Primal-Dual Interior Point algorithms for Linear Programming

George Tzallas-Regas

Interior point methods were widely used in the past in the form of barrier methods. In linear programming, the *simplex method* dominated, mainly due to inefficiencies of barrier methods. Interior point methods became quit popular again after 1984, when Karmarkar announced a fast *polynomial-time* interior method for nonlinear programming [Karmarkar, 1984]. In this section we present *primal-dual interior point methods* for linear programming.

1 Linear Programming and Optimality Conditions

In linear programming, the problem to solve in *standard form* is

$$\begin{array}{ll} \underset{\mathbf{x}}{\text{minimize}} & \mathbf{c}^T\mathbf{x} \\ \text{subject to} & \mathbf{A}\mathbf{x} = \mathbf{b} \qquad \mathbf{x} \geq \mathbf{0}, \end{array} \tag{PP}$$

where $\mathbf{c}, \mathbf{x} \in \mathbb{R}^n$, $\mathbf{b} \in \mathbb{R}^m$ and $\mathbf{A}$ is an $m \times n$ matrix. This problem is called the *primal problem*. Associated with it, is the *dual problem*, which can be formulated as

$$\begin{array}{ll} \underset{\mathbf{y}}{\text{maximize}} & \mathbf{b}^T\mathbf{y} \\ \text{subject to} & \mathbf{A}^T\mathbf{y} \leq \mathbf{c}, \end{array} \tag{1}$$

or, in standard form

$$\begin{array}{ll} \underset{(\mathbf{y},\mathbf{s})}{\text{maximize}} & \mathbf{b}^T\mathbf{y} \\ \text{subject to} & \mathbf{A}^T\mathbf{y} + \mathbf{s} = \mathbf{c}, \qquad \mathbf{s} \geq \mathbf{0}, \end{array} \tag{DP}$$

where $\mathbf{y} \in \mathbb{R}^m$ and $\mathbf{s} \in \mathbb{R}^n$. For some more terminology, vector $\mathbf{s}$ is called the *dual slack*. The solution $\mathbf{y}^*$ of the dual is the Lagrange multiplier of the primal problem, and vice versa. The quantity

$$\mathbf{c}^T\mathbf{x} - \mathbf{b}^T\mathbf{y} \tag{DG}$$

is called the *duality gap*, it is nonnegative, and it is zero only at $(\mathbf{x}^*, \mathbf{y}^*)$. The value of the duality gap is a common termination criterion in interior LP methods.

The optimality conditions can be written as

$$\begin{aligned} \mathbf{A}^T\mathbf{y} + \mathbf{s} &= \mathbf{c} && \text{(2a)} \\ \mathbf{A}\mathbf{x} &= \mathbf{b} && \text{(2b)} \\ \mathbf{X}\mathbf{S}\mathbf{e} &= \mathbf{0} && \text{(2c)} \\ (\mathbf{x}, \mathbf{s}) &\geq \mathbf{0} && \text{(2d)} \end{aligned}$$

where $\mathbf{X}$ is used to denote the diagonal matrix with diagonal $\mathbf{x}$

$$\mathbf{X} = \begin{pmatrix} [x]_1 & & & \\ & [x]_2 & & \\ & & \ddots & \\ & & & [x]_n \end{pmatrix}$$

and analogous notation is used for other quantities. Vector $\mathbf{e}$ is used to denote a vector of all ones, whose dimension usually varies depending on the context. Condition (2a) is known as *dual feasibility*, condition (2b) is known as *primal feasibility*, and (2c) corresponds to *complementarity*.

A point $\mathbf{x}^*$ is a solution of problem PP if and only if there exist vectors $\mathbf{s}^*$ and $\mathbf{y}^*$ such that conditions (2) hold for $(\mathbf{x}, \boldsymbol{y}, \boldsymbol{s}) = (\mathbf{x}^*, \boldsymbol{y}^*, \boldsymbol{s}^*)$. On the other hand, a point $(\mathbf{y}^*, \mathbf{s}^*)$ is a solution of problem DP if and only if there exists vector $\mathbf{x}^*$ such that conditions (2) hold for $(\mathbf{x}, \boldsymbol{y}, \boldsymbol{s}) = (\mathbf{x}^*, \boldsymbol{y}^*, \boldsymbol{s}^*)$. Vector $(\mathbf{x}^*, \boldsymbol{y}^*, \boldsymbol{s}^*)$ is called a *primal-dual solution*.

2 Duality gap

The duality gap (DG) is the difference between the objective function of the primal problem (PP) and the dual problem (DP). The sign of the duality gap tells us about the relationship between these two problems. The sign of the duality gap is explained by *Duality theory* in terms of the feasible set and the solution set for the two problems.

For instance, given any feasible vectors $\bar{\mathbf{x}}$ for (PP) and $(\bar{\mathbf{y}}, \bar{\mathbf{s}})$ for (DP), we have that

$$\mathbf{b}^T\bar{\mathbf{y}} \leq \mathbf{c}^T\bar{\mathbf{x}} \tag{3}$$

This is easy to verify from the KKT conditions (2). From (2d) we have that $[\bar{x}]_i, [\bar{s}]_i \geq 0$ therefore

$$0 \leq \bar{\mathbf{s}}^T\bar{\mathbf{x}}.$$

From (2a) we have that $\bar{\mathbf{s}} = \mathbf{c} - \mathbf{A}^T\bar{\mathbf{y}}$, therefore

$$\bar{\mathbf{s}}^T\bar{\mathbf{x}} = \left(\mathbf{c} - \mathbf{A}^T\bar{\mathbf{y}}\right)^T\bar{\mathbf{x}} = \mathbf{c}^T\bar{\mathbf{x}} - \bar{\mathbf{y}}^T(\mathbf{A}\bar{\mathbf{x}}).$$

Eq. (2b) states that $\mathbf{A}\bar{\mathbf{x}} = \mathbf{b}$, therefore

$$\bar{\mathbf{y}}^T(\mathbf{A}\bar{\mathbf{x}}) = \bar{\mathbf{y}}^T\mathbf{b} = \mathbf{b}^T\bar{\mathbf{y}}.$$

Putting everything together we obtain

$$0 \leq \mathbf{c}^T\bar{\mathbf{x}} - \mathbf{b}^T\bar{\mathbf{y}},$$

which is inequality (3).

If we have a point $(\mathbf{x}^*, \boldsymbol{y}^*, \boldsymbol{s}^*)$ that satisfies (2), it follows from (2c) that

$$0 = \mathbf{s}^{*T}\mathbf{x}^* = \mathbf{c}^T\mathbf{x}^* - \mathbf{b}^T\mathbf{y}^*,$$

therefore in the optimal solution the optimality gap vanishes and the optimal values for problems (PP) and (DP) coincide.

3 Primal Dual variants

Primal-dual interior-point methods find primal-dual solutions $(\mathbf{x}^*, \boldsymbol{y}^*, \boldsymbol{s}^*)$ by applying variants of Newton's method to conditions (2a)–(2c), and modifying search directions and step lengths, so that (2d) is strictly satisfied at each iteration. It is more convenient to write the optimality conditions in terms of a mapping, namely $\mathbf{F} : \mathbb{R}^{2n+m} \rightarrow \mathbb{R}^{2n+m}$:

$$\mathbf{F}(\mathbf{x}, \boldsymbol{y}, \boldsymbol{s}) = \begin{bmatrix} \mathbf{A}^T\mathbf{y} + \mathbf{s} - \mathbf{c} \\ \mathbf{A}\mathbf{x} - \mathbf{b} \\ \mathbf{XSe} \end{bmatrix} = \mathbf{0}, \qquad (\mathbf{x}, \mathbf{s}) \geq \mathbf{0} \tag{4}$$

Newton's method forms a linear model for $\mathbf{F}$ around the current iterate $(\mathbf{x}^{(k)}, \boldsymbol{y}^{(k)}, \boldsymbol{s}^{(k)})$, and obtains a search direction $(\delta\mathbf{x}, \delta\mathbf{y}, \delta\mathbf{s})$ by solving

$$\mathbf{F}'(\mathbf{x}, \boldsymbol{y}, \boldsymbol{s}) \begin{bmatrix} \delta\mathbf{x} \\ \delta\mathbf{y} \\ \delta\mathbf{s} \end{bmatrix} = -\mathbf{F}(\mathbf{x}, \boldsymbol{y}, \boldsymbol{s}), \tag{5}$$

where $\mathbf{F}'$ is the Jacobian of $\mathbf{F}$. Newton's method is applicable fot systems of equalities not inequalities. If we assume that we have strictly feasible iterates, the previous equation becomes

$$\begin{bmatrix} \mathbf{0} & \mathbf{A}^T & \mathbf{I} \\ \mathbf{A} & \mathbf{0} & \mathbf{0} \\ \mathbf{S} & \mathbf{0} & \mathbf{X} \end{bmatrix} \begin{bmatrix} \delta\mathbf{x} \\ \delta\mathbf{y} \\ \delta\mathbf{s} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ -\mathbf{XSe} \end{bmatrix} \tag{6}$$

and the next feasible iterate is obtained as

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

for a line search parameter $\alpha \in (0, 1]$. The line search parameter is chosen so that the next iterate is kept strictly feasible. If the next iterate is strictly feasible, then we can apply Newton's method iteratively.

4 Central path

The *central path* is an important aspect of interior point methods, that will help us to establish a general algorithm for primal-dual methods.

A central path is an arc of strictly feasible points that is parameterized by a positive scalar ρ, and each point $(\mathbf{x}(\rho), \boldsymbol{\lambda}(\rho), \mathbf{s}(\rho))$ that belongs to the path, solves

$$\mathbf{A}^T\mathbf{y} + \mathbf{s} = \mathbf{c} \tag{7a}$$

$$\mathbf{Ax} = \mathbf{b} \tag{7b}$$

$$[x]_i \, [s]_i = \rho, \quad i = 1, \dots, n \tag{7c}$$

$$(\mathbf{x}, \mathbf{s}) > \mathbf{0} \tag{7d}$$

Eqs. (7) differ from Eqs. (2) only in the complementarity condition (7c), where the condition now is that products $[x]_i \, [s]_i$ have the same value for all $i = 1, \ldots, n$. But as ρ goes to zero, then Eq. (7) approximate Eq. (2). Therefore points on the central path converge to a primal-dual solution of the linear problem, as $\rho \to 0^-$. The use of the parameter ρ, and the good behaviour we get from the central path when it approaches zero, is reminiscent of barrier methods, and this is not a coincidence.

5 The Barrier Problem

Let's go back to the primal problem (PP). The nonnegativity constraints $\mathbf{x} \geq 0$ can be replaced, by adding a *barrier term* in the objective function, that looks like

$$B(\mathbf{x}) = \sum_{j=1}^{n} \log x_j$$

This barrier term is called the logarithmic barrier term, it is finite as long as x_j is positive, and approaches negative infinity as x_j approaches zero. The primal problem can now be rewritten as

$$\begin{array}{ll} \underset{\mathbf{x}}{\text{maximize}} & \mathbf{c}^T\mathbf{x} - \rho B(\mathbf{x}) \\ \text{subject to} & \mathbf{A}\mathbf{x} = \mathbf{b}. \end{array} \tag{PLBP}$$

for $\mathbf{x} > \mathbf{0}$. The Lagrangian for this problem is

$$\mathcal{L}(\mathbf{x}, \mathbf{y}) = \mathbf{c}^T\mathbf{x} - \rho B(\mathbf{x}) + \mathbf{y}^T(\mathbf{b} - \mathbf{A}x).$$

The KKT conditions for problem (PLBP) are

$$\begin{aligned} \mathbf{c} - \rho\mathbf{X}^{-1}\mathbf{e} - \mathbf{A}^T\mathbf{y} &= \mathbf{0} \\ \mathbf{b} - \mathbf{A}\mathbf{x} &= \mathbf{0} \end{aligned}$$

for $\mathbf{x} > 0$. If we introduce an extra vector defined as $\mathbf{s} = \rho\mathbf{X}^{-1}\mathbf{e}$, we can rewrite the KKT conditions as

$$\begin{aligned} \mathbf{A}^T\mathbf{y} + \mathbf{s} &= \mathbf{c} \\ \mathbf{A}\mathbf{x} &= \mathbf{b} \\ \mathbf{s} &= \rho\mathbf{X}^{-1}\mathbf{e} \end{aligned}$$

or, if we mutliply the last equation to the left by $\mathbf{X}$

$$\mathbf{A}^T\mathbf{y} + \mathbf{s} = \mathbf{c} \tag{8a}$$

$$\mathbf{A}\mathbf{x} = \mathbf{b} \tag{8b}$$

$$\mathbf{X}\mathbf{S}\mathbf{e} = \rho\mathbf{e} \tag{8c}$$

for $(\mathbf{x}, \mathbf{s}) > \mathbf{0}$. This system of equations is the same as Eqs. (7).

6 Primal-Dual Interior Point Algorithm

The Newton equations (6) have been derived for the first order optimality conditions of problem (PP), under the assumption that the iterates $(\mathbf{x}, \mathbf{s})$ are strictly feasible. In a similar fashion, the Newton equations for the logarithmic barrier reformulation of problem (PP) are

$$\begin{bmatrix} \mathbf{0} & \mathbf{A}^T & \mathbf{I} \\ \mathbf{A} & \mathbf{0} & \mathbf{0} \\ \mathbf{S} & \mathbf{0} & \mathbf{X} \end{bmatrix} \begin{bmatrix} \delta\mathbf{x} \\ \delta\mathbf{y} \\ \delta\mathbf{s} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ -\mathbf{XSe} + \tau\gamma\mathbf{e} \end{bmatrix}. \tag{9}$$

In these equations we have used $\rho = \tau\gamma$, where $\tau \in [0, 1]$ is a *centering parameter* and

$$\gamma = \frac{\mathbf{x}^T\mathbf{s}}{n} \tag{DM}$$

is a *duality measure*, *i.e.* it measures the average value of the complementarity products $[x]_i\,[s]_i$.

Algorithm 1 General primal-dual interior point method

Determine $(\mathbf{x}^{(0)}, \boldsymbol{y}^{(0)}, \boldsymbol{s}^{(0)})$ strictly feasible
Set $k := 0$
repeat
    Set $\tau^{(k)} \in [0, 1]$ and $\gamma^{(k)} = \dfrac{\mathbf{x}^T\mathbf{s}}{n}$
    Solve system (9) to obtain $(\delta\mathbf{x}^{(k)}, \delta\mathbf{y}^{(k)}, \delta\mathbf{s}^{(k)})$
    Set

$$(\mathbf{x}^{(k+1)}, \boldsymbol{y}^{(k+1)}, \boldsymbol{s}^{(k+1)}) = (\mathbf{x}^{(k)}, \boldsymbol{y}^{(k)}, \boldsymbol{s}^{(k)}) + \alpha^{(k)}(\delta\mathbf{x}^{(k)}, \delta\mathbf{y}^{(k)}, \delta\mathbf{s}^{(k)})$$

    choosing $\alpha^{(k)}$ so that $(\mathbf{x}^{(k+1)}, \mathbf{s}^{(k+1)}) > \mathbf{0}$.
    Set $k := k + 1$
until Convergence

Algorithm (1) gives a general primal-dual interior-point method. There are two major variants of this algorithm

1. *Path-following methods* restrict iterates to a neighbourhood of the central path, and follow the central path to a solution of the linear problem. This class of algorithms include

    - *short-step path-following methods.* They choose values of the centering parameter τ close to 1, and make slow progress towards the solution
    - *long-step path-following methods.* The centering parameter τ varies more, and make more rapid progress than their short-step counterparts
    - *predictor-corrector path-following methods.* They alternate the value of the centering parameter τ, to obtain a *predictor step*, ($\tau = 0$) which moves along the *affine scaling direction*, and then a corrector step ($\tau = 1$), which moves along the *centering direction*.

2. *Potential reduction methods* which are more or less the same as path-following methods, but do not follow the central path explicitly

We do not give further details on the methods and the aforementioned directions. Bertsekas [1995, Section 4.1 and Section 4.4], Vanderbei [1998, Part 3] provide a detailed treatment of path-following methods, and Wright [1997, chapters 4–5] also presents both classes of algorithms in an excellent manner. Wright [1992] has written an excellent monograph on interior methods, and El-Bakry, Tapia, Tsuchiya, and Zhang [1996] have given a formulation of interior-point methods for nonlinear programming.

References

Dimitri P. Bertsekas. *Nonlinear Programming.* Athena Scientific, Belmont, MA, 1995. 2nd edition 1999.

A. S. El-Bakry, Richard A. Tapia, T. Tsuchiya, and Yin Zhang. On the formulation and theory of the Newton interior-point method for nonlinear programming. *Journal of Optimization Theory and Applications*, 89(3): 507–541, 1996. ISSN 0022-3239.

Narendra Karmarkar. A new polynomial-time algorithm for linear programming. *Combinatorica*, 4(4):373–395, 1984.

Robert J. Vanderbei. *Linear Programming: Foundations and Extensions*. Kluwer Academic Publishers, Dordrecht, 1998.

Margaret H. Wright. Interior methods for constrained optimization. *Acta Numerica*, pages 341–407, 1992. URL `ftp://netlib.att.com/netlib/att/cs/doc/91/4-10.ps.Z`.

Stephen J. Wright. *Primal-Dual Interior-Point Methods*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 1997. ISBN 0-89871-382-X.