Lecture 16: CS395T Numerical Optimization for Graphics and AI — Linear Programming (Simplex Method II and Interior Point Method I

Qixing Huang
The University of Texas at Austin
huangqx@cs.utexas.edu

Disclaimer

This note is adapted from

- Section 13 and Section 14 of *Numerical Optimization* by Jorge Nocedal and Stephen J. Wright. Springer series in operations research and financial engineering. Springer, New York, NY, 2. ed. edition, (2006)

1 Simplex Methods

In class we will discuss the following items regarding the simplex method:

Dual Program and Optimality Conditions. The primal problem of LP

$$\begin{aligned} &\min_{\boldsymbol{x}\in\mathbb{R}^n} \ \boldsymbol{c}^T\boldsymbol{x} \\ &\text{subject to } A\boldsymbol{x} = \boldsymbol{b} \\ &\qquad \boldsymbol{x} \geq 0. \end{aligned} \tag{1}$$

The dual problem of LP

$$\begin{aligned} &\max_{\lambda,\boldsymbol{s}} \ \boldsymbol{b}^T\lambda \\ &\text{subject to } A^T\lambda + \boldsymbol{s} = \boldsymbol{c}, \ \boldsymbol{s} \geq 0 \\ &\qquad \boldsymbol{x} \geq 0. \end{aligned} \tag{2}$$

An important property of LP is that the KKT conditions are sufficient for optimality:

$$A^T\lambda + \boldsymbol{s} = \boldsymbol{c}, \tag{3}$$

$$A\boldsymbol{x} = \boldsymbol{b}, \tag{4}$$

$$\boldsymbol{x} \geq 0, \tag{5}$$

$$\boldsymbol{s} \geq 0, \tag{6}$$

$$x_i s_i = 0, \ i = 1, 2, \cdots, n. \tag{7}$$

Geometry of Feasible Set.

Definition 1.1. *A vector* $\boldsymbol{x}$ *is a basic feasible point if it is feasible and if there exists a subset* $\mathcal{B}$ *of the index set* $\{1, \cdots, n\}$ *such that*

- *$\mathcal{B}$ contains exactly m indices;*
- *$i \notin \mathcal{B} \to x_i = 0$ (that is, the bound $x_i \geq 0$ can be inactive only if $i \in \mathcal{B}$);*
- *The $m \times m$ matrix B defined by*
$$B = [A_i]_{i \in \mathcal{B}}$$
is non-singular, where A_i is the i-th column of A.

A set $\mathcal{B}$ satisfying these properties is called a basis *for the problem (1). The corresponding matrix B is called the* basis matrix.

Theorem 1.1. • *If (1) has a nonempty feasible region, then there is at least one basic feasible point;*

- *If (1) has solutions, then at least one such solution is a basic optimal point.*
- *If (1) is feasible and bounded, then it has an optimal solution.*

Theorem 1.2. *All basic feasible points for (1) are vertices of the feasible polytope $\{\boldsymbol{x} | A\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq 0\}$, and vice versa.*

Definition 1.2. *A basis $\mathcal{B}$ is said to be degenerate if $x_i = 0$ for some $i \in \mathcal{B}$, where $\boldsymbol{x}$ is the basic feasible solution corresponding to $\mathcal{B}$. A linear program is said to be degenerate if it has at least one degenerate basis.*

Two-Phase Procedure for the Simplex Method. The first phase solves the following linear program to obtain an initial solution:

$$\min \boldsymbol{e}^T \boldsymbol{z} \quad \text{subject to} \quad A\boldsymbol{x} + E\boldsymbol{z} = \boldsymbol{b}, \ (\boldsymbol{x}, \boldsymbol{z}) \geq 0, \tag{8}$$

where $\boldsymbol{z} \in \mathbb{R}^m$, $\boldsymbol{e} = (1, \cdots, 1)^T$, and E is a diagonal matrix whose diagonal elements are

$$E_{jj} = 1, \text{ if } b_j \geq 0, \quad E_{jj} = -1, \text{ if } b_j = 0.$$

The nice thing about this formulation is that there exists a very simple basic feasible solution for (8):

$$\boldsymbol{x} = 0, \quad z_j = |b_j|, j = 1, 2, \cdots, m.$$

The second phase solves the following linear program:

$$\min \boldsymbol{c}^T \boldsymbol{x} \quad \text{subject to} \quad A\boldsymbol{x} + \boldsymbol{z} = \boldsymbol{b}, \ \boldsymbol{x} \geq 0, \quad 0 \geq \boldsymbol{z} \quad 0. \tag{9}$$

It is easy to modify the simplex method for solving (9).

2 Interior Point Method (Primal-Dual Methods)

2.1 Outline

Primal-dual methods find solutions $(\boldsymbol{x}^\star, \lambda^\star, \boldsymbol{s}^\star)$ of this system by applying variants of Newton's method to the three equalities (3),(4) and (7) and modifying the search directions and step lengths so that the inequalities $(\boldsymbol{x}, \boldsymbol{s}) \geq 0$ are satisfied strictly at every iteration. The equations (3),(4) and (7) are linear or only mildly nonlinear and so are not difficult to solve by themselves. However, the problem becomes much more difficult when we add the nonnegativity requirement $(\boldsymbol{x}, \boldsymbol{s}) \geq 0$, which gives rise to all the complications in the design and analysis of interior-point methods.

To derive primal-dual interior-point methods we restate the optimality conditions in a slightly different form by means of a mapping F from $\mathbb{R}^{2n+m}$ to $\mathbb{R}^{2n+m}$:

$$F(\boldsymbol{x}, \lambda, \boldsymbol{s}) = 0, \tag{10}$$
$$(\boldsymbol{x}, \boldsymbol{s}) \geq 0, \tag{11}$$

where

$$X = \text{diag}(x_1, x_2, \cdots, x_n), \quad S = \text{diag}(s_1, s_2, \cdots, s_n),$$

and $\boldsymbol{e} = (1, 1, \cdots, 1)^T$. Primal-dual methods generate iterates $(\boldsymbol{x}^k, \lambda^k, \boldsymbol{x}^k)$ that satisfy the bounds (11) strictly, that is, $\boldsymbol{x}^k > 0$ and $\boldsymbol{s}^k > 0$. This property is the origin of the term interior-point. By respecting these bounds, the methods avoid spurious solutions, that is, points that satisfy $F(\boldsymbol{x}, \lambda, \boldsymbol{s}) = 0$ but not $(\boldsymbol{x}, \boldsymbol{s}) \geq 0$. Like most iterative algorithms in optimization, primal-dual interior-point methods have two basic ingredients: a procedure for determining the step and a measure of the desirability of each point in the search space. An important component of the measure of desirability is the average value of the pairwise products $x_i s_i, i = 1, 2, \cdots, n$, which are all positive when $\boldsymbol{x} > 0$ and $\boldsymbol{s} > 0$. This quantity is known as the duality measure and is defined as follows:

$$\mu = \frac{1}{n} \sum_{i=1}^{n} x_i s_i = \frac{\boldsymbol{x}^T \boldsymbol{s}}{n}. \tag{12}$$

The procedure for determining the search direction has its origins in Newton's method for the nonlinear equations (10). Newton's method forms a linear model for F around the current point and obtains the search direction $(\delta \boldsymbol{x}, \delta \lambda, \delta \boldsymbol{s})$ by solving the following system of linear equations:

$$J(\boldsymbol{x}, \lambda, \boldsymbol{s}) \begin{bmatrix} \delta \boldsymbol{x} \\ \delta \lambda \\ \delta \boldsymbol{s} \end{bmatrix} = -F(\boldsymbol{x}, \lambda, \boldsymbol{s}),$$

where J is the Jacobian of F. If we use the notation $\boldsymbol{r}_c$ and $\boldsymbol{r}_b$ for the first two block rows in F, that is,

$$\boldsymbol{r}_b = A\boldsymbol{x} - \boldsymbol{b}, \quad \boldsymbol{r}_c = A^T \lambda + \boldsymbol{s} - \boldsymbol{c}, \tag{13}$$

we can write the Newton equations as follows:

$$\begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \delta \boldsymbol{x} \\ \delta \lambda \\ \delta \boldsymbol{s} \end{bmatrix} = \begin{bmatrix} -\boldsymbol{r}_c \\ -\boldsymbol{r}_b \\ -XS\boldsymbol{e} \end{bmatrix}. \tag{14}$$

Usually, a full step along this direction would violate the bound $(\boldsymbol{x}, \boldsymbol{s}) \geq 0$, so we perform a line search along the Newton direction and define the new iterate as

$$(\boldsymbol{x}, \lambda, \boldsymbol{s}) + \alpha(\delta \boldsymbol{x}, \delta \lambda, \delta \boldsymbol{s}),$$

for some line search parameter $\alpha \in (0, 1]$. We often can take only a small step along this direction $\alpha << 1$ before violating the condition $(\boldsymbol{x}, \boldsymbol{s}) > 0$. Hence, the pure Newton direction (14), sometimes known as the affine scaling direction, often does not allow us to make much progress toward a solution. Most primal-dual methods use a less aggressive Newton direction, one that does not aim directly for a solution but rather for a point whose pairwise products $x_i s_i$ are reduced to a lower average value – not all the way to zero. Specifically, we take a Newton step toward the a point for which $x_i s_i = \sigma \mu$, where μ is the current duality measure and $\sigma \in [0, 1]$ is the reduction factor that we wish to achieve in the duality measure on this step. The modified step equation is then

$$\begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \delta \boldsymbol{x} \\ \delta \lambda \\ \delta \boldsymbol{s} \end{bmatrix} = \begin{bmatrix} -\boldsymbol{r}_c \\ -\boldsymbol{r}_b \\ -XS\boldsymbol{e} + \sigma \mu \boldsymbol{e} \end{bmatrix}. \tag{15}$$

We call σ the centering parameter, for reasons to be discussed later in this class. When $\sigma > 0$, it usually is possible to take a longer step α along the direction defined by (15) before violating the bounds $(\boldsymbol{x}, \boldsymbol{s}) \geq 0$. At this point, we have specified most of the elements of a path-following primal-dual interior-point method.

The choices of centering parameter σ_k and step length α_k are crucial to the performance of the method. Techniques for controlling these parameters, directly and indirectly, give rise to a wide variety of methods with diverse properties.