Linear Programming

Lecture 07B - Primal-Dual Interior-Point Algorithm

John Bagterp Jørgensen

Department of Applied Mathematics and Computer Sicence
Technical University of Denmark

02612 Constrained Optimization

Learning objectives

- Describe and explain the primal-dual interior-point algorithm for linear programs (in standard form)
- Implement and test primal-dual interior-point algorithms

Nocedal & Wright: Chap. 14

Outline

Optimality Conditions

- Linear Program in Standard Form

$$\min_{x \in \mathbb{R}^n} \quad g'x \tag{1a}$$
$$s.t. \quad Ax = b \tag{1b}$$
$$x \geq 0 \tag{1c}$$

- Lagrange function

$$L(x, \mu, \lambda) = g'x - \mu'(Ax - b) - \lambda' x \tag{2}$$

- Optimality Conditions

$$\nabla_x L(x, \mu, \lambda) = g - A'\mu - \lambda = 0 \tag{3a}$$
$$Ax = b \tag{3b}$$
$$x \geq 0 \tag{3c}$$
$$\lambda \geq 0 \tag{3d}$$
$$x_i \lambda_i = 0 \quad i = 1, 2, \ldots, n \tag{3e}$$

Optimality Conditions

- Notation

$$X = \begin{bmatrix} x_1 & & & \\ & x_2 & & \\ & & \ddots & \\ & & & x_n \end{bmatrix} \quad \Lambda = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} \quad e = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \tag{4}$$

- The complementarity conditions

$$x_i \lambda_i = 0 \quad i = 1, 2, \ldots, n \tag{5}$$

can be expressed as

$$X \Lambda e = 0 \tag{6}$$

Optimality Conditions

- The optimality conditions can be expressed as

$$r_L = \nabla_x L = g - A'\mu - \lambda = 0 \tag{7a}$$
$$r_A = Ax - b = 0 \tag{7b}$$
$$X\Lambda e = 0 \tag{7c}$$
$$x \geq 0, \quad \lambda \geq 0 \tag{7d}$$

- which is the same as

$$F(x, \mu, \lambda) = \begin{bmatrix} g - A'\mu - \lambda \\ Ax - b \\ X\Lambda e \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{8a}$$
$$(x, \lambda) \geq 0 \tag{8b}$$

- Solve $F(x, \mu, \lambda) = 0$ such that $(x, \lambda) \geq 0$ using some Newton-like method.

Newton's Method

- Nonlinear equation

$$F(x,\mu,\lambda) = \begin{bmatrix} g - A'\mu - \lambda \\ Ax - b \\ X\Lambda e \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = 0 \tag{9}$$

- Newton's method

$$J(x,\mu,\lambda) \begin{bmatrix} \Delta x \\ \Delta \mu \\ \Delta \lambda \end{bmatrix} = -F(x,\mu,\lambda) \tag{10}$$

- expressed as

$$\begin{bmatrix} 0 & -A' & -I \\ A & 0 & 0 \\ \Lambda & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \mu \\ \Delta \lambda \end{bmatrix} = - \begin{bmatrix} r_L = g - A'\mu - \lambda \\ r_A = Ax - b \\ X\Lambda e \end{bmatrix} \tag{11}$$

- Step

$$\begin{bmatrix} x \\ \mu \\ \lambda \end{bmatrix} \leftarrow \begin{bmatrix} x \\ \mu \\ \lambda \end{bmatrix} + \alpha \begin{bmatrix} \Delta x \\ \Delta \mu \\ \Delta \lambda \end{bmatrix} \quad (x,\lambda) \geq 0 \tag{12}$$

Primal-Dual Interior Point Methods

Primal-dual interior-point methods modify the basic Newton procedure in two important ways:

1. They bias the search direction toward the interior of the non-negative orthant $(x, \lambda) \geq 0$. This allows us to move further along the search direction before one of the components of (x, λ) becomes negative
2. They keep the components of (x, λ) from moving to close to the boundary of the non-negative orthant.

$$\begin{bmatrix} 0 & -A' & -I \\ A & 0 & 0 \\ \Lambda & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \mu \\ \Delta \lambda \end{bmatrix} = - \begin{bmatrix} g - A'\mu - \lambda \\ Ax - b \\ X\Lambda e \end{bmatrix} = - \begin{bmatrix} r_L \\ r_A \\ X\Lambda e \end{bmatrix} \quad (13)$$

$$\begin{bmatrix} x \\ \mu \\ \lambda \end{bmatrix} \leftarrow \begin{bmatrix} x \\ \mu \\ \lambda \end{bmatrix} + \alpha \begin{bmatrix} \Delta x \\ \Delta \mu \\ \Delta \lambda \end{bmatrix} \quad (x, \lambda) \geq 0 \quad (14)$$

Central Path

- Equations defining the central path

$$
\begin{aligned}
r_L &= g - A'\mu - \lambda = 0 && \text{(15a)} \\
r_A &= Ax - b = 0 && \text{(15b)} \\
x_i\lambda_i &= \tau \quad i = 1, 2, \ldots, n && \text{(15c)} \\
(x, \lambda) &> 0 && \text{(15d)}
\end{aligned}
$$

- The only difference from the KKT-conditions is $x_i\lambda_i = \tau$ rather than $x_i\lambda_i = 0$.
- The central path

$$\mathcal{C} = \{(x_\tau, \mu_\tau, \lambda_\tau) : \tau > 0\} \tag{16}$$

Central Path

- Points on the central path

$$F(x_\tau, \mu_\tau, \lambda_\tau) = \begin{bmatrix} g - A'\mu_\tau - \lambda_\tau \\ Ax_\tau - b \\ X_\tau \Lambda_\tau e \end{bmatrix} = \begin{bmatrix} r_L \\ r_A \\ X_\tau \Lambda_\tau e \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \tau e \end{bmatrix} \tag{17}$$

$$(x_\tau, \lambda_\tau) > 0 \tag{18}$$

- Newton step

$$\begin{bmatrix} 0 & -A' & -I \\ A & 0 & 0 \\ \Lambda & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \mu \\ \Delta \lambda \end{bmatrix} = - \begin{bmatrix} r_L \\ r_A \\ X\Lambda e \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \tau e \end{bmatrix} = \begin{bmatrix} -r_L \\ -r_A \\ -X\Lambda e + \tau e \end{bmatrix} \tag{19}$$

$$\begin{bmatrix} x_\tau \\ \mu_\tau \\ \lambda_\tau \end{bmatrix} \leftarrow \begin{bmatrix} x_\tau \\ \mu_\tau \\ \lambda_\tau \end{bmatrix} + \alpha \begin{bmatrix} \Delta x \\ \Delta \mu \\ \Delta \lambda \end{bmatrix} \quad (x_\tau, \lambda_\tau) > 0 \tag{20}$$

Central Path

- Duality measure

$$s = \frac{1}{n}\sum_{i=1}^{n} x_i\lambda_i = \frac{x'\lambda}{n} \tag{21}$$

- Centering parameter, $\sigma \in [0, 1]$

$$\tau = \sigma s \tag{22}$$

- Step

$$\begin{bmatrix} 0 & -A' & -I \\ A & 0 & 0 \\ \Lambda & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \mu \\ \Delta \lambda \end{bmatrix} = \begin{bmatrix} -r_L \\ -r_A \\ -X\Lambda e + \sigma s e \end{bmatrix} \tag{23}$$

$$\begin{bmatrix} x_\tau \\ \mu_\tau \\ \lambda_\tau \end{bmatrix} \leftarrow \begin{bmatrix} x_\tau \\ \mu_\tau \\ \lambda_\tau \end{bmatrix} + \alpha \begin{bmatrix} \Delta x \\ \Delta \mu \\ \Delta \lambda \end{bmatrix} \quad (x_\tau, \lambda_\tau) > 0 \tag{24}$$

- Centering step: $\sigma = 1$. Affine step: $\sigma = 0$.

Primal-Dual Framework

Require: (x^0, μ^0, λ^0) such that $(x^0, \lambda^0) > 0$

while not STOP **do**

Compute the duality measure

$$s^k = \frac{(x^k)'\lambda^k}{n} \tag{25}$$

Set the centering parameter: $\sigma^k \in [0, 1]$

Compute residuals

$$r_L = g - A'\mu^k - \lambda^k \tag{26}$$

$$r_A = Ax^k - b \tag{27}$$

Solve

$$\begin{bmatrix} 0 & -A' & -I \\ A & 0 & 0 \\ \Lambda^k & 0 & X^k \end{bmatrix} \begin{bmatrix} \Delta x^k \\ \Delta \mu^k \\ \Delta \lambda^k \end{bmatrix} = \begin{bmatrix} -r_L \\ -r_A \\ -X^k \Lambda^k e + \sigma^k s^k e \end{bmatrix} \tag{28}$$

Choose the step-length α^k such that

$$\begin{bmatrix} x^k \\ \lambda^k \end{bmatrix} + \alpha^k \begin{bmatrix} \Delta x^k \\ \Delta \lambda^k \end{bmatrix} \geq 0 \tag{29}$$

Update

$$\begin{bmatrix} x^{k+1} \\ \mu^{k+1} \\ \lambda^{k+1} \end{bmatrix} = \begin{bmatrix} x^k \\ \mu^k \\ \lambda^k \end{bmatrix} + \alpha^k \begin{bmatrix} \Delta x^k \\ \Delta \mu^k \\ \Delta \lambda^k \end{bmatrix} \tag{30}$$

end while

Mehrotra's Modifications

1. Addition of a **corrector step** to the search direction $\Rightarrow$ The algorithm more closely follows a trajectory to the primal-dual solution set.
2. Adaptive choice of the centering parameter σ.

Motivation:
Shift the central path $\mathcal{C}$ such that it starts at our current iterate (x, μ, λ) and ends at the set of solution points Ω.

1. Predictor step.
2. Adaptive choice of the centering parameter σ.
3. Centering step.
4. Corrector step.

Mehrotra's Modifications

$Ax = b + c$ can be solved as $x = y + z$ in which $Ay = b$ and $Az = c$.

$$\begin{bmatrix} 0 & -A' & -I \\ A & 0 & 0 \\ \Lambda & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \mu \\ \Delta \lambda \end{bmatrix} = \begin{bmatrix} -r_L \\ -r_A \\ -X\Lambda e + \sigma s e \end{bmatrix} \tag{31}$$

Affine direction ($\sigma = 0$)

$$\begin{bmatrix} 0 & -A' & -I \\ A & 0 & 0 \\ \Lambda & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x^{aff} \\ \Delta \mu^{aff} \\ \Delta \lambda^{aff} \end{bmatrix} = \begin{bmatrix} -r_L \\ -r_A \\ -X\Lambda e \end{bmatrix} \tag{32}$$

Center direction

$$\begin{bmatrix} 0 & -A' & -I \\ A & 0 & 0 \\ \Lambda & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x^{cen} \\ \Delta \mu^{cen} \\ \Delta \lambda^{cen} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \sigma s e \end{bmatrix} \tag{33}$$

Combine

$$\begin{bmatrix} \Delta x \\ \Delta \mu \\ \Delta \lambda \end{bmatrix} = \begin{bmatrix} \Delta x^{aff} \\ \Delta \mu^{aff} \\ \Delta \lambda^{aff} \end{bmatrix} + \begin{bmatrix} \Delta x^{cen} \\ \Delta \mu^{cen} \\ \Delta \lambda^{cen} \end{bmatrix} \tag{34}$$

Mehrotra's Modifications

- If the affine step is good, make σ small, i.e. close to zero.
- If the affine step is bad, set $\sigma = 1$

Define α^{aff} and β^{aff} as the largest values satisfying

$$x + \alpha^{aff} \Delta x^{aff} \geq 0 \tag{35}$$

$$\lambda + \beta^{aff} \Delta \lambda^{aff} \geq 0 \tag{36}$$

Duality gap for affine step

$$s^{aff} = \frac{(x + \alpha^{aff} \Delta x^{aff})'(\lambda + \beta^{aff} \lambda^{aff})}{n} \tag{37}$$

Centering parameter

$$\sigma = \left(\frac{s^{aff}}{s} \right)^3 \tag{38}$$

Mehrotra's Modifications

Corrector step

$$\begin{bmatrix} 0 & -A' & -I \\ A & 0 & 0 \\ \Lambda & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x^{cor} \\ \Delta \mu^{cor} \\ \Delta \lambda^{cor} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -\Delta X^{aff} \Delta S^{aff} e \end{bmatrix} \tag{39}$$

Center step

$$\begin{bmatrix} 0 & -A' & -I \\ A & 0 & 0 \\ \Lambda & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x^{cen} \\ \Delta \mu^{cen} \\ \Delta \lambda^{cen} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \sigma s e \end{bmatrix} \tag{40}$$

Mehrotra's direction

$$\begin{bmatrix} \Delta x \\ \Delta \mu \\ \Delta \lambda \end{bmatrix} = \begin{bmatrix} \Delta x^{aff} \\ \Delta \mu^{aff} \\ \Delta \lambda^{aff} \end{bmatrix} + \begin{bmatrix} \Delta x^{cor} \\ \Delta \mu^{cor} \\ \Delta \lambda^{cor} \end{bmatrix} + \begin{bmatrix} \Delta x^{cen} \\ \Delta \mu^{cen} \\ \Delta \lambda^{cen} \end{bmatrix} \tag{41}$$

Mehrotra's Modifications

Mehrotra's direction

$$\begin{bmatrix} 0 & -A' & -I \\ A & 0 & 0 \\ \Lambda & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \mu \\ \Delta \lambda \end{bmatrix} = \begin{bmatrix} -r_L \\ -r_A \\ -X\Lambda e - \Delta X^{aff} \Delta \Lambda^{aff} e + \sigma s e \end{bmatrix} \tag{42}$$

Step length. Largest α and β such that

$$x + \alpha \Delta x \geq 0 \tag{43}$$

$$\lambda + \beta \Delta \lambda \geq 0 \tag{44}$$

Update

$$\begin{bmatrix} x \\ \mu \\ \lambda \end{bmatrix} = \begin{bmatrix} x + \eta\alpha\Delta x \\ \mu + \eta\beta\Delta\mu \\ \lambda + \eta\beta\Delta\lambda \end{bmatrix} \tag{45}$$

Stopping Criteria

The optimality conditions can be expressed as

$$r_L = \nabla_x L = g - A'\mu - \lambda = 0 \tag{46a}$$

$$r_A = Ax - b = 0 \tag{46b}$$

$$X\Lambda e = 0 \tag{46c}$$

$$x \geq 0, \quad \lambda \geq 0 \tag{46d}$$

Dual gap

$$s = \frac{x'\lambda}{n} \tag{47}$$

Hence, simple stopping criteria

$$\|r_L\| \leq \varepsilon \quad \|r_A\| \leq \varepsilon \quad |s| \leq \varepsilon \tag{48}$$

and iteration number less than maximum number of iterations

Predictor-Corrector Interior-Point Algorithm

Require: (x, μ, λ) with $(x, \lambda) > 0$.
while not STOP **do**
Solve (affine direction)

$$\begin{bmatrix} 0 & -A' & -I \\ A & 0 & 0 \\ \Lambda & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x^{aff} \\ \Delta \mu^{aff} \\ \Delta \lambda^{aff} \end{bmatrix} = \begin{bmatrix} -(g - A'\mu - \lambda) \\ -(Ax - b) \\ -X\Lambda e \end{bmatrix} \tag{49}$$

Compute the largest α^{aff} and β^{aff} such that

$$x + \alpha^{aff} \Delta x^{aff} \geq 0 \qquad \lambda + \beta^{aff} \Delta \lambda^{aff} \geq 0 \tag{50}$$

Compute the affine duality gap: $s^{aff} = (x + \alpha^{aff} \Delta x^{aff})'(\lambda + \beta^{aff} \Delta \lambda^{aff})/n$
Compute the centering parameter: $\sigma = (s^{aff}/s)^3$ with $s = x'\lambda/n$
Solve

$$\begin{bmatrix} 0 & -A' & -I \\ A & 0 & 0 \\ \Lambda & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \mu \\ \Delta \lambda \end{bmatrix} = \begin{bmatrix} -(g - A'\mu - \lambda) \\ -(Ax - b) \\ -X\Lambda e - \Delta X^{aff} \Delta \Lambda^{aff} e + \sigma s e \end{bmatrix} \tag{51}$$

Compute the largest α and β such that

$$x + \alpha \Delta x \geq 0 \qquad \lambda + \beta \Delta \lambda \geq 0 \tag{52}$$

$\eta = 0.995$. Update

$$\begin{bmatrix} x \\ \mu \\ \lambda \end{bmatrix} \leftarrow \begin{bmatrix} x + \eta \alpha \Delta x \\ \mu + \eta \beta \Delta \mu \\ \lambda + \eta \beta \Delta \lambda \end{bmatrix} \tag{53}$$

end while

Matrix Factorization

The linear system is

$$\begin{bmatrix} 0 & -A' & -I \\ A & 0 & 0 \\ \Lambda & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \mu \\ \Delta \lambda \end{bmatrix} = \begin{bmatrix} -r_L \\ -r_A \\ -r_C \end{bmatrix} \tag{54}$$

The third equation gives

$$\Lambda \Delta x + X \Delta \lambda = -r_C \quad \Leftrightarrow \quad \Delta \lambda = -X^{-1} r_C - X^{-1} \Lambda \Delta x \tag{55}$$

Substitution in the first equation gives

$$-r_L = -A' \Delta \mu - \Delta \lambda = -A' \Delta \mu + X^{-1} r_c + X^{-1} \Lambda \Delta x \tag{56}$$

and

$$X^{-1} \Lambda \Delta x - A' \Delta \mu = -r_L - X^{-1} r_C \tag{57}$$

Augmented form

$$\begin{bmatrix} X^{-1} \Lambda & -A' \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \mu \end{bmatrix} = \begin{bmatrix} -r_L - X^{-1} r_C \\ -r_A \end{bmatrix} \tag{58}$$

Matrix Factorization

The equation

$$X^{-1}\Lambda\Delta x - A'\Delta\mu = -r_L - X^{-1}r_C \tag{59}$$

gives

$$\Delta x = \Lambda^{-1}XA'\Delta\mu - \Lambda^{-1}Xr_L - \Lambda^{-1}r_C \tag{60}$$

Substitution in

$$A\Delta x = -r_C \tag{61}$$

gives the **normal equations**

$$\left(A\Lambda^{-1}XA'\right)\Delta\mu = -r_C - A\left(-\Lambda^{-1}Xr_L - \Lambda^{-1}r_C\right) \tag{62}$$

Matlab code

```
function [x,info,mu,lambda,iter] = LPippd(g,A,b,x)
% LPIPPD   Primal-Dual Interior-Point LP Solver
%
%          min  g'*x
%           x
%          s.t. A x  = b      (Lagrange multiplier: mu)
%                 x >= 0      (Lagrange multiplier: lamba)
%
% Syntax: [x,info,mu,lambda,iter] = LPippd(g,A,b,x)
%
%         info = true   : Converged
%              = false  : Not Converged
```