PRIMAL-DUAL INTERIOR-POINT ALGORITHMS: AN INTRODUCTION *

Spring 1995

We consider linear programming problems in the following primal (P) and dual (D) forms:

$$
\begin{array}{llrcl}
(P) & \text{maximize} & c^T x & & \\
 & & Ax & = & b, \\
 & & x & \geq & 0,
\end{array}
$$

$$
\begin{array}{llrcrcl}
(D) & \text{minimize} & b^T y & & & & \\
 & & A^T y & - & s & = & c, \\
 & & & & s & \geq & 0,
\end{array}
$$

where $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, and $c \in \mathbb{R}^n$ (all vectors are column vectors). Note that we have explicitly included the slack variables s. We will assume $rank(A) = m$ (if not, we can do Gaussian elimination). For now, we will also assume that there exist **interior** solutions for both problems, i.e. there exist $\bar{x}$ and $(\bar{y}, \bar{s})$ such that $A\bar{x} = b$, $\bar{x}_j > 0$ for all j and $A^T\bar{y} - \bar{s} = c$, $\bar{s}_j > 0$ for all j. So, an **interior** solution is a feasible solution which satisfies all the inequality constraints strictly.

By the weak duality relation we have

$$b^T\bar{y} - c^T\bar{x} \geq 0.$$

In fact, we can write the duality gap in terms of $\bar{x}$ and $\bar{s}$:

$$b^T\bar{y} - c^T\bar{x} \;=\; \bar{x}^T A^T \bar{y} - c^T\bar{x} \;=\; \bar{x}^T(A^T\bar{y} - c) \;=\; \bar{x}^T\bar{s} \;>\; 0.$$

If $\bar{x}$ and $\bar{s}$ had been optimal solutions, we would have had $\bar{x}^T\bar{s} = 0$. Recall the necessary and sufficient conditions for optimality:

*Please address your questions and comments to e-mail: ltuncel@math.uwaterloo.ca.

$$
\begin{aligned}
Ax &= b, \quad x \geq 0, \quad \text{(primal feasibility)} && (1)\\
A^T y - s &= c, \quad s \geq 0, \quad \text{(dual feasibility)} && (2)\\
x_j s_j &= 0 \text{ for all } j \text{ (complementary slackness).} && (3)
\end{aligned}
$$

Now, we would like to find a new solution x, y, s so that x is feasible for (P), (y, s) is feasible for (D) and the duality gap $b^T y - c^T x = x^T s$ is smaller than the current duality gap: $b^T \bar{y} - c^T \bar{x} = \bar{x}^T \bar{s}$. Recall that primal simplex algorithms keep primal feasibility and complementary slackness and strive for dual feasibility. Dual simplex algorithms keep dual feasibility and complementary slackness and strive for primal feasibility. Here, we are keeping primal feasibility and dual feasibility and striving for complementary slackness. So, we would like to find directions $\mathbf{dx}$ in primal and $\mathbf{dy}$, $\mathbf{ds}$ in the dual such that for some **step size** α,

1. $\bar{x} + \alpha\mathbf{dx}$ is feasible in (P)
2. $\bar{y} + \alpha\mathbf{dy}$ and $\bar{s} + \alpha\mathbf{ds}$ are feasible in (D)
3. $(\bar{x} + \alpha\mathbf{dx})^T(\bar{s} + \alpha\mathbf{ds}) < \bar{x}^T\bar{s}$.

The first item above implies $A\mathbf{dx} = 0$, the second item implies $A^T\mathbf{dy} - \mathbf{ds} = 0$, and the third one implies $(\bar{x}_j + \alpha\mathbf{dx}_j)(\bar{s}_j + \alpha\mathbf{ds}_j) = 0$. The first two systems of equations are linear in $\mathbf{dx}$, $\mathbf{dy}$, and $\mathbf{ds}$; however, the third one is

$$
\alpha\bar{s}_j\mathbf{dx}_j + \alpha\bar{x}_j\mathbf{ds}_j + \alpha^2\mathbf{dx}_j\mathbf{ds}_j = -\bar{x}_j\bar{s}_j
$$

which is quadratic (we have the term $\mathbf{dx}_j\mathbf{ds}_j$). Of course, if we could solve these three groups of equations at once, we would have optimal solutions of (P) and (D). Since we do not know how to solve these equations (without solving the linear programming problem), we will ignore the quadratic term (yes, the idea is very closely related to Newton's Method for finding zeros of nonlinear functions) and try to solve $\bar{s}_j\mathbf{dx}_j + \bar{x}_j\mathbf{ds}_j = -\bar{x}_j\bar{s}_j$ instead.

So, we solve the following system of linear equations. Note that the last group of equalities is trying to force a solution $\mathbf{dx}$, $\mathbf{ds}$ such that $(\bar{x} + \alpha\mathbf{dx})^T(\bar{s} + \alpha\mathbf{ds})$ is close to zero.

$$
\begin{aligned}
A\mathbf{dx} &= 0, && (4)\\
A^T\mathbf{dy} - \mathbf{ds} &= 0, && (5)\\
\bar{s}_j\mathbf{dx}_j + \bar{x}_j\mathbf{ds}_j &= -\bar{x}_j\bar{s}_j, \quad \text{for all } j. && (6)
\end{aligned}
$$

This is a system of linear equations and we know how to solve it efficiently (e.g. by Gaussian Elimination). In fact, under our assumptions ($rank(A) = m$ and $\bar{x} > 0$, $\bar{s} > 0$), this linear system has a unique solution.

One point we neglected so far is the nonnegativity constraints. Since we would like to iterate until we reach an optimal solution, we should pick α such that $\bar{x} + \alpha\mathbf{dx} > 0$ and $\bar{s} + \alpha\mathbf{ds} > 0$. We can choose, for instance,

$$\alpha = 0.9 \min\left\{\frac{x_j}{-\mathbf{dx}_j}, \frac{s_j}{-\mathbf{ds}_j} : \mathbf{dx}_j < 0,\ \mathbf{ds}_j < 0, \right\}.$$

(In practice, we may even want to choose different step sizes in the primal and the dual problems, e.g. $\alpha_P = 0.9 \min\left\{\frac{x_j}{-\mathbf{dx}_j} : \mathbf{dx}_j < 0, \right\}$ and $\alpha_D = 0.9 \min\left\{\frac{s_j}{-\mathbf{ds}_j} : \mathbf{ds}_j < 0, \right\}$ and then we update $(\bar{x}, \bar{y}, \bar{s})$ using $(\bar{x} + \alpha_P\mathbf{dx}, \bar{y} + \alpha_D\mathbf{dy}, \bar{s} + \alpha_D\mathbf{ds})$.) There are many different ways of choosing the step size α (some have practical and/or theoretical advantages) see the references in the last page.

Now, we show that the duality gap decreases from iteration to iteration. Here we are able to show that the decrease is proportional to $(1-\alpha)$. That is, when the step size is large, we get a large decrease in the duality gap. Indeed, if $\bar{x} + \mathbf{dx} \geq 0$ (i.e. $\alpha = 1$ yields a feasible solution in (P)) and $\bar{s} + \mathbf{ds} \geq 0$ (i.e. $\alpha = 1$ also yields a feasible solution in (D)), then they ($\bar{x} + \mathbf{dx}$ and $\bar{s} + \mathbf{ds}$) will be optimal. Let's calculate the new duality gap:

$$\begin{aligned}
\text{new duality gap} &= \sum_{j=1}^{n} (\bar{x}_j + \alpha\mathbf{dx}_j)(\bar{s}_j + \alpha\mathbf{ds}_j), &(7)\\
&= \sum_{j=1}^{n} \bar{x}_j\bar{s}_j + \alpha\sum_{j=1}^{n}(\bar{s}_j\mathbf{dx}_j + \bar{x}_j\mathbf{ds}_j) + \alpha^2\sum_{j=1}^{n}\mathbf{dx}_j\mathbf{ds}_j. &(8)
\end{aligned}$$

¿From (5) we have $\mathbf{ds} = A^T\mathbf{dy}$. So,

$$\sum_{j=1}^{n}\mathbf{dx}_j\mathbf{ds}_j = (\mathbf{dx})^T A^T \mathbf{dy} = (\mathbf{dy})^T(A\mathbf{dx}) = 0(\text{since } A\mathbf{dx} = 0).$$

Also note that using (6), we have $(\bar{s}_j\mathbf{dx}_j + \bar{x}_j\mathbf{ds}_j) = -\bar{x}_j\bar{s}_j$. Hence,

$$\begin{aligned}
(\text{new duality gap}) &= \sum_{j=1}^{n}\bar{x}_j\bar{s}_j + \alpha\sum_{j=1}^{n}(\bar{s}_j\mathbf{dx}_j + \bar{x}_j\mathbf{ds}_j) &(9)\\
&= \sum_{j=1}^{n}\bar{x}_j\bar{s}_j - \alpha\sum_{j=1}^{n}\bar{x}_j\bar{s}_j &(10)\\
&= (1-\alpha)\sum_{j=1}^{n}\bar{x}_j\bar{s}_j &(11)\\
&= (1-\alpha)\ (\text{current duality gap}). &(12)
\end{aligned}$$

Note that unless $\alpha = 1$ at some iteration, we never reach an optimal solution. So, we might choose to stop iterating when the duality gap is below some small positive number ϵ (say 10^{-8}).

Now, we can state the algorithm (this is called primal-dual affine-scaling algorithm):

Algorithm:

Given (x^o, y^o, s^o) so that $x^o > 0$ and $s^o > 0$, satisfying the equality constraints, set $k := 0$.

While $((x^k)^T s^k > \epsilon)$ do

$(\bar{x}, \bar{y}, \bar{s}) := (x^k, y^k, s^k)$

compute $(\mathbf{dx}, \mathbf{dy}, \mathbf{ds})$ from (4)-(6)

if $(x + \mathbf{dx}, y + \mathbf{dy}, s + \mathbf{ds})$ is feasible, let $(x^{k+1}, y^{k+1}, s^{k+1}) := (x + \mathbf{dx}, y + \mathbf{dy}, s + \mathbf{ds})$ and STOP (in this case, $(x^{k+1}, y^{k+1}, s^{k+1})$ is an optimal solution of (P) and (D)).

choose step size $\alpha_k \in (0, 1)$ such that $x + \alpha_k \mathbf{dx} > 0$, and $s + \alpha_k \mathbf{ds} > 0$

set $(x^{k+1}, y^{k+1}, s^{k+1}) := (x + \alpha_k \mathbf{dx}, y + \alpha_k \mathbf{dy}, s + \alpha_k \mathbf{ds})$

$k := k + 1$

end{while}

It turns out that the solution $\mathbf{dx}, \mathbf{dy}$, and $\mathbf{ds}$ of the linear system (which is solved at each iteration) can be written explicitly in terms of A and the current solution $\bar{x}$ and $\bar{s}$ (note that $A\bar{x} = b$ and $A^T\bar{y} + \bar{s} = c$). Define X to be the diagonal matrix with entries $\bar{x}_1$, $\bar{x}_2$, $\ldots$, $\bar{x}_n$ down the diagonal (and S is the diagonal matrix with entries $\bar{s}_1$, $\bar{s}_2$, $\ldots$, $\bar{s}_n$ down the diagonal). Then the solution vectors $\mathbf{dx}, \mathbf{dy}$, and $\mathbf{ds}$ are given by the following formulas:

$$\mathbf{dy} = -(AXS^{-1}A^T)^{-1}b, \tag{13}$$

$$\mathbf{ds} = A^T\mathbf{dy}, \tag{14}$$

$$\mathbf{dx} \;=\; -\bar{x} - XS^{-1}\mathbf{ds}. \tag{15}$$

Example Consider the LP given in the form of (P) with

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 1 & 1 & 0 \\ 3 & 1 & -1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \end{pmatrix} \qquad \mathbf{b} \;=\; \begin{pmatrix} 5 \\ 8 \\ 3 \end{pmatrix} \mathbf{c} \;=\; \begin{pmatrix} 4 \\ 2 \\ 1 \\ 1 \\ -3 \end{pmatrix}$$

Note that we can choose,
$\bar{x} \;=\; (1,1,1,1,5)^T$, and $\bar{s} \;=\; (4,9,3,4,4)^T$,
so that $\bar{x}$ is feasible in (P) and $\bar{s}$ is feasible in (D) (there are also ways of starting from infeasible solutions). These solutions are not optimal, as the duality gap, $\bar{x}^T\bar{s} = 40$, is not zero. We compute
$\mathbf{dy} \;=\; -(AXS^{-1}A^T)^{-1}b = (4.0824, -3.2583, -10.0983)^T$.
$\mathbf{ds} \;=\; A^T\mathbf{dy} = (-5.6926, -5.1918, -2.7575, -6.0158, -3.2583)^T$.
$\mathbf{dx} \;=\; -\bar{x} - XS^{-1}\mathbf{ds} = (0.4231, -0.4231, -0.0808, 0.5040, -0.9271)^T$.
We choose $\alpha > 0$ (using the formula in page 4) so that it is as large as possible (without violating or hitting the constraints): $\alpha = 0.6649$. We update our solution:
$\bar{x} \;=\; (1.2812, 0.7188, 0.9463, 1.3349, 4.3839)^T$,
$\bar{y} \;=\; (7.7131, -1.1654, -6.7111)^T$,
$\bar{s} \;=\; (0.2169, 5.5497, 1.1674, 0.0020, 1.8346)^T$.
New duality gap, is $\bar{x}^T\bar{s} = 13.41$. We keep repeating these steps, using current solutions.

Next (iteration 2) we get
$\bar{x} \;=\; (1.5885, 0.4115, 0.4207, 2.1678, 3.2437)^T$,
$\bar{y} \;=\; (8.2153, -1.3773, -7.2153)^T$,
$\bar{s} \;=\; (0.0834, 5.8380, 1.3773, 0.0000001, 1.6227)^T$.
Duality gap: $\bar{x}^T\bar{s} = 8.3776$.

Iteration 3:
$\bar{x} \;=\; (1.9163, 0.0837, 0.0002, 2.9161, 2.1675)^T$,
$\bar{y} \;=\; (9.6372, -1.8748, -8.6372)^T$,
$\bar{s} \;=\; (0.0130, 6.7625, 1.8748, 0.00000002, 1.6227)^T$.
Duality gap: $\bar{x}^T\bar{s} = 3.02998$.

Iteration 4:
$\bar{x} \;=\; (1.98678, 0.01322, 0.0000001, 2.98678, 2.02643)^T$,
$\bar{y} \;=\; (11.5995, -2.5318, -10.5995)^T$,
$\bar{s} \;=\; (0.00407, 8.06770, 2.5318, 0.000000005, 0.46818)^T$.
Duality gap: $\bar{x}^T\bar{s} = 1.06345$.

Iteration 5:
$\bar{x} \;=\; (1.9992, 0.00077, 0.00000000005, 2.99923, 2.00154)^T$,

$\bar{y} = (12.76625, -2.92188, -11.76625)^T,$
$\bar{s} = (0.00060, 8.84437, 2.92188, 0.0000000008, 0.0781)^T.$
Duality gap: $\bar{x}^T\bar{s} = 0.16437$.

Iteration 6:
$\bar{x} = (1.99999, 0.000007, 0.00000000000003, 2.99999, 2.00001)^T,$
$\bar{y} = (12.9936, -2.9979, -11.9936)^T,$
$\bar{s} = (0.000016, 8.9958, 2.9979, 0.00000000002, 0.002130)^T.$
Duality gap: $\bar{x}^T\bar{s} = 0.004$.
Now, I think we have an idea of an optimal basis for the primal. Namely,

$$B = \{1, 4, 5\}.$$

The corresponding basic solution is

$$x^* = (2, 0, 0, 3, 2)^T.$$

Using the complementary slackness theorem we see that the corresponding dual solution is

$$y^* = (13, -3, -12)^T, \quad s^* = (0, 9, 3, 0, 0)^T.$$

Since $c^T x^* = b^T y^*$ (or in other words, $(x^*)^T s^* = 0$), x^* is an optimal solution of (P) and y^* and s^* make an optimal solution of (D).

References

[1] N. Karmarkar, "A new polynomial time algorithm for linear programming," *Combinatorica* 4, pp. 373-395, 1984.

[2] M. Kojima, S. Mizuno, and A. Yoshise, "A Primal-Dual Interior Point Algorithm for Linear Programming," in: N. Megiddo, ed., *Progress in Mathematical Programming, Interior Point and Related Methods* (Springer–Verlag, New York, 1988) pp. 29–47.

[3] I. J. Lustig, R. E. Marsten, and D. F. Shanno, Computation Experience with a Primal-Dual Interior Point Method for Linear Programming, *Linear Algebra and Its Applications* 152 (1991), pp. 191-222.

[4] N. Megiddo, "Pathways to the Optimal Set in Linear Programming," in: N. Megiddo, ed., *Progress in Mathematical Programming, Interior Point and Related Methods* (Springer–Verlag, New York, 1988) pp. 131–158.

[5] R. D. C. Monteiro, I. Adler and M. G. C. Resende, "A Polynomial Time Primal-Dual Affine Scaling Algorithm for Linear and Convex Quadratic Programming and Its Power Series Extension," *Mathematics of Operations Research* 15 (1990) 191–214.

[6] L. Tunçel, Constant Potential Primal-Dual Algorithms: A Framework, *Math. Programming* 66 (1994), pp. 145–159.