Part I

Foundations

1

The Karmarkar Revolution

Optimization problems seek points that maximize or minimize stated objective functions over feasible regions that are defined herein by given sets of equality and inequality constraints. Equation-solving problems seek points that simultaneously satisfy given sets of equality constraints. We restrict attention to problems where the underlying spaces are *real* and *finite-dimensional* and the functions defining objectives and constraints are *differentiable*. Problems of this type arise in all areas of science and engineering, and our focus in this monograph is on the *unified study of algorithms*[1] for solving them.

1.1 Classical Portrait of the Field

The genealogy of algorithmic optimization and equation-solving can be traced to the works of venerated mathematicians—Cauchy, Euler, Fourier, Gauss, Kantorovich, Lagrange, Newton, Poincaré, and others. But it was only after a very long gestation period that the subject was truly born in the mid nineteen forties with Dantzig's discovery of the wide-ranging practical applicability of the linear programming model—the flagship of the

[1]Algorithmic techniques for solving differentiable problems over R^n are also important because they provide a foundation for solving more general classes of optimization problems defined over finite- or infinite-dimensional spaces, for example, problems of nondifferentiable programming, stochastic programming, semidefinite programming, and optimal control.

field—and his invention of its main solution engine, the simplex method. Other areas of mathematical programming developed rapidly in what may fittingly be termed the *Dantzig Modeling-and-Algorithmic Revolution*. Its history is nicely told in the foreword of the recent book by Dantzig and Thapa [1997]. See also the linear programming classic Dantzig [1963].

The basic models of differentiable optimization and equation-solving within the classical treatment will now be itemized. In the discussion that follows, let f, f_i, F, h_i, and g_j denote smooth, nonlinear functions from $R^n \to R$:

$$\begin{aligned} &\textit{Linear Programming Model} \\ &\text{LP:} \quad \text{minimize}_{\mathbf{x}\in R^n} \quad \mathbf{c}^T\mathbf{x} \\ &\text{s.t.} \quad \mathbf{a}_i^T\mathbf{x} = b_i, \quad i = 1, \ldots, m, \\ &l_j \le x_j \le u_j, \quad j = 1, \ldots, n, \end{aligned} \tag{1.1}$$

where m and n are positive integers, $m \le n$, $\mathbf{x}$ is an n-vector with components x_j, $\mathbf{c}$ and $\mathbf{a}_i$ are n-vectors, and $\mathbf{b}$ is an m-vector with components b_i. The quantities l_j and u_j are lower and upper bounds, which are also permitted to assume the values $-\infty$ and $+\infty$, respectively.

The special structure of the feasible region of the LP model, namely, a convex polytope, and the fact that its preeminent solution engine, the simplex algorithm, proceeds along a path of vertices of the feasible polytope, imparted the flavor of combinatorial or *discrete* mathematics to the subject of linear programming. As a result, the development of algorithmic optimization during the four decades following the Dantzig revolution was characterized by a *prominent watershed between linear programming and the remainder of the subject, nonlinear programming (NLP)*. Many of the NLP solution algorithms are rooted in analytic techniques named for mathematicians mentioned above, and the subject retained the flavor of *continuous* mathematics.

Nonlinear programming, in its classical development, itself displayed a *secondary* but equally marked *watershed* between the following:

- problems with no constraints or only linear constraints, together with the closely allied areas of unconstrained nonlinear least squares and the solution of systems of nonlinear equations;
- problems with more general nonlinear equality and/or inequality constraints, including the special case where the functions defining the objective and constraints are convex.

More specifically, the problems on one side of this secondary watershed are as follows:

$$\begin{aligned} &\textit{Unconstrained Minimization Model} \\ &\text{UM:} \quad \text{minimize}_{\mathbf{x}\in R^n} \quad f(\mathbf{x}). \end{aligned} \tag{1.2}$$

Joined to UM in the classical treatment was the problem of finding a global minimum of a nonlinear function, which we will identify by the acronym GUM. An important special case of UM and GUM is the choice $n = 1$ or unidimensional optimization.

Nonlinear Least-Squares Model

$$\text{NLSQ:} \quad \text{minimize}_{\mathbf{x} \in R^n} \quad F(\mathbf{x}) = \sum_{i=1}^{l} [f_i(\mathbf{x})]^2, \tag{1.3}$$

where l is a positive integer. NLSQ is obviously a particular instance of the UM problem.

Nonlinear Equations Model

$$\text{NEQ:} \quad \text{Solve}_{\mathbf{x} \in R^n} \quad h_i(\mathbf{x}) = 0, \quad i = 1, \ldots, n. \tag{1.4}$$

The choice $n = 1$ is an important special case. A solution of NEQ is a *global* minimizing point of $\sum_{i=1}^{n} [h_i(\mathbf{x})]^2$, thereby providing the link to NLSQ, UM, and GUM.

The traditional treatment of the UM, NLSQ, and NEQ areas can be found, for example, in Dennis and Schnabel [1983].

When *linear* equality or inequality constraints are present, the algorithms for UM, NLSQ, and GUM can be extended to operate within appropriately defined affine subspaces, and the resulting subject remains closely allied. The problem is as follows:

Linear Convex-Constrained Programming Model

$$\begin{aligned} &\text{LCCP:} \quad \text{minimize}_{\mathbf{x} \in R^n} \quad f(\mathbf{x}) \\ &\text{s.t.} \quad \mathbf{a}_i^T \mathbf{x} = b_i, \quad i = 1, \ldots, m, \\ &\quad l_j \leq x_j \leq u_j, \quad j = 1, \ldots, n, \end{aligned} \tag{1.5}$$

where the associated quantities are defined as in LP above. An important special case arises when the objective function is a convex or nonconvex quadratic function, a subject known as quadratic programming (QP).

On the *other side* of the secondary watershed within nonlinear programming that was mentioned above, we have the following problem areas:

Nonlinear Equality-Constrained Programming Model

$$\begin{aligned} &\text{NECP:} \quad \text{minimize}_{\mathbf{x} \in R^n} \quad f(\mathbf{x}) \\ &\text{s.t.} \quad h_i(\mathbf{x}) = 0, \quad i = 1, \ldots, m. \end{aligned} \tag{1.6}$$

Convex-Constrained Programming Model

$$\begin{aligned} &\text{CCP:} \quad \text{minimize}_{\mathbf{x} \in R^n} \quad f(\mathbf{x}) \\ &\text{s.t.} \quad \mathbf{a}_i^T \mathbf{x} = b_i, \quad i = 1, \ldots, m, \\ &\quad g_j^c(\mathbf{x}) \leq 0, \quad j = 1, \ldots, \bar{n}, \end{aligned} \tag{1.7}$$

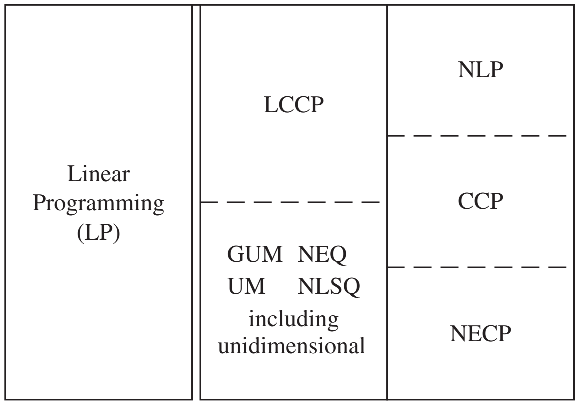


FIGURE 1.1 Classical differentiable optimization and equation-solving.

where $\bar{n}$ is an integer and $g_j^c : R^n \to R$ are smooth, convex functions.

General Nonlinear Programming Model

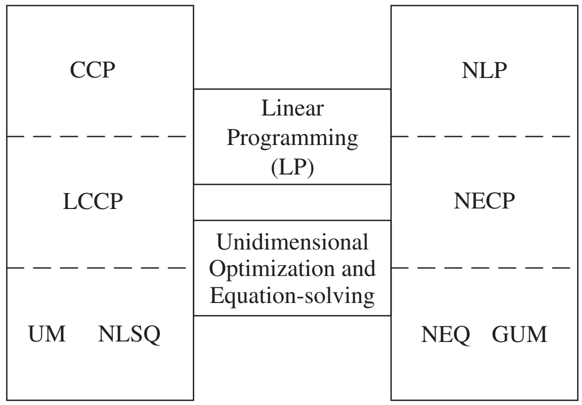
$$
\begin{aligned}
\text{NLP:} \quad & \text{minimize}_{\mathbf{x} \in R^n} \quad f(\mathbf{x}) \\
\text{s.t.} \quad & h_i(\mathbf{x}) = 0, \quad i = 1, \ldots, m, \\
& g_j(\mathbf{x}) \leq 0, \quad j = 1, \ldots, \bar{n}.
\end{aligned}
\tag{1.8}
$$

Thus, using the acronyms in (1.1)–(1.8), we can summarize the organizational picture of the subject that emerged in the four decades following the Dantzig modeling-and-algorithmic revolution by Figure 1.1.

1.2 Modern Portrait of the Field

Following a decade of intense research by many investigators worldwide, motivated by Karmarkar's 1984 article, a new portrait has emerged that has left virtually none of the foregoing problem areas untouched. Again using the acronyms defined in (1.1)–(1.8), this new portrait is summarized in Figure 1.2, and its main features are as follows:

- **UM**, NLSQ, LCCP, CCP: Algorithms for UM seek a *local* minimum, i.e., a point of local convexity of an arbitrary, smooth objective function $f(\mathbf{x})$ over the convex region R^n. The NLSQ problem is a special case of UM. The LCCP and CCP problems are natural extensions where R^n is replaced by a convex polytope and a more general convex region, respectively.
- **NEQ**, GUM, NECP, NLP: Points that satisfy a set of nonlinear equations (equality constraints) form a set that is *not* necessarily convex.



"The great watershed in optimization isn't between linearity and nonlinearity but convexity and nonconvexity" (Rockafellar [1993]).

FIGURE 1.2 Modern differentiable optimization and equation-solving.

Thus, the NEQ problem belongs naturally on the nonconvex side of the watershed. The GUM problem must take the nonconvexity of the objective function into consideration, and it also belongs naturally here. When the number of equality constraints in NEQ is fewer than the number of variables and an objective function is optimized over the resulting feasible region, one obtains the NECP problem. When inequality constraints are introduced, one obtains the NLP problem in full generality.

- **1-D:** Unidimensional optimization and equation-solving bridge the watershed depicted in Figure 1.2. Algorithms for solving UM and NEQ for the special case $n = 1$ are much more closely interrelated than their multidimensional counterparts. In the classical treatment, false lessons were learned from the 1-D case, which contributed to the incorrect grouping of UM and NEQ under a common umbrella, as depicted in Figure 1.1.
- **LP:** Linear programming is the other bridge that traverses the convexity/nonconvexity watershed. Some of its key algorithms are derived, in a natural way, by approaching the subject from the convexity side. Others are best approached from the nonconvexity side.

In this monograph, we focus on the *foundations* of our subject, namely, the **UM**, **NEQ**, **1-D**, and **LP** problems and the unified study of algorithms for solving them.[2] Because the areas of linear programming, nonlinear pro-

[2]Other problems in Figure 1.2 build on these foundations. For a comprehensive and up-to-date treatment that fits the overall pattern of Figure 1.2, see Bertsekas [1999].

gramming, and nonlinear equation-solving are so closely interconnected in the post-Karmarkar treatment, it is appropriate that the entire subject be called *differentiable optimization and equation-solving*, or, more compactly, *differentiable programming*.[3]

1.3 Overview of Monograph

Our discussion falls into five parts, each comprising three chapters, as follows:

Part I: In the next two chapters, we consider algorithms for the unconstrained minimization and nonlinear equation-solving problems, respectively.

In Chapter 2, we show that a single underlying technique, the Newton–Cauchy (NC) method, lies at the heart of most unconstrained minimization algorithms in current use. It combines a model in the spirit of Newton's method and a metric in the spirit of Cauchy's.

In Chapter 3, we first give examples to illustrate the well-known fact that solving nonlinear equations via nonlinear least squares is useful in a practical sense, but inherently flawed from a conceptual standpoint. Homotopy-based Euler–Newton (or predictor/corrector) techniques, neglected in the classical treatment in favor of techniques based on the sum-of-squares merit function (Figure 1.1), must be given a more prominent role at the algorithmic foundations of the subject. A natural association between nonlinear equation-solving and nonlinear equality-constrained programming (NECP) reveals a fundamental alternative, the Lagrange–NC method. It employs Lagrangian-based potentials derived from standard NECP techniques, and uses them, in conjunction with Newton–Cauchy algorithms, to attack the nonlinear equation-solving problem.

Part II, in three chapters, covers lessons that can be learned from unidimensional programming.

The first, Chapter 4, explains why algorithms for solving unconstrained minimization and nonlinear equation-solving for the special case $n = 1$ are much more closely interrelated than their multidimensional counterparts.

Next, the key role of the unidimensional line search in the fortune of the Fletcher–Reeves nonlinear conjugate-gradient algorithm is the topic of Chapter 5.

A third unidimensional lesson is considered in Chapter 6. Here, a link is established between classical golden-section search and the generic Nelder–

[3] These alternative names also provide a nice counterpoint to the widely studied area of nondifferentiable (or nonsmooth) optimization and equation-solving.

Mead (NM) direct-search algorithm restricted to $n = 1$, in order to obtain a conceptually more satisfactory approach to the multidimensional NM algorithm.

Part III addresses the linear programming problem from the perspective of the left half of the watershed of Figure 1.2.

Both the simplex algorithm of Dantzig and the affine-scaling interior algorithm of Dikin can be conveniently approached from the convexity side. The simplex method is a highly specialized form of active-set method for linearly constrained convex programming. And the NC model/metric approach to unconstrained minimization of Chapter 2, restricted to appropriate diagonal matrices, carries across to linear programming and yields Dikin's algorithm. These key algorithms along with dual and primal–dual variants are the focus of Chapters 7 and 8. The complexity of linear programming algorithms discussed in these two chapters is *not* known to be polynomial.

Diagonal metrics for *nonlinear* unconstrained minimization problems, under the rubric quasi-Cauchy (QC), are also considered in this part of the monograph, in its concluding Chapter 9.

Part IV again addresses the linear programming problem, now approached from the right half of the watershed of Figure 1.2.

The homotopy-based Euler–Newton method of Chapter 3, applied to the Karush–Kuhn–Tucker optimality equations of a linear program, yields a variety of basic path-following interior algorithms discussed in Chapter 10.

Extracting the fundamental connection with the classical logarithmic-barrier approach to nonlinear programming is the topic of Chapter 11.

Barrier functions, in turn, motivate potential functions—the centerpiece of Karmarkar's algorithm—as discussed in Chapter 12.

Linear programming algorithms discussed in this part of the monograph usually exhibit polynomial complexity.

Part V: Parts I–IV of the monograph provide a vehicle for considering larger issues, the topic of the last three chapters.

Within the specific setting of the unconstrained minimization problem, Chapter 13 considers basic algorithmic principles on which variable-metric and related algorithms are premised. Implications, in general, for other families of algorithms are also discussed.

Chapter 14 develops a new paradigm for optimization and equation-solving based on the fundamental Darwinian ideas of population-thinking and variation.

Finally, Chapter 15 concludes the monograph with a philosophy and vision for the emerging discipline of algorithmic science.

1.4 Notes

Section 1.2: Figure 1.2 and the post-Karmarkar restructuring are based on Nazareth [2000].

http://www.springer.com/978-0-387-95572-8

Differentiable Optimization and Equation Solving
A Treatise on Algorithmic Science and the Karmarkar Revolution
Nazareth, J.

2003, XVII, 256 p. 2 illus., Hardcover
ISBN: 978-0-387-95572-8