] T. Cover and J. Thomas, *Elements of Information Theory.* New York: Wiley, 1991.

[25] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, November 1999.

[26] R. G. Bland, D. Goldfarb, and M. J. Todd, "The ellipsoid method: a survey," *Operations Research*, vol. 29, no. 6, pp. 1039–1091, November 1981.

[27] Z.-Q. Luo and W. Yu, "An introduction to convex optimization for communications and signal processing," *IEEE Journal of Selected Topics in Signal Processing*, vol. 24, no. 8, pp. 1426–1438, August 2006.

[28] Y. Ye and S. Zhang, "New results on quadratic minimization," *SIAM J. Optim.*, vol. 14, pp. 245–267, 2003.