Now we sum the bracketed partial sum of a geometric series to get

$$\begin{aligned}(1-t)^{k-1}\left[\left(\frac{1-\tilde{t}}{1-t}\right)^{k-1}+\cdots+\frac{1-\tilde{t}}{1-t}+1\right] &= (1-t)^{k-1}\frac{1-\left(\frac{1-\tilde{t}}{1-t}\right)^{k}}{1-\frac{1-\tilde{t}}{1-t}} \\ &= \frac{(1-\tilde{t})^k-(1-t)^k}{t-\tilde{t}}.\end{aligned}$$

Recalling that $\tilde{t}=t(1-\delta)$ and dropping the second term in the numerator, we get

$$\frac{(1-\tilde{t})^k-(1-t)^k}{t-\tilde{t}} \le \frac{(1-\tilde{t})^k}{\delta t}.$$

Putting this all together, we see that

$$\gamma^{(k)} \le (1-\tilde{t})^k\left(\gamma^{(0)}+\frac{\tilde{M}}{\delta t}\right).$$

Denoting the parenthesized expression by $\bar{M}$ completes the proof. □

Theorem 18.1 is only a partial convergence result because it depends on the assumption that the step lengths remain bounded away from zero. To show that the step lengths do indeed have this property requires that the algorithm be modified and that the starting point be carefully selected. The details are rather technical and hence omitted (see the Notes at the end of the chapter for references).

Also, before we leave this topic, note that the primal and dual infeasibilities go down by a factor of $1-t$ at each iteration, whereas the duality gap goes down by a smaller amount $1-\tilde{t}$. The fact that the duality gap converges more slowly that the infeasibilities is also readily observed in practice.

Exercises

18.1 Starting from $(x,w,y,z)=(e,e,e,e)$, and using $\delta=1/10$, and $r=9/10$, compute (x,w,y,z) after one step of the path-following method for the problem given in
 (a) Exercise 2.3.
 (b) Exercise 2.4.
 (c) Exercise 2.5.
 (d) Exercise 2.10.

18.2 Let $\{(x_\mu,w_\mu,y_\mu,z_\mu):\mu\ge 0\}$ denote the central trajectory. Show that

$$\lim_{\mu\to\infty} b^Ty_\mu - c^Tx_\mu = \infty.$$

Hint: look at (18.5).

18.3 Consider a linear programming problem whose feasible region is bounded and has nonempty interior. Use the result of Exercise 18.2 to show that the dual problem's feasible set is unbounded.

18.4 *Scale invariance*. Consider a linear program and its dual:

$$(P) \quad \begin{array}{rl} \max & c^T x \\ \text{s.t.} & Ax + w = b \\ & x, w \geq 0 \end{array} \qquad (D) \quad \begin{array}{rl} \min & b^T y \\ \text{s.t.} & A^T y - z = c \\ & y, z \geq 0. \end{array}$$

Let R and S be two given diagonal matrices having positive entries along their diagonals. Consider the *scaled* reformulation of the original problem and its dual:

$$(\bar{P}) \quad \begin{array}{rl} \max & (Sc)^T \bar{x} \\ \text{s.t.} & RAS\bar{x} + \bar{w} = Rb \\ & \bar{x}, \bar{w} \geq 0 \end{array} \qquad (\bar{D}) \quad \begin{array}{rl} \min & (Rb)^T \bar{y} \\ \text{s.t.} & SA^T R\bar{y} - \bar{z} = Sc \\ & \bar{y}, \bar{z} \geq 0. \end{array}$$

Let (x^k, w^k, y^k, z^k) denote the sequence of solutions generated by the primal–dual interior-point method applied to (P)–(D). Similarly, let $(\bar{x}^k, \bar{w}^k, \bar{y}^k, \bar{z}^k)$ denote the sequence of solutions generated by the primal–dual interior-point method applied to $(\bar{P})$–$(\bar{D})$. Suppose that we have the following relations among the starting points:

$$\bar{x}^0 = S^{-1}x^0, \quad \bar{w}^0 = Rw^0, \quad \bar{y}^0 = R^{-1}y^0, \quad \bar{z}^0 = Sz^0.$$

Show that these relations then persist. That is, for each $k \geq 1$,

$$\bar{x}^k = S^{-1}x^k, \quad \bar{w}^k = Rw^k, \quad \bar{y}^k = R^{-1}y^k, \quad \bar{z}^k = Sz^k.$$

18.5 *Homotopy method*. Let $\bar{x}$, $\bar{y}$, $\bar{z}$, and $\bar{w}$ be given componentwise positive "initial" values for x, y, z, and w, respectively. Let t be a parameter between 0 and 1. Consider the following nonlinear system:

$$\begin{aligned} Ax + w &= tb + (1-t)(A\bar{x} + \bar{w}) \\ A^T y - z &= tc + (1-t)(A^T\bar{y} - \bar{z}) \\ XZe &= (1-t)\bar{X}\bar{Z}e \\ YWe &= (1-t)\bar{Y}\bar{W}e \\ x, y, z, w &> 0. \end{aligned} \tag{18.12}$$

(a) Use Exercise 17.3 to show that this nonlinear system has a unique solution for each $0 \leq t < 1$. Denote it by $(x(t), y(t), z(t), w(t))$.
(b) Show that $(x(0), y(0), z(0), w(0)) = (\bar{x}, \bar{y}, \bar{z}, \bar{w})$.
(c) Assuming that the limit

$$(x(1), y(1), z(1), w(1)) = \lim_{t \to 1} (x(t), y(t), z(t), w(t))$$

exists, show that it solves the standard-form linear programming problem.

(d) The family of solutions $(x(t), y(t), z(t), w(t))$, $0 \le t < 1$, describes a curve in "primal–dual" space. Show that the tangent to this curve at $t = 0$ coincides with the path-following step direction at $(\bar{x}, \bar{y}, \bar{z}, \bar{w})$ computed with $\mu = 0$; that is,

$$\left(\frac{dx}{dt}(0), \frac{dy}{dt}(0), \frac{dz}{dt}(0), \frac{dw}{dt}(0)\right) = (\Delta x, \Delta y, \Delta z, \Delta w),$$

where $(\Delta x, \Delta y, \Delta z, \Delta w)$ is the solution to (18.1)–(18.4).

18.6 *Higher-order methods*. The previous exercise shows that the path-following step direction can be thought of as the direction one gets by approximating a homotopy path with its tangent line:

$$x(t) \approx x(0) + \frac{dx}{dt}(0)t.$$

By using more terms of the Taylor's series expansion, one can get a better approximation:

$$x(t) \approx x(0) + \frac{dx}{dt}(0)t + \frac{1}{2}\frac{d^2x}{dt^2}(0)t^2 + \cdots + \frac{1}{k!}\frac{d^kx}{dt^k}(0)t^k.$$

(a) Differentiating the equations in (18.12) twice, derive a linear system for $(d^2x/dt^2(0), d^2y/dt^2(0), d^2z/dt^2(0), d^2w/dt^2(0))$.

(b) Can the same technique be applied to derive linear systems for the higher-order derivatives?

18.7 *Linear Complementarity Problem*. Given a $k \times k$ matrix M and a k-vector q, a vector x is said to solve the linear complementarity problem if

$$\begin{aligned} -Mx + z &= q \\ XZe &= 0 \\ x, z &\ge 0 \end{aligned}$$

(note that the first equation can be taken as the definition of z).

(a) Show that the optimality conditions for linear programming can be expressed as a linear complementarity problem with

$$M = \begin{bmatrix} 0 & -A \\ A^T & 0 \end{bmatrix}.$$

(b) The path-following method introduced in this chapter can be extended to cover linear complementarity problems. The main step in the derivation is to replace the complementarity condition $XZe = 0$ with a μ-complementarity condition $XZe = \mu e$ and then to use Newton's method to derive step directions Δx and Δz. Carry out this procedure and indicate the system of equations that define Δx and Δz.

(c) Give conditions under which the system derived above is guaranteed to have a unique solution.
(d) Write down the steps of the path-following method for the linear complementarity problem.
(e) Study the convergence of this algorithm by adapting the analysis given in Section 18.5.

18.8 Consider again the L^1-regression problem:

$$\text{minimize } \|b - Ax\|_1.$$

Complete the following steps to derive the step direction vector Δx associated with the primal-dual affine-scaling method for solving this problem.

(a) Show that the L^1-regression problem is equivalent to the following linear programming problem:

$$\begin{array}{lll} & \text{minimize} & e^T(t_+ + t_-) \\ (18.13) & \text{subject to} & Ax + t_+ - t_- = b \\ & & t_+, t_- \geq 0. \end{array}$$

(b) Write down the dual of (18.13).
(c) Add slack and/or surplus variables as necessary to reformulate the dual so that all inequalities are simple nonnegativities of variables.
(d) Identify all primal-dual pairs of complementary variables.
(e) Write down the nonlinear system of equations consisting of: (1) the primal equality constraints, (2) the dual equality constraints, (3) all complementarity conditions (using $\mu = 0$ since we are looking for an affine-scaling algorithm).
(f) Apply Newton's method to the nonlinear system to obtain a linear system for step directions for all of the primal and dual variables.
(g) We may assume without loss of generality that both the initial primal solution and the initial dual solution are feasible. Explain why.
(h) The linear system derived above is a 6×6 block matrix system. But it is easy to solve most of it by hand. First eliminate those step directions associated with the nonnegative variables to arrive at a 2×2 block matrix system.
(i) Next, solve the 2×2 system. Give an explicit formula for Δx.
(j) How does this primal-dual affine-scaling algorithm compare with the iteratively reweighted least squares algorithm defined in Section 12.5?

18.9

(a) Let ξ_j, $j = 1, 2, \ldots$, denote a sequence of real numbers between zero and one. Show that $\prod_j (1 - \xi_j) = 0$ if $\sum_j \xi_j = \infty$.

(b) Use the result of part a to prove the following convergence result: if the sequences $\|x^{(k)}\|_\infty$, $k = 1, 2, \ldots$, and $\|y^{(k)}\|_\infty$, $k = 1, 2, \ldots$, are bounded and $\sum_k \theta^{(k)} = \infty$, then

$$\lim_{k \to \infty} \|\rho^{(k)}\|_1 = 0$$
$$\lim_{k \to \infty} \|\sigma^{(k)}\|_1 = 0$$
$$\lim_{k \to \infty} \gamma^{(k)} = 0.$$

Notes

The path-following algorithm introduced in this chapter has its origins in a paper by Kojima et al. (1989). Their paper assumed an initial feasible solution and therefore was a true interior-point method. The method given in this chapter does not assume the initial solution is feasible—it is a one-phase algorithm. The simple yet beautiful idea of modifying the Kojima–Mizuno–Yoshise primal–dual algorithm to make it into a one-phase algorithm is due to Lustig (1990).

Of the thousands of papers on interior-point methods that have appeared in the last decade, the majority have included convergence proofs for some version of an interior-point method. Here, we only mention a few of the important papers. The first polynomial-time algorithm for linear programming was discovered by Khachian (1979). Khachian's algorithm is fundamentally different from any algorithm presented in this book. Paradoxically, it proved in practice to be inferior to the simplex method. N.K. Karmarkar's pathbreaking paper (Karmarkar 1984) contained a detailed convergence analysis. His claims, based on preliminary testing, that his algorithm is uniformly substantially faster than the simplex method sparked a revolution in linear programming. Unfortunately, his claims proved to be exaggerated, but nonetheless interior-point methods have been shown to be competitive with the simplex method and usually superior on very large problems. The convergence proof for a primal–dual interior-point method was given by Kojima et al. (1989). Shortly thereafter, Monteiro & Adler (1989) improved on the convergence analysis. Two recent survey papers, Todd (1995) and Anstreicher (1996), give nice overviews of the current state of the art. Also, a soon-to-be-published book by Wright (1996) should prove to be a valuable reference to the reader wishing more information on convergence properties of these algorithms.

The homotopy method outlined in Exercise 18.5 is described in Nazareth (1986) and Nazareth (1996). Higher-order path-following methods are described (differently) in Carpenter et al. (1993).