Semidefinite Programming

Yinyu Ye

December 2002
Preface

This is a monograph for MS&E 314, “Semidefinite Programming”, which I am teaching at Stanford.

Information, supporting materials, and computer programs related to this book may be found at the following address on the World-Wide Web:

http://www.stanford.edu/class/msande314

Please report any question, comment and error to the address:

yinyu-ye@stanford.edu

YINYU YE

Stanford, 2002
Chapter 1

Introduction and Preliminaries

1.1 Introduction

Semidefinite Programming, hereafter SDP, is a natural extension of Linear programming (LP) that is a central decision model in Management Science and Operations Research. LP plays an extremely important role in the theory and application of Optimization. In one sense it is a continuous optimization problem in minimizing a linear objective function over a convex polyhedron; but it is also a combinatorial problem involving selecting an extreme point among a finite set of possible vertices. Businesses, large and small, use linear programming models to optimize communication systems, to schedule transportation networks, to control inventories, to adjust investments, and to maximize productivity.

In LP, the variables form a vector which is required to be nonnegative, where in SDP they are components of a matrix and it is constrained to be positive semidefinite. Both of them may have linear equality constraints as well. One thing in common is that interior-point algorithms developed in past two decades for LP are naturally applied to solving SDP.

Interior-point algorithms are continuous iterative algorithms. Computation experience with sophisticated procedures suggests that the number of iterations necessarily grows much more slowly than the dimension grows. Furthermore, they have an established worst-case polynomial iteration bound, providing the potential for dramatic improvement in computation effectiveness.

The goal of the monograph is to provide a text book for teaching
Semidefinite Programming, a modern Linear Programming decision model and its applications in other scientific and engineering fields.

The monograph is organized as follows. In Chapter 1, we discuss some necessary mathematical preliminaries. We also present several decision and optimization problems and several basic numerical procedures used throughout the text.

Chapter 2 is devoted to studying the theories and geometries of matrix inequalities, convexity, and semidefinite programming. Almost all interior-point methods exploit the rich geometric properties of linear and matrix inequalities, such as “center,” “volume,” “potential,” etc. These geometries are also helpful for teaching, learning, and research.

Chapter 3 is focused on interior-point algorithms. Here, we select two types algorithms: the path-following algorithm and the potential reduction algorithm. Each algorithm has three forms, the primal, the dual and the primal-dual form. We analyze the worst-case complexity bound for them, where we will use the real number computation model in our analysis because of the continuous nature of interior-point algorithms. We also compare the complexity theory with the convergence rate used in numerical analysis.

Not only has the convergence speed of SDP algorithms been significantly improved during the last decade, but also the problem domain applicable by SDP has dramatically widened. Chapters 4, 5, 6, and 7 would describe some of SDP applications and new established results in Engineering, Combinatory Optimization, Robust Optimization and Quantum Computation.

Finally, we discuss major computational issues in Chapter 8. We discuss several effective implementation techniques frequently used in interior-point SDP software, such as the sparse linear system, the predictor and corrector step, and the homogeneous and self-dual formulation. We also present major difficulties and challenges faced by SDP.

### 1.2 Mathematical Preliminaries

This section summarizes mathematical background material for linear algebra, linear programming, and nonlinear optimization.

#### 1.2.1 Basic notations

The notation described below will be followed in general. There may be some deviation where appropriate.

By \( \mathcal{R} \) we denote the set of real numbers. \( \mathcal{R}_+ \) denotes the set of non-negative real numbers, and \( \mathcal{R}_+^\circ \) denotes the set of positive numbers. For
1.2. MATHEMATICAL PRELIMINARIES

a natural number $n$, the symbol $\mathcal{R}^n$ ($\mathcal{R}^n_+$, $\mathcal{R}^n_0$) denotes the set of vectors with $n$ components in $\mathcal{R}$ ($\mathcal{R}_+$, $\mathcal{R}_0$). We call $\mathcal{R}^n_0$ the interior of $\mathcal{R}^n_+$. The vector inequality $x \geq y$ means $x_j \geq y_j$ for $j = 1, 2, \ldots, n$. Zero represents a vector whose entries are all zeros and $e$ represents a vector whose entries are all ones, where their dimensions may vary according to other vectors in expressions. A vector is always considered as a column vector, unless otherwise stated. Upper-case letters will be used to represent matrices. Greek letters will typically be used to represent scalars. For convenience, we sometime write a column vector $x$ as

$$x = (x_1; x_2; \ldots; x_n)$$

and a row vector as

$$x = (x_1, x_2, \ldots, x_n).$$

Addition of matrices and multiplication of matrices with scalars are standard. The superscript “$T$” denotes transpose operation. The inner product in $\mathcal{R}^n$ is defined as follows:

$$\langle x, y \rangle := x^T y = \sum_{j=1}^{n} x_j y_j \quad \text{for} \quad x, y \in \mathcal{R}^n.$$  

The $l_2$ norm of a vector $x$ is given by

$$\|x\|_2 = \sqrt{x^T x},$$

and the $l_\infty$ norm is

$$\|x\|_\infty = \max\{|x_1|, |x_2|, \ldots, |x_n|\}.$$  

In general, the $p$ norm is

$$\|x\|_p = \left( \sum_{j=1}^{n} |x_j|^p \right)^{1/p}, \quad p = 1, 2, \ldots$$

The dual of the $p$ norm, denoted by $\|\cdot\|^*$, is the $q$ norm, where

$$\frac{1}{p} + \frac{1}{q} = 1.$$  

In this book, $\|\cdot\|$ generally represents the $l_2$ norm.

For natural numbers $m$ and $n$, $\mathcal{R}^{m \times n}$ denotes the set of real matrices with $m$ rows and $n$ columns. For $A \in \mathcal{R}^{m \times n}$, we assume that the row index
set of $A$ is $\{1, 2, ..., m\}$ and the column index set is $\{1, 2, ..., n\}$. The $i$th row of $A$ is denoted by $a_i$, and the $j$th column of $A$ is denoted by $a_{ij}$. If $I$ is a subset of the row index set and $J$ is a subset of the column index set, then $A_I$ denotes the submatrix of $A$ whose rows belong to $I$, $A_J$ denotes the submatrix of $A$ whose columns belong to $J$, $A_{IJ}$ denotes the submatrix of $A$ induced by those components of $A$ whose indices belong to $I$ and $J$, respectively.

The identity matrix is denoted by $I$. The null space of $A$ is denoted $\mathcal{N}(A)$ and the range of $A$ is $\mathcal{R}(A)$. The determinant of an $n \times n$-matrix $A$ is denoted by $\det(A)$. The trace of $A$, denoted by $\text{tr}(A)$, is the sum of the diagonal entries in $A$. The operator norm of $A$, denoted by $\|A\|$, is

$$
\|A\|^2 := \max_{0 \neq x \in \mathbb{R}^n} \frac{\|Ax\|^2}{\|x\|^2}.
$$

For a vector $x \in \mathbb{R}^n$, $D_x$ represents a diagonal matrix in $\mathbb{R}^{n \times n}$ whose diagonal entries are the entries of $x$, i.e.,

$$
D_x = \text{diag}(x).
$$

A matrix $Q \in \mathbb{R}^{n \times n}$ is said to be positive definite (PD), denoted by $Q \succ 0$, if

$$
x^T Q x > 0, \quad \text{for all } x \neq 0,
$$

and positive semi-definite (PSD), denoted by $Q \succeq 0$, if

$$
x^T Q x \geq 0, \quad \text{for all } x.
$$

If $Q \succ 0$, then $-Q$ is called negative definite (ND), denoted by $Q \prec 0$; if $Q \succeq 0$, then $-Q$ is called negative semi-definite (NSD), denoted by $Q \preceq 0$. If $Q$ is symmetric, then its eigenvalues are all real numbers; furthermore, $Q$ is PSD if and only if all its eigenvalues are non-negative, and $Q$ is PD if and only if all its eigenvalue are positive. Given a PD matrix $Q$ we can define a $Q$-norm, $\|\cdot\|_Q$, for vector $x$ as

$$
\|x\|_Q = \sqrt{x^T Q x}.
$$

$\mathcal{M}^n$ denotes the space of symmetric matrices in $\mathbb{R}^{n \times n}$. The inner product in $\mathcal{M}^n$ is defined as follows:

$$
\langle X, Y \rangle := X \bullet Y = \text{tr} X^T Y = \sum_{i,j} X_{i,j} Y_{i,j} \quad \text{for } X, Y \in \mathcal{M}^n.
$$
This is a generalization of the vector inner product to matrices. The matrix norm associated with the inner product is called Frobenius norm:

\[ \|X\|_f = \sqrt{\text{tr}X^TX} . \]

\( \mathcal{M}_n^+ \) denote the set of positive semi-definite matrices in \( \mathcal{M}^n \). \( \mathcal{M}_n^{\circ} \) denotes the set of positive definite matrices in \( \mathcal{M}^n \). We call \( \mathcal{M}_n^{\circ} \) the interior of \( \mathcal{M}_n^+ \).

\( \{x^k\}_0^\infty \) is an ordered sequence \( x^0, x^1, x^2, \ldots, x^k, \ldots \). A sequence \( \{x^k\}_0^\infty \) is convergent to \( \bar{x} \), denoted \( x^k \to \bar{x} \), if

\[ \|x^k - \bar{x}\| \to 0. \]

A point \( x \) is a limit point of \( \{x^k\}_0^\infty \) if there is a subsequence of \( \{x^k\} \) convergent to \( x \).

If \( g(x) \geq 0 \) is a real valued function of a real nonnegative variable, the notation \( g(x) = O(x) \) means that \( g(x) \leq \bar{c}x \) for some constant \( \bar{c} \); the notation \( g(x) = \Omega(x) \) means that \( g(x) \geq c_\bar{x} \) for some constant \( c_\bar{x} \); the notation \( g(x) = \theta(x) \) means that \( c_\bar{x}x \leq g(x) \leq \bar{c}x \). Another notation is \( g(x) = o(x) \), which means that \( g(x) \) goes to zero faster than \( x \) does:

\[ \lim_{x \to 0} \frac{g(x)}{x} = 0. \]

### 1.2.2 Convex sets

If \( x \) is a member of the set \( \Omega \), we write \( x \in \Omega \); if \( y \) is not a member of \( \Omega \), we write \( y \notin \Omega \). The union of two sets \( S \) and \( T \) is denoted \( S \cup T \); the intersection of them is denoted \( S \cap T \). A set can be specified in the form \( \Omega = \{x : P(x)\} \) as the set of all elements satisfying property \( P \).

For \( y \in \mathbb{R}^n \) and \( \epsilon > 0 \), \( B(y, \epsilon) = \{x : \|x - y\| \leq \epsilon\} \) is the ball or sphere of radius \( \epsilon \) with center \( y \). In addition, for a positive definite matrix \( Q \) of dimension \( n \), \( E(y, Q) = \{x : (x - y)^TQ(x - y) \leq 1\} \) is called an ellipsoid. The vector \( y \) is the center of \( E(y, Q) \).

A set \( \Omega \) is closed if \( x^k \to x \), where \( x^k \in \Omega \), implies \( x \in \Omega \). A set \( \Omega \) is open if around every point \( y \in \Omega \) there is a ball that is contained in \( \Omega \), i.e., there is an \( \epsilon > 0 \) such that \( B(y, \epsilon) \subset \Omega \). A set is bounded if it is contained within a ball with finite radius. A set is compact if it is both closed and bounded. The (topological) interior of any set \( \Omega \), denoted \( \hat{\Omega} \), is the set of points in \( \Omega \) which are the centers of some balls contained in \( \Omega \). The closure of \( \Omega \), denoted \( \tilde{\Omega} \), is the smallest closed set containing \( \Omega \). The boundary of \( \Omega \) is the part of \( \hat{\Omega} \) that is not in \( \tilde{\Omega} \).
A set $C$ is said to be convex if for every $x^1, x^2 \in C$ and every real number $\alpha$, $0 < \alpha < 1$, the point $\alpha x^1 + (1 - \alpha)x^2 \in C$. The convex hull of a set $\Omega$ is the intersection of all convex sets containing $\Omega$.

A set $C$ is a cone if $x \in C$ implies $\alpha x \in C$ for all $\alpha > 0$. A cone that is also convex is a convex cone. For a cone $C \subset \mathcal{E}$, the dual of $C$ is the cone

$$C^* := \{ y : \langle x, y \rangle \geq 0 \text{ for all } x \in C \},$$

where $\langle \cdot, \cdot \rangle$ is an inner product operation for space $\mathcal{E}$.

**Example 1.1** The $n$-dimensional non-negative orthant, $\mathbb{R}^n_+ = \{ x \in \mathbb{R}^n : x \geq 0 \}$, is a convex cone. The dual of the cone is also $\mathbb{R}^n_+$; it is self-dual.

**Example 1.2** The set of all positive semi-definite matrices in $\mathcal{M}^n$, $\mathcal{M}^n_+$, is a convex cone, called the positive semi-definite matrix cone. The dual of the cone is also $\mathcal{M}^n_+$; it is self-dual.

**Example 1.3** The set $\{(t; x) \in \mathbb{R}^{n+1} : t \geq \|x\|\}$ is a convex cone in $\mathbb{R}^{n+1}$, called the second-order cone. The dual of the cone is also the second-order cone in $\mathbb{R}^{n+1}$; it is self-dual.

A cone $C$ is (convex) polyhedral if $C$ can be represented by

$$C = \{ x : Ax \leq 0 \}$$

for some matrix $A$ (Figure 1.1).

**Figure 1.1:** Polyhedral and nonpolyhedral cones.
Example 1.4 The non-negative orthant is a polyhedral cone, and neither the positive semi-definite matrix cone nor the second-order cone is polyhedral.

The most important type of convex set is a hyperplane. Hyperplanes dominate the entire theory of optimization. Let $a$ be a nonzero $n$-dimensional vector, and let $b$ be a real number. The set

$$H = \{ x \in \mathbb{R}^n : a^T x = b \}$$

is a hyperplane in $\mathbb{R}^n$ (Figure 1.2). Relating to hyperplane, positive and negative closed half spaces are given by

$$H_+ = \{ x : a^T x \geq b \}$$
$$H_- = \{ x : a^T x \leq b \}.$$

A set which can be expressed as the intersection of a finite number of closed half spaces is said to be a convex polyhedron:

$$P = \{ x : Ax \leq b \}.$$

A bounded polyhedron is called polytope.

Let $P$ be a polyhedron in $\mathbb{R}^n$, $F$ is a face of $P$ if and only if there is a vector $c$ for which $F$ is the set of points attaining max $\{ c^T x : x \in P \}$ provided the this maximum is finite. A polyhedron has only finite many faces; each face is a nonempty polyhedron.

The most important theorem about the convex set is the following separating theorem (Figure 1.3).
Theorem 1.1 (Separating hyperplane theorem) Let $C \subset \mathcal{E}$, where $\mathcal{E}$ is either $\mathbb{R}^n$ or $\mathcal{M}^n$, be a closed convex set and let $y$ be a point exterior to $C$. Then there is a vector $a \in \mathcal{E}$ such that

$$\langle a, y \rangle < \inf_{x \in C} \langle a, x \rangle.$$ 

The geometric interpretation of the theorem is that, given a convex set $C$ and a point $y$ outside of $C$, there is a hyperplane containing $y$ that contains $C$ in one of its open half spaces.

Example 1.5 Let $C$ be a unit circle centered at the point $(1; 1)$. That is, $C = \{x \in \mathbb{R}^2 : (x_1 - 1)^2 + (x_2 - 1)^2 \leq 1\}$. If $y = (2; 0)$, $a = (-1; 1)$ is a separating hyperplane vector. If $y = (0; -1)$, $a = (0; 1)$ is a separating hyperplane vector. It is worth noting that these separating hyperplanes are not unique.

We use the notation $\mathcal{E}$ to represent either $\mathbb{R}^n$ or $\mathcal{M}^n$, depending on the context, throughout this book, because all our decision and optimization problems take variables from one or both of these two vector spaces.

1.2.3 Real functions

The real function $f(x)$ is said to be continuous at $x$ if $x_k \to x$ implies $f(x_k) \to f(x)$. The real function $f(x)$ is said to be continuous on set $\Omega \subset \mathcal{E}$, where recall that $\mathcal{E}$ is either $\mathbb{R}^n$ or $\mathcal{M}^n$, if $f(x)$ is continuous at $x$ for every $x \in \Omega$. 
A function \( f(x) \) is called homogeneous of degree \( k \) if \( f(\alpha x) = \alpha^k f(x) \) for all \( \alpha \geq 0 \).

**Example 1.6** Let \( c \in \mathbb{R}^n \) be given and \( x \in \mathbb{R}^n_+ \). Then \( c^T x \) is homogeneous of degree \( 1 \) and

\[
\mathcal{P}(x) = n \log(c^T x) - \sum_{j=1}^n \log x_j
\]

is homogeneous of degree \( 0 \), where \( \log \) is the natural logarithmic function.

Let \( C \in \mathcal{M}^n \) be given and \( X \in \mathcal{M}_+^n \). Then \( x^T C x \) is homogeneous of degree \( 2 \), \( C \cdot X \) and \( \det(X) \) are homogeneous of degree \( 1 \), and

\[
\mathcal{P}(X) = n \log(C \cdot X) - \log \det(X)
\]

is homogeneous of degree \( 0 \).

A set of real-valued function \( f_1, f_2, ..., f_m \) defined on \( \mathcal{E} \) can be written as a single vector function \( f = (f_1, f_2, ..., f_m)^T \in \mathbb{R}^m \). If \( f \) has continuous partial derivatives of order \( p \), we say \( f \in C^p \). The gradient vector or matrix of a real-valued function \( f \in C^1 \) is a vector or matrix

\[
\nabla f(x) = \{ \partial f/\partial x_{ij} \}, \quad \text{for } i, j = 1, ..., n.
\]

If \( f \in C^2 \), we define the Hessian of \( f \) to be the \( n^2 \)-dimensional symmetric matrix

\[
\nabla^2 f(x) = \left\{ \frac{\partial^2 f}{\partial x_{ij} \partial x_{kl}} \right\}, \quad \text{for } i, j, k, l = 1, ..., n.
\]

If \( f = (f_1, f_2, ..., f_m)^T \in \mathbb{R}^m \), the Jacobian matrix of \( f \) is

\[
\nabla f(x) = \begin{pmatrix}
\nabla f_1(x) \\
... \\
\nabla f_m(x)
\end{pmatrix}.
\]

\( f \) is a (continuous) convex function if and only if for \( 0 \leq \alpha \leq 1 \),

\[
f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y).
\]

\( f \) is a (continuous) quasi-convex function if and only if for \( 0 \leq \alpha \leq 1 \),

\[
f(\alpha x + (1-\alpha)y) \leq \max[f(x), f(y)].
\]

Thus, a convex function is a quasi-convex function. The level set of \( f \) is given by

\[
L(z) = \{ x : f(x) \leq z \}.
\]
f is a quasi-convex function implies that the level set of f is convex for any given \( z \) (see Exercise 1.9).

A group of results that are used frequently in analysis are under the heading of Taylor’s theorem or the mean-value theorem. The theorem establishes the linear and quadratic approximations of a function.

**Theorem 1.2 (Taylor expansion)** Let \( f \in C^1 \) be in a region containing the line segment \([x, y]\). Then there is a \( \alpha, 0 \leq \alpha \leq 1 \), such that

\[
f(y) = f(x) + \nabla f(\alpha x + (1 - \alpha)y)(y - x).
\]

Furthermore, if \( f \in C^2 \) then there is a \( \alpha, 0 \leq \alpha \leq 1 \), such that

\[
f(y) = f(x) + \nabla f(x)(y - x) + (1/2)(y - x)^T \nabla^2 f(\alpha x + (1 - \alpha)y)(y - x).
\]

We also have several propositions for real functions. The first indicates that the linear approximation of a convex function is a under-estimate.

**Proposition 1.3** Let \( f \in C^1 \). Then \( f \) is convex over a convex set \( \Omega \) if and only if

\[
f(y) \geq f(x) + \nabla f(x)(y - x)
\]

for all \( x, y \in \Omega \).

The following proposition states that the Hessian of a convex function is positive semi-definite.

**Proposition 1.4** Let \( f \in C^2 \). Then \( f \) is convex over a convex set \( \Omega \) if and only if the Hessian matrix of \( f \) is positive semi-definite throughout \( \Omega \).

### 1.2.4 Inequalities

There are several important inequalities that are frequently used in algorithm design and complexity analysis.

- **Cauchy-Schwarz**: given \( x, y \in \mathbb{R}^n \), then
  \[
x^T y \leq \|x\|\|y\|.
\]

- **Arithmetic-geometric mean**: given \( x \in \mathbb{R}^n_+ \),
  \[
  \frac{\sum x_j}{n} \geq \left(\prod x_j\right)^{1/n}.
\]

- **Harmonic**: given \( x \in \mathbb{R}^n_+ \),
  \[
  \left(\sum x_j\right)\left(\sum 1/x_j\right) \geq n^2.
\]

- **Hadamard**: given \( A \in \mathbb{R}^{m \times n} \) with columns \( a_1, a_2, \ldots, a_n \), then
  \[
  \sqrt{\det(A^TA)} \leq \prod \|a_j\|.
  \]
1.3 Some Basic Decision and Optimization Problems

A decision or optimization problem has a form that is usually characterized by the decision variables and the constraints. A problem, \( \mathcal{P} \), consists of two sets, data set \( \mathcal{Z}_p \) and solution set \( \mathcal{S}_p \). In general, \( \mathcal{S}_p \) can be implicitly defined by the so-called optimality conditions. The solution set may be empty, i.e., problem \( \mathcal{P} \) may have no solution.

**Theorem 1.5** Weierstrass theorem A continuous function \( f \) defined on a compact set (bounded and closed) \( \Omega \subset \mathcal{E} \) has a minimizer in \( \Omega \); that is, there is an \( x^* \in \Omega \) such that for all \( x \in \Omega \), \( f(x) \geq f(x^*) \).

In what follows, we list several decision and optimization problems. More problems will be listed later when we address them.

### 1.3.1 System of linear equations

Given \( A \in \mathbb{R}^{m \times n} \) and \( b \in \mathbb{R}^m \), the problem is to solve \( m \) linear equations for \( n \) unknowns:

\[
Ax = b.
\]

The data and solution sets are

\[
\mathcal{Z}_p = \{ A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m \} \quad \text{and} \quad \mathcal{S}_p = \{ x \in \mathbb{R}^n : Ax = b \}.
\]

\( \mathcal{S}_p \) in this case is an affine set. Given an \( x \), one can easily check to see if \( x \) is in \( \mathcal{S}_p \) by a matrix-vector multiplication and a vector-vector comparison. We say that a solution of this problem is easy to recognize.

To highlight the analogy with the theories of linear inequalities and linear programming, we list several well-known results of linear algebra. The first theorem provides two basic representations, the null and row spaces, of a linear subspaces.

**Theorem 1.6** Each linear subspace of \( \mathbb{R}^n \) is generated by finitely many vectors, and is also the intersection of finitely many linear hyperplanes; that is, for each linear subspace of \( L \) of \( \mathbb{R}^n \) there are matrices \( A \) and \( C \) such that \( L = \mathcal{N}(A) = \mathcal{R}(C) \).

The following theorem was observed by Gauss. It is sometimes called the fundamental theorem of linear algebra. It gives an example of a characterization in terms of necessary and sufficient conditions, where necessity is straightforward, and sufficiency is the key of the characterization.
Theorem 1.7 Let $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. The system \( \{ x : Ax = b \} \) has a solution if and only if that $A^Ty = 0$ implies $b^Ty = 0$.

A vector $y$, with $A^Ty = 0$ and $b^Ty = 1$, is called an infeasibility certificate for the system \( \{ x : Ax = b \} \).

Example 1.7 Let $A = (1; -1)$ and $b = (1; 1)$. Then, $y = (1/2; 1/2)$ is an infeasibility certificate for \( \{ x : Ax = b \} \).

1.3.2 Linear least-squares problem

Given $A \in \mathbb{R}^{m \times n}$ and $c \in \mathbb{R}^n$, the system of equations $A^Ty = c$ may be over-determined or have no solution. Such a case usually occurs when the number of equations is greater than the number of variables. Then, the problem is to find an $y \in \mathbb{R}^m$ or $s \in \mathbb{R}(A^T)$ such that $\|A^Ty - c\|$ or $\|s - c\|$ is minimized. We can write the problem in the following format:

\[
(\text{LS}) \quad \text{minimize} \quad \|A^Ty - c\|^2 \\
\text{subject to} \quad y \in \mathbb{R}^m,
\]

or

\[
(\text{LS}) \quad \text{minimize} \quad \|s - c\|^2 \\
\text{subject to} \quad s \in \mathbb{R}(A^T).
\]

In the former format, the term $\|A^Ty - c\|^2$ is called the objective function, $y$ is called the decision variable. Since $y$ can be any point in $\mathbb{R}^m$, we say this (optimization) problem is unconstrained. The data and solution sets are

\[\mathbb{Z}_p = \{ A \in \mathbb{R}^{m \times n}, c \in \mathbb{R}^n \}\]

and

\[\mathbb{S}_p = \{ y \in \mathbb{R}^m : \|A^Ty - c\|^2 \leq \|A^Tx - c\|^2 \text{ for every } x \in \mathbb{R}^m \}\]

Given a $y$, to see if $y \in \mathbb{S}_p$ is as the same as the original problem. However, from a projection theorem in linear algebra, the solution set can be characterized and represented as

\[\mathbb{S}_p = \{ y \in \mathbb{R}^m : AA^Ty = Ac \},\]

which becomes a system of linear equations and always has a solution. The vector $s = A^Ty = A^T(AA^T)^+Ac$ is the projection of $c$ onto the range of $A^T$, where $AA^T$ is called normal matrix and $(AA^T)^+$ is called pseudoinverse. If $A$ has full row rank then $(AA^T)^+ = (AA^T)^{-1}$, the standard inverse of full rank matrix $AA^T$. If $A$ is not of full rank, neither is $AA^T$ and $(AA^T)^+AA^T x = x$ only for $x \in \mathbb{R}(A^T)$. 

1.3. SOME BASIC DECISION AND OPTIMIZATION PROBLEMS

The vector $c - A^T y = (I - A^T (AA^T)^+ A) c$ is the projection of $c$ onto the null space of $A$. It is the solution of the following least-squares problem:

\[
(\text{LS}) \quad \text{minimize} \quad \|x - c\|^2 \\
\text{subject to} \quad x \in \mathcal{N}(A).
\]

In the full rank case, both matrices $A^T (AA^T)^{-1} A$ and $I - A^T (AA^T)^{-1} A$ are called projection matrices. These symmetric matrices have several desired properties (see Exercise 1.15).

1.3.3 System of linear inequalities

Given $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$, the problem is to find a solution $x \in \mathbb{R}^n$ satisfying $Ax \leq b$ or prove that the solution set is empty. The inequality problem includes other forms such as finding an $x$ that satisfies the combination of linear equations $Ax = b$ and inequalities $x \geq 0$. The data and solution sets of the latter are

\[
Z_p = \{ A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m \} \quad \text{and} \quad S_p = \{ x \in \mathbb{R}^n : Ax = b, x \geq 0 \}.
\]

Traditionally, a point in $S_p$ is called a feasible solution, and a strictly positive point in $S_p$ is called a strictly feasible or interior feasible solution.

The following results are Farkas’ lemma and its variants.

**Theorem 1.8** (Farkas’ lemma) Let $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. Then, the system \( \{x : Ax = b, x \geq 0 \} \) has a feasible solution $x$ if and only if that $A^T y \leq 0$ implies $b^T y \leq 0$.

A vector $y$, with $A^T y \leq 0$ and $b^T y = 1$, is called a (primal) infeasibility certificate for the system \( \{x : Ax = b, x \geq 0 \} \). Geometrically, Farkas’ lemma means that if a vector $b \in \mathbb{R}^m$ does not belong to the cone generated by $a_1, \ldots, a_n$, then there is a hyperplane separating $b$ from cone($a_1, \ldots, a_n$).

**Example 1.8** Let $A = (1, 1)$ and $b = -1$. Then, $y = -1$ is an infeasibility certificate for \( \{x : Ax = b, x \geq 0 \} \).

**Theorem 1.9** (Farkas’ lemma variant) Let $A \in \mathbb{R}^{m \times n}$ and $c \in \mathbb{R}^n$. Then, the system \( \{y : A^T y \leq c\} \) has a solution $y$ if and only if that $Ax = 0$ and $x \geq 0$ imply $c^T x \geq 0$.

Again, a vector $x \geq 0$, with $Ax = 0$ and $c^T x = -1$, is called a (dual) infeasibility certificate for the system \( \{y : A^T y \leq c\} \).
Chapter 1. Introduction and Preliminaries

Example 1.9 Let $A = (1; -1)$ and $c = (1; -2)$. Then, $x = (1; 1)$ is an infeasibility certificate for $\{ y : A^T y \leq c \}$.

We say $\{ x : Ax = b, x \geq 0 \}$ or $\{ y : A^T y \leq c \}$ is approximately feasible in the sense that we have an approximate solution to the equations and inequalities. In this case we can show that any certificate proving their infeasibility must have large norm. Conversely, if $\{ x : Ax = b, x \geq 0 \}$ or $\{ y : A^T y \leq c \}$ is “approximately infeasible” in the sense that we have an approximate certificate in Farkas’ lemma, then any feasible solution must have large norm.

Example 1.10 Given $\epsilon > 0$ but small. Let $A = (1, 1)$ and $b = -\epsilon$. Then, $x = (0; 0)$ is approximately feasible for $\{ x : Ax = b, x \geq 0 \}$, and the infeasibility certificate $y = -1/\epsilon$ has a large norm.

Let $A = (1; -1)$ and $c = (1; -1 - \epsilon)$. Then, $y = 1$ is approximately feasible for $\{ y : A^T y \leq c \}$, and the infeasibility certificate $x = (1/\epsilon; 1/\epsilon)$ has a large norm.

1.3.4 Linear programming (LP)

Given $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ and $c, l, u \in \mathbb{R}^n$, the linear programming (LP) problem is the following optimization problem:

$$
\begin{align*}
\text{minimize} & \quad c^T x \\
\text{subject to} & \quad Ax = b, \ l \leq x \leq u,
\end{align*}
$$

where some elements in $l$ may be $-\infty$ meaning that the associated variables are unbounded from below, and some elements in $u$ may be $\infty$ meaning that the associated variables are unbounded from above. If a variable is unbounded either from below or above, then it is called a “free” variable.

The standard form linear programming problem is given below, which we will use throughout this book:

$$
\begin{align*}
(\text{LP}) \quad \text{minimize} & \quad c^T x \\
\text{subject to} & \quad Ax = b, \ x \geq 0.
\end{align*}
$$

The linear function $c^T x$ is called the objective function, and $x$ is called the decision variables. In this problem, $Ax = b$ and $x \geq 0$ enforce constraints on the selection of $x$. The set $\mathcal{F}_b = \{ x : Ax = b, x \geq 0 \}$ is called feasible set or feasible region. A point $x \in \mathcal{F}_b$ is called a feasible point, and a feasible point $x^*$ is called an optimal solution if $c^T x^* \leq c^T x$ for all feasible points $x$. If there is a sequence $\{ x^k \}$ such that $x^k$ is feasible and $c^T x^k \to -\infty$, then (LP) is said to be unbounded.
1.3. SOME BASIC DECISION AND OPTIMIZATION PROBLEMS

The data and solution sets for (LP), respectively, are
\[ Z_p = \{A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m, c \in \mathbb{R}^n\} \]
and
\[ S_p = \{x \in \mathcal{F}_p : c^T x \leq c^T y, \text{ for every } y \in \mathcal{F}_p\}. \]
Again, given an \( x \), to see if \( x \in S_p \) is as difficult as the original problem. However, due to the duality theorem, we can simplify the representation of the solution set significantly.

With every (LP), another linear program, called the dual (LD), is the following problem:
\[
(\text{LD}) \quad \text{maximize} \quad b^T y \\
\text{subject to} \quad A^T y + s = c, \quad s \geq 0,
\]
where \( y \in \mathbb{R}^m \) and \( s \in \mathbb{R}^n \). The components of \( s \) are called dual slacks. Denote by \( \mathcal{F}_d \) the sets of all \( (y, s) \) that are feasible for the dual. We see that (LD) is also a linear programming problem where \( y \) is a “free” vector.

The following theorems give us an important relation between the two problems.

**Theorem 1.10** (Weak duality theorem) Let \( \mathcal{F}_p \) and \( \mathcal{F}_d \) be non-empty. Then,
\[ c^T x \geq b^T y \quad \text{where} \quad x \in \mathcal{F}_p, \ (y, s) \in \mathcal{F}_d. \]
This theorem shows that a feasible solution to either problem yields a bound on the value of the other problem. We call \( c^T x - b^T y \) the duality gap. From this we have important results.

**Theorem 1.11** (Strong duality theorem) Let \( \mathcal{F}_p \) and \( \mathcal{F}_d \) be non-empty. Then, \( x^* \) is optimal for (LP) if and only if the following conditions hold:

i) \( x^* \in \mathcal{F}_p \);

ii) there is \( (y^*, s^*) \in \mathcal{F}_d \);

iii) \( c^T x^* = b^T y^* \).

**Theorem 1.12** (LP duality theorem) If (LP) and (LD) both have feasible solutions then both problems have optimal solutions and the optimal objective values of the objective functions are equal.

If one of (LP) or (LD) has no feasible solution, then the other is either unbounded or has no feasible solution. If one of (LP) or (LD) is unbounded then the other has no feasible solution.
The above theorems show that if a pair of feasible solutions can be found to the primal and dual problems with equal objective values, then these are both optimal. The converse is also true; there is no “gap.” From this condition, the solution set for (LP) and (LD) is

\[
S_p = \left\{ (x, y, s) \in (\mathbb{R}^n_+, \mathbb{R}^m, \mathbb{R}^n_+) : \begin{array}{c}
c^T x - b^T y = 0 \\
Ax = b \\
-A^T y - s = -c
\end{array} \right\}, \quad (1.1)
\]

which is a system of linear inequalities and equations. Now it is easy to verify whether or not a pair \((x, y, s)\) is optimal.

For feasible \(x\) and \((y, s)\), \(x^T s = x^T (c - A^T y) = c^T x - b^T y\) is called the complementarity gap. If \(x^T s = 0\), then we say \(x\) and \(s\) are complementary to each other. Since both \(x\) and \(s\) are nonnegative, \(x^T s = 0\) implies that \(x_j s_j = 0\) for all \(j = 1, \ldots, n\). Thus, one equation plus nonnegativity are transformed into \(n\) equations. Equations in (1.1) become

\[
\begin{align*}
D_x s &= 0 \\
Ax &= b \\
-A^T y - s &= -c.
\end{align*}
\]

This system has total \(2n + m\) unknowns and \(2n + m\) equations including \(n\) nonlinear equations.

The following theorem plays an important role in analyzing LP interior-point algorithms. It give a unique partition of the LP variables in terms of complementarity.

**Theorem 1.13** (Strict complementarity theorem) If (LP) and (LD) both have feasible solutions then both problems have a pair of strictly complementary solutions \(x^* \geq 0\) and \(s^* \geq 0\) meaning

\[X^* s^* = 0 \quad \text{and} \quad x^* + s^* > 0.\]

Moreover, the supports

\[P^* = \{j : x_j^* > 0\} \quad \text{and} \quad Z^* = \{j : s_j^* > 0\}\]

are invariant for all pairs of strictly complementary solutions.

Given (LP) or (LD), the pair of \(P^*\) and \(Z^*\) is called the (strict) complementarity partition. \(\{x : A_P x_P^* = b, x_P^* \geq 0, x_{Z^*} = 0\}\) is called the primal optimal face, and \(\{y : c_{Z^*} - A_{Z^*}^T y \geq 0, c_{P^*} - A_{P^*}^T y = 0\}\) is called the dual optimal face.
Select $m$ linearly independent columns, denoted by the index set $B$, from $A$. Then matrix $A_B$ is nonsingular and we may uniquely solve

$$A_B x_B = b$$

for the $m$-vector $x_B$. By setting the variables, $x_N$, of $x$ corresponding to the remaining columns of $A$ equal to zero, we obtain a solution $x$ such that

$$A x = b.$$ 

Then, $x$ is said to be a (primal) basic solution to (LP) with respect to the basis $A_B$. The components of $x_B$ are called basic variables. A dual vector $y$ satisfying

$$A_B^T y = c_B$$

is said to be the corresponding dual basic solution. If a basic solution $x \geq 0$, then $x$ is called a basic feasible solution. If the dual solution is also feasible, that is

$$s = c - A^T y \geq 0,$$

then $x$ is called an optimal basic solution and $A_B$ an optimal basis. A basic feasible solution is a vertex on the boundary of the feasible region. An optimal basic solution is an optimal vertex of the feasible region.

If one or more components in $x_B$ has value zero, that basic solution $x$ is said to be (primal) degenerate. Note that in a nondegenerate basic solution the basic variables and the basis can be immediately identified from the nonzero components of the basic solution. If all components, $s_N$, in the corresponding dual slack vector $s$, except for $s_B$, are non-zero, then $y$ is said to be (dual) nondegenerate. If both primal and dual basic solutions are nondegenerate, $A_B$ is called a nondegenerate basis.

**Theorem 1.14** (LP fundamental theorem) Given (LP) and (LD) where $A$ has full row rank $m$,

i) if there is a feasible solution, there is a basic feasible solution;

ii) if there is an optimal solution, there is an optimal basic solution.

The above theorem reduces the task of solving a linear program to that searching over basic feasible solutions. By expanding upon this result, the simplex method, a finite search procedure, is derived. The simplex method is to proceed from one basic feasible solution (an extreme point of the feasible region) to an adjacent one, in such a way as to continuously decrease the value of the objective function until a minimizer is reached. In contrast, interior-point algorithms will move in the interior of the feasible region and reduce the value of the objective function, hoping to by-pass many extreme points on the boundary of the region.
1.3.5 Quadratic programming (QP)

Given $Q \in \mathbb{R}^{n \times n}$, $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^{m}$ and $c \in \mathbb{R}^{n}$, the quadratic programming (QP) problem is the following optimization problem:

$$\text{(QP)} \quad \text{minimize} \quad q(x) := \frac{1}{2} x^T Q x + c^T x$$
subject to $Ax = b$, $x \geq 0$.

We may denote the feasible set by $F_p$. The data and solution sets for (QP) are

$$Z_p = \{Q \in \mathbb{R}^{n \times n}, A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^{m}, c \in \mathbb{R}^{n}\}$$

and

$$S_p = \{x \in F_p : q(x) \leq q(y) \quad \text{for every} \quad y \in F_p\}.$$

A feasible point $x^*$ is called a KKT point, where KKT stands for Karush-Kuhn-Tucker, if the following KKT conditions hold: there exists $(y^* \in \mathbb{R}^{m}, s^* \in \mathbb{R}^{n})$ such that $(x^*, y^*, s^*)$ is feasible for the following dual problem:

$$\text{(QD)} \quad \text{maximize} \quad d(x, y) := b^T y - (1/2)x^T Q x$$
subject to $A^T y + s - Q x = c$, $x, s \geq 0$,

and satisfies the complementarity condition

$$(x^*)^T s^* = (1/2)(x^*)^T Q x^* + c^T x^* - (b^T y^* - (1/2)(x^*)^T Q x^* = 0.$$ 

Similar to LP, we can write the KKT condition as:

$$(x, y, s) \in (\mathbb{R}_+^n, \mathbb{R}^m, \mathbb{R}_+^n)$$

and

$$D_x s = 0 \quad Ax = b \quad -A^T y + Q x - s = -c. \quad (1.3)$$

Again, this system has total $2n + m$ unknowns and $2n + m$ equations including $n$ nonlinear equations.

The above condition is also called the first-order necessary condition. If $Q$ is positive semi-definite, then $x^*$ is an optimal solution for (QP) if and only if $x^*$ is a KKT point for (QP). In this case, the solution set for (QP) is characterized by a system of linear inequalities and equations. One can see (LP) is a special case of (QP).
1.4 Algorithms and Computations

An algorithm is a list of instructions to solve a problem. For every instance of problem $\mathcal{P}$, i.e., for every given data $Z \in \mathbb{Z}_p$, an algorithm for solving $\mathcal{P}$ either determines that $\mathcal{S}_p$ is empty or generates an output $x$ such that $x \in \mathcal{S}_p$ or $x$ is close to $\mathcal{S}_p$ in certain measure. The latter $x$ is called an approximate solution.

Let us use $\mathcal{A}_p$ to denote the collection of all possible algorithm for solving every instance in $\mathcal{P}$. Then, the (operation) complexity of an algorithm $A \in \mathcal{A}_p$ for solving an instance $Z \in \mathbb{Z}_p$ is defined as the total arithmetic operations: $+, -, \times, /$, and comparison on real numbers. Denote it by $c_o(A, Z)$. Sometimes it is convenient to define the iteration complexity, denoted by $c_i(A, Z)$, where we assume that each iteration costs a polynomial number (in $m$ and $n$) of arithmetic operations. In most iterative algorithms, each iteration can be performed efficiently both sequentially and in parallel, such as solving a system of linear equations, rank-one updating the inversion of a matrix, pivoting operation of a matrix, multiplying a matrix by a vector, etc.

In the real number model, we introduce $\epsilon$, the error for an approximate solution as a parameter. Let $c(A, Z, \epsilon)$ be the total number of operations of algorithm $A$ for generating an $\epsilon$-approximate solution, with a well-defined measure, to problem $\mathcal{P}$. Then,

$$c(A, \epsilon) := \sup_{Z \in \mathbb{Z}_p} c(A, Z, \epsilon) \leq f_A(m, n, \epsilon) \text{ for any } \epsilon > 0.$$ 

We call this complexity model error-based. One may also view an approximate solution an exact solution to a problem $\epsilon$-near to $\mathcal{P}$ with a well-defined measure in the data space. This is the so-called backward analysis model in numerical analysis.

If $f_A(m, n, \epsilon)$ is a polynomial in $m$, $n$, and $\log(1/\epsilon)$, then algorithm $A$ is a polynomial algorithm and problem $\mathcal{P}$ is polynomially solvable. Again, if $f_A(m, n, \epsilon)$ is independent of $\epsilon$ and polynomial in $m$ and $n$, then we say algorithm $A$ is a strongly polynomial algorithm. If $f_A(m, n, \epsilon)$ is a polynomial in $m$, $n$, and $(1/\epsilon)$, then algorithm $A$ is a polynomial approximation scheme or pseudo-polynomial algorithm. For some optimization problems, the complexity theory can be applied to prove not only that they cannot be solved in polynomial-time, but also that they do not have polynomial approximation schemes. In practice, approximation algorithms are widely used and accepted in practice.

Example 1.11 There is a strongly polynomial algorithm for sorting a vector in descending or ascending order, for matrix-vector multiplication, and for computing the norm of a vector.
Example 1.12 Consider the bisection method to locate a root of a continuous function \( f(x) : \mathbb{R} \to \mathbb{R} \) within interval \([0, 1]\), where \( f(0) > 0 \) and \( f(1) < 0 \). The method calls the oracle to evaluate \( f(1/2) \) (counted as one operation). If \( f(1/2) > 0 \), we throw away \([0, 1/2]\); if \( f(1/2) < 0 \), we throw away \((1/2, 1]\). Then we repeat this process on the remaining half interval. Each step of the method halves the interval that contains the root. Thus, in \( \log(1/\epsilon) \) steps, we must have an approximate root whose distance to the root is less than \( \epsilon \). Therefore, the bisection method is a polynomial algorithm.

We have to admit that the criterion of polynomiality is somewhat controversial. Many algorithms may not be polynomial but work fine in practice. This is because polynomiality is built upon the worst-case analysis. However, this criterion generally provides a qualitative statement: if a problem is polynomial solvable, then the problem is indeed relatively easy to solve regardless of the algorithm used. Furthermore, it is ideal to develop an algorithm with both polynomiality and practical efficiency.

1.4.1 Convergence rate

Most algorithms are iterative in nature. They generate a sequence of ever-improving points \( x^0, x^1, \ldots, x^k, \ldots \) approaching the solution set. For many optimization problems and/or algorithms, the sequence will never exactly reach the solution set. One theory of iterative algorithms, referred to as local or asymptotic convergence analysis, is concerned with the rate at which the optimality error of the generated sequence converges to zero.

Obviously, if each iteration of competing algorithms requires the same amount of work, the speed of the convergence of the error reflects the speed of the algorithm. This convergence rate, although it may hold locally or asymptotically, provides evaluation and comparison of different algorithms. It has been widely used by the nonlinear optimization and numerical analysis community as an efficiency criterion. In many cases, this criterion does explain practical behavior of iterative algorithms.

Consider a sequence of real numbers \( \{r^k\} \) converging to zero. One can define several notions related to the speed of convergence of such a sequence.

**Definition 1.1.** Let the sequence \( \{r^k\} \) converge to zero. The order of convergence of \( \{r^k\} \) is defined as the supremum of the nonnegative numbers \( p \) satisfying

\[
0 \leq \limsup_{k \to \infty} \frac{|r^{k+1}|}{|r^k|^p} < \infty.
\]
1.4. ALGORITHMS AND COMPUTATIONS

Definition 1.2. Let the sequence \( \{r^k\} \) converge to zero such that
\[
\limsup_{k \to \infty} \frac{|r^{k+1}|}{|r^k|^2} < \infty.
\]
Then, the sequence is said to converge quadratically to zero.

It should be noted that the order of convergence is determined only by the properties of the sequence that holds as \( k \to \infty \). In this sense we might say that the order of convergence is a measure of how good the tail of \( \{r^k\} \) is. Large values of \( p \) imply the faster convergence of the tail.

Definition 1.3. Let the sequence \( \{r^k\} \) converge to zero such that
\[
\limsup_{k \to \infty} \frac{|r^{k+1}|}{|r^k|} = \beta < 1.
\]
Then, the sequence is said to converge linearly to zero with convergence ratio \( \beta \).

Linear convergence is the most important type of convergence behavior. A linearly convergence sequence, with convergence ratio \( \beta \), can be said to have a tail that converges to zero at least as fast as the geometric sequence \( c\beta^k \) for a fixed number \( c \). Thus, we also call linear convergence geometric convergence.

As a rule, when comparing the relative effectiveness of two competing algorithms both of which produce linearly convergent sequences, the comparison is based on their corresponding convergence ratio—the smaller the ratio, the faster the algorithm. The ultimate case where \( \beta = 0 \) is referred to as superlinear convergence.

Example 1.13 Consider the conjugate gradient algorithm for minimizing
\[
\frac{1}{2} x^T Q x + c.
\]
Starting from an \( x^0 \in \mathbb{R}^n \) and \( d^0 = Q x^0 + c \), the method uses iterative formula
\[
x^{k+1} = x^k - \alpha^k d^k
\]
where
\[
\alpha^k = \frac{(d^k)^T (Q x^k + c)}{\|d^k\|_Q^2},
\]
and
\[
d^{k+1} = Q x^{k+1} - \theta^k d^k
\]
where
\[
\theta^k = \frac{(d^k)^T Q (Q x^{k+1} + c)}{\|d^k\|_Q^2}.
\]
This algorithm is superlinearly convergent (in fact, it converges in finite number of steps).
1.5 Basic Computational Procedures

There are several basic numerical problems frequently solved by interior-point algorithms.

1.5.1 Gaussian elimination method

Probably the best-known algorithm for solving a system of linear equations is the Gaussian elimination method. Suppose we want to solve

$$Ax = b.$$ 

We may assume $a_{11} \neq 0$ after some row switching, where $a_{ij}$ is the component of $A$ in row $i$ and column $j$. Then we can subtract appropriate multiples of the first equation from the other equations so as to have an equivalent system:

$$
\begin{pmatrix}
    a_{11} & A_1 \\
    0 & A'
\end{pmatrix}
\begin{pmatrix}
    x_1 \\
    x'
\end{pmatrix}
= 
\begin{pmatrix}
    b_1 \\
    b'
\end{pmatrix}.
$$

This is a pivot step, where $a_{11}$ is called a *pivot*, and $A'$ is called a *Schur complement*. Now, recursively, we solve the system of the last $m-1$ equations for $x'$. Substituting the solution $x'$ found into the first equation yields a value for $x_1$. The last process is called *back-substitution*.

In matrix form, the Gaussian elimination method transforms $A$ into the form

$$
\begin{pmatrix}
    U & C \\
    0 & 0
\end{pmatrix}
$$

where $U$ is a nonsingular, upper-triangular matrix,

$$A = L \begin{pmatrix}
    U & C \\
    0 & 0
\end{pmatrix},$$

and $L$ is a nonsingular, lower-triangular matrix. This is called the *LU-decomposition*.

Sometimes, the matrix is transformed further to a form

$$
\begin{pmatrix}
    D & C \\
    0 & 0
\end{pmatrix}
$$

where $D$ is a nonsingular, diagonal matrix. This whole procedure uses about $nm^2$ arithmetic operations. Thus, it is a strong polynomial-time algorithm.
1.5.2 Choleski decomposition method

Another useful method is to solve the least squares problem:

\[(LS) \ \text{minimize} \ |A^T y - c|.
\]

The theory says that \(y^*\) minimizes \(|A^T y - c|\) if and only if

\[AA^T y^* = Ac.
\]

So the problem is reduced to solving a system of linear equations with a symmetric semi-positive definite matrix.

One method is Choleski’s decomposition. In matrix form, the method transforms \(AA^T\) into the form

\[AA^T = L\Lambda L^T,
\]

where \(L\) is a lower-triangular matrix and \(\Lambda\) is a diagonal matrix. (Such a transformation can be done in about \(nm^2\) arithmetic operations as indicated in the preceding section.) \(L\) is called the Choleski factor of \(AA^T\). Thus, the above linear system becomes

\[L\Lambda L^T y^* = Ac,
\]

and \(y^*\) can be obtained by solving two triangle systems of linear equations.

1.5.3 The Newton method

The Newton method is used to solve a system of nonlinear equations: given \(f(x) : \mathbb{R}^n \to \mathbb{R}^n\), the problem is to solve \(n\) equations for \(n\) unknowns such that

\[f(x) = 0.
\]

The idea behind Newton’s method is to use the Taylor linear approximation at the current iterate \(x^k\) and let the approximation be zero:

\[f(x) \simeq f(x^k) + \nabla f(x^k)(x - x^k) = 0.
\]

The Newton method is thus defined by the following iterative formula:

\[x^{k+1} = x^k - \alpha(\nabla f(x^k))^{-1} f(x^k),
\]

where scalar \(\alpha \geq 0\) is called step-size. Rarely, however, is the Jacobian matrix \(\nabla f\) inverted. Generally the system of linear equations

\[\nabla f(x^k)dx = -f(x^k).
\]
is solved and $x^{k+1} = x^k + \alpha d_x$ is used. The direction vector $d_x$ is called a 
Newton step, which can be carried out in strongly polynomial time.

A modified or quasi Newton method is defined by 

$$x^{k+1} = x^k - \alpha M^k f(x^k),$$

where $M^k$ is an $n \times n$ symmetric matrix. In particular, if $M^k = I$, the 
method is called the gradient method, where $f$ is viewed as the gradient 
vector of a real function.

The Newton method has a superior asymptotic convergence order equal 
2 for $\|f(x^k)\|$. It is frequently used in interior-point algorithms, and believed 
to be the key to their effectiveness.

### 1.5.4 Solving ball-constrained linear problem

The ball-constrained linear problem has the following form:

\[(BP) \quad \text{minimize } c^T x \quad \text{subject to } \quad Ax = 0, \quad \|x\|^2 \leq 1,\]

or

\[(BD) \quad \text{minimize } b^T y \quad \text{subject to } \quad \|A^T y\|^2 \leq 1.\]

$x^*$ minimizes (BP) if and only if there always exists a $y$ such that they 
satisfy

$$AA^T y = Ac,$$

and if $c - A^T y \neq 0$ then

$$x^* = -(c - A^T y)/\|c - A^T y\|;$$

otherwise any feasible $x$ is a solution. The solution $y^*$ for (BD) is given as 
follows: Solve

$$AA^T \bar{y} = b,$$

and if $\bar{y} \neq 0$ then set

$$y^* = -\bar{y}/\|A^T \bar{y}\|;$$

otherwise any feasible $y$ is a solution. So these two problems can be reduced 
to solving a system of linear equations.
1.5.5 Solving ball-constrained quadratic problem

The ball-constrained quadratic problem has the following form:

\[(BP) \quad \text{minimize} \quad \frac{1}{2} x^T Q x + c^T x \]

subject to \(Ax = 0, \|x\|^2 \leq 1,\)

or simply

\[(BD) \quad \text{minimize} \quad \frac{1}{2} y^T Q y + b^T y \]

subject to \(\|y\|^2 \leq 1.\)

This problem is used by the classical trust region method for nonlinear optimization. The optimality conditions for the minimizer \(y^*\) of (BD) are

\[(Q + \mu^* I)y^* = -b, \quad \mu^* \geq 0, \quad \|y^*\|^2 \leq 1, \quad \mu^*(1 - \|y^*\|^2) = 0,\]

and

\[(Q + \mu^* I) \succeq 0.\]

These conditions are necessary and sufficient. This problem can be solved in polynomial time \(\log(1/\epsilon)\) and \(\log(\log(1/\epsilon))\) by the bisection method or a hybrid of the bisection and Newton methods, respectively. In practice, several trust region procedures have been very effective in solving this problem.

The ball-constrained quadratic problem will be used as a sub-problem by several interior-point algorithms in solving complex optimization problems. We will discuss them later in the book.

1.6 Notes

The term “complexity” was introduced by Hartmanis and Stearns [180]. Also see Garey and Johnson [131] and Papadimitriou and Steiglitz [337]. The NP theory was due to Cook [88] and Karp [219]. The importance of \(P\) was observed by Edmonds [106].

Linear programming and the simplex method were introduced by Dantzig [93]. Other inequality problems and convexity theories can be seen in Gritzmann and Klee [169], Grötschel, Lovász and Schrijver [170], Grünbaum [171], Rockafellar [364], and Schrijver [373]. Various complementarity problems can be found found in Cottle, Pang and Stone [91]. The positive semidefinite programming, an optimization problem in nonpolyhedral cones, and its applications can be seen in Nesterov and Nemirovskii [327], Alizadeh [9], and Boyd, Ghaoui, Feron and Balakrishnan [72]. Recently, Goemans and Williamson [141] obtained several breakthrough results on approximation algorithms using positive semi-definite programming. The
KKT condition for nonlinear programming was given by Karush, Kuhn and Tucker [240].

It was shown by Klee and Minty [223] that the simplex method is not a polynomial-time algorithm. The ellipsoid method, the first polynomial-time algorithm for linear programming with rational data, was proven by Khachiyan [221]; also see Bland, Goldfarb and Todd [63]. The method was devised independently by Shor [380] and by Nemirovskii and Yudin [321]. The interior-point method, another polynomial-time algorithm for linear programming, was developed by Karmarkar. It is related to the classical barrier-function method studied by Frisch [126] and Fiacco and McCormick [116]; see Gill, Murray, Saunders, Tomlin and Wright [139], and Anstreicher [23]. For a brief LP history, see the excellent article by Wright [457].

The real computation model was developed by Blum, Shub and Smale [66] and Nemirovskii and Yudin [321]. The average setting can be seen in Traub, Wasilkowski and Wozniakowski [420]. The asymptotic convergence rate and ratio can be seen in Luenberger [248], Ortega and Rheinboldt [334], and Traub [419]. Other complexity issues in numerical optimization were discussed in Vavasis [450].

Many basic numerical procedures listed in this chapter can be found in Golub and Van Loan [154]. The ball-constrained quadratic problem and its solution methods can be seen in Moré [308], Sorenson [387], and Dennis and Schnable [96]. The complexity result of the ball-constrained quadratic problem was proved by Vavasis [450] and Ye [469, 473].

1.7 Exercises

1.1 Let $Q \in \mathbb{R}^{n \times n}$ be a given nonsingular matrix, and $a$ and $b$ be given $\mathbb{R}^n$ vectors. Show

$$
(Q + ab^T)^{-1} = Q^{-1} - \frac{1}{1 + b^T Q^{-1} a} Q^{-1} ab^T Q^{-1}.
$$

This formula is called the Sherman-Morrison-Woodbury formula.

1.2 Prove that the eigenvalues of all matrices $Q \in \mathbb{M}^{n \times n}$ are real. Furthermore, show that $Q$ is PSD if and only if all its eigenvalues are non-negative, and $Q$ is PD if and only if all its eigenvalues are positive.

1.3 Using the ellipsoid representation in Section 1.2.2, find the matrix $Q$ and vector $y$ that describes the following ellipsoids:

1. The 3-dimensional sphere of radius 2 centered at the origin;
2. The 2-dimensional ellipsoid centered at \((1; 2)\) that passes the points \((0; 2), (1; 0), (2; 2),\) and \((1; 4)\);

3. The 2-dimensional ellipsoid centered at \((1; 2)\) with axes parallel to the line \(y = x\) and \(y = -x\), and passing through \((-1; 0), (3; 4), (0; 3),\) and \((2; 1)\).

1.4 Show that the biggest coordinate-aligned ellipsoid that is entirely contained in \(\mathbb{R}^n_+\) and has its center at \(x^0 \in \mathbb{R}^n_+\) can be written as:

\[
E(x^0) = \{ x \in \mathbb{R}^n : \parallel (X^0)^{-1}(x - x^0) \parallel \leq 1 \}.
\]

1.5 Show that the non-negative orthant, the positive semi-definite cone, and the second-order cone are all self-dual.

1.6 Consider the convex set \(C = \{ x \in \mathbb{R}^2 : (x_1 - 1)^2 + (x_2 - 1)^2 \leq 1 \}\) and let \(y \in \mathbb{R}^2\). Assuming \(y \notin C\),

1. Find the point in \(C\) that is closest to \(y\);

2. Find a separating hyperplane vector as a function of \(y\).

1.7 Using the idea of Exercise 1.6, prove the separating hyperplane theorem 1.1.

1.8 Given an \(m \times n\) matrix \(A\) and a vector \(c \in \mathbb{R}^n\), consider the function \(B(y) = \sum_{j=1}^n \log s_j\) where \(s = c - A^T y > 0\). Find \(\nabla B(y)\) and \(\nabla^2 B(y)\) in terms of \(s\).

1.9 Prove that the level set of a quasi-convex function is convex.

1.10 Prove Propositions 1.3 and 1.4 for convex functions in Section 1.2.3.

1.11 Let \(f_1, \ldots, f_m\) be convex functions. Then, the function \(\bar{f}(x)\) defined below is also convex:

- \[
\max_{i=1,\ldots,m} f_i(x)
\]
- \[
\sum_{i=1}^m f_i(x)
\]

1.12 Prove the Harmonic inequality described in Section 1.2.4.
1.13 Prove Farkas’ lemma 1.7 for linear equations.

1.14 Prove the linear least-squares problem always has a solution.

1.15 Let \( P = A^T (AA^T)^{-1} A \) or \( P = I - A^T (AA^T)^{-1} A \). Then prove
1. \( P = P^2 \).
2. \( P \) is positive semi-definite.
3. The eigenvalues of \( P \) are either 0 or 1.

1.16 Using the separating theorem, prove Farkas’ lemmas 1.8 and 1.9.

1.17 If a system \( A^T y \leq c \) of linear inequalities in \( m \) variables has no solution, show that \( A^T y \leq c \) has a subsystem \((A')^T y \leq c'\) of at most \( m + 1 \) inequalities having no solution.

1.18 Prove the LP fundamental theorem 1.14.

1.19 If \((LP)\) and \((LD)\) have a nondegenerate optimal basis \( A_B \), prove that the strict complementarity partition in Theorem 1.13 is
\[
P^* = B.
\]

1.20 If \( Q \) is positive semi-definite, prove that \( x^* \) is an optimal solution for \((QP)\) if and only if \( x^* \) is a KKT point for \((QP)\).

1.21 Prove \( X \bullet S \geq 0 \) if both \( X \) and \( S \) are positive semi-definite matrices.

1.22 Prove that two positive semi-definite matrices are complementary to each other, \( X \bullet S = 0 \), if and only if \( XS = 0 \).

1.23 Let both \((LP)\) and \((LD)\) for a given data set \((A,b,c)\) have interior feasible points. Then consider the level set
\[
\Omega(z) = \{ y : c - A^T y \geq 0, -z + b^T y \geq 0 \}
\]
where \( z < z^* \) and \( z^* \) designates the optimal objective value. Prove that \( \Omega(z) \) is bounded and has an interior for any finite \( z < z^* \), even \( \mathcal{F}_d \) is unbounded.

1.24 Given an \((LP)\) data set \((A,b,c)\) and an interior feasible point \( x^0 \), find the feasible direction \( d_x \) (\( Ad_x = 0 \)) that achieves the steepest decrease in the objective function.
1.25 Given an (LP) data set \((A, b, c)\) and a feasible point \((x^0, y^0, s^0) \in (\mathbb{R}^n_+, \mathbb{R}^m, \mathbb{R}^n_+)\) for the primal and dual, and ignoring the nonnegativity condition, write the systems of linear equations used to calculate the Newton steps for finding points that satisfy the optimality equations (1.2) and (1.3), respectively.

1.26 Show the optimality conditions for the minimizer \(y^*\) of (BD) in Section 1.5.5:

\[(Q + \mu^* I)y^* = -b, \quad \mu^* \geq 0, \quad \|y^*\| \leq 1, \quad \mu^*(1 - \|y^*\|) = 0,\]

and

\[(Q + \mu^* I) \succeq 0,\]

are necessary and sufficient.