Ch 14: LP: Interior-point methods

14.1 Primal dual methods.

Consider $\min c^T x$ subj. to $Ax=b$, $x \geq 0$.
where A is $m \times n$ with full row rank.

The dual problem is:

$$\max b^T \lambda$$

subject to $A^T\lambda + s = c$, $s \geq 0$.

Summarize KKT by (13.4):

$$(*) \begin{cases} A^T\lambda + s = c \\ Ax = b \\ \left.\begin{array}{l} x \geq 0 \\ s \geq 0 \end{array}\right\} \text{summarized as: } (x,s) \geq 0. \\ x_i s_i = 0 \end{cases}$$

from $\mathcal{L}(x,\lambda,s) = c^T x - \lambda^T(Ax-b) - s^T x$

Problem: $(x,s) \geq 0$, rest ok!

Restate $(*)$ using $F: \mathbb{R}^{2n+m} \longrightarrow \mathbb{R}^{2n+m}$ with:

$$F(x,\lambda,s) = \begin{bmatrix} A^T\lambda + s - c \\ Ax - b \\ XSe \end{bmatrix} = 0, \quad (x,s) \geq 0, \quad \begin{cases} X = \operatorname{diag}(x_1, \dots, x_n) \\ S = \operatorname{diag}(s_1, \dots, s_n) \\ e = (1, 1, \dots, 1)^T \end{cases}$$

Basic idea:

Solve $F(x,\lambda,s)=0$ with $x,s>0$ strictly.

This generates (x^k, s^k, λ^k).

Problem: $F(x,\lambda,s)=0$ but not $x,s \geq 0$.
$\Rightarrow$ They are not useful!

Evaluate iterate using a duality measure:

$$0 < \mu = 1/n \sum_{i=1}^{n} x_i s_i = \frac{x^T s}{n}$$ "average positivity".

From ch 11, algorithm 11.1, we have that

$$J(x_k)\, p_k = -r(x_k)$$

where:

$$J(x_k) = \begin{bmatrix} \nabla r_1(x_k)^T \\ \nabla r_2(x_k)^T \\ \vdots \\ \nabla r_n(x_k)^T \end{bmatrix}$$

is used to solve $r(x_k)=0$, to drive the residuals to zero.

We have a similar problem here, where we want $F(x,\lambda,s)=0$:

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

<u>Define</u>: $r_b = Ax - b = r_b(x)$

$r_c = A^T\lambda + s - c = r_c(\lambda, s)$

<u>Then</u>, define the order for $\nabla F_i(x,\lambda,s)$ by:

$$\nabla F_i(x,\lambda,s) = \begin{bmatrix} \nabla_x F_i \\ \nabla_\lambda F_i \\ \nabla_s F_i \end{bmatrix}, \quad i=1,2,3$$

This ordering leads to:

$$\nabla F_1 = \begin{bmatrix} \nabla_x F_1 \\ \nabla_\lambda F_1 \\ \nabla_s F_1 \end{bmatrix} = \begin{bmatrix} \nabla_x r_c \\ \nabla_\lambda r_c \\ \nabla_s r_c \end{bmatrix} = \begin{bmatrix} 0 \\ A^T \\ I \end{bmatrix}$$

$$\nabla F_2 = \begin{bmatrix} \nabla_x r_b \\ \nabla_\lambda r_b \\ \nabla_s r_c \end{bmatrix} = \begin{bmatrix} A \\ 0 \\ 0 \end{bmatrix}$$

$$\nabla F_3 = \begin{bmatrix} S \\ 0 \\ X \end{bmatrix}$$

We now have the basic equation for taking the error to zero:

$$\begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \lambda \\ \Delta s \end{bmatrix} = \begin{bmatrix} -r_c \\ -r_b \\ -XSe \end{bmatrix}$$

from:

$$\begin{bmatrix} \nabla F_1^T \\ \nabla F_2^T \\ \nabla F_3^T \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \lambda \\ \Delta s \end{bmatrix} = -F$$

Here, we require $(x, s) \geq 0$ that leads to a line search along:

$$(x, \lambda, s) + \alpha(\Delta x, \Delta \lambda, \Delta s)$$

for $\alpha \in (0, 1]$.

Problem:

* Often, $\alpha << 1$ before violating $(x, s) > 0$

Approach:

* do not try to bring (x, s) down to zero.

Modify the equation to

$x_i s_i = \sigma \mu$, μ is the current measure

and $\sigma \in [0,1]$.

This modifies RHS as follows:

$$\begin{bmatrix} -r_c \\ -r_b \\ -XSe \end{bmatrix} \xrightarrow{\text{Becomes}} \begin{bmatrix} -r_c \\ -r_b \\ -XSe + \sigma \mu e \end{bmatrix}$$

Where σ is a centering parameter.

Framework 14.1 (Primal-dual path following)

Given (x^0, λ^0, s^0) with $(x^0, s^0) > 0$

for $k = 0, 1, 2, \ldots$

Choose $\sigma_k \in [0,1]$ and solve

$$\begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S^k & 0 & X^k \end{bmatrix} \begin{bmatrix} \Delta x^k \\ \Delta \lambda^k \\ \Delta s^k \end{bmatrix} = \begin{bmatrix} -r_c^k \\ -r_b^k \\ -X^k S^k e + \sigma_k \mu_k e \end{bmatrix}$$

where $\mu_k = (x^k)^T s^k / n$;

$(x^{k+1}, \lambda^{k+1}, s^{k+1}) = (x^k, \lambda^k, s^k) + \alpha_k (\Delta x^k, \Delta \lambda^k, \Delta s^k)$

choosing α_k so that $(x^{k+1}, s^{k+1}) > 0$

end

See section 14.2 algorithm for solving this without starting from a feasible initial point.

The Central Path

Define: Primal-dual feasible set $\mathcal{F}$:

$$\mathcal{F} = \{(x, \lambda, s) \mid Ax = b,\ A^T\lambda + s = c,\ (x, s) \geq 0\}$$

Define: Primal-dual strictly feasible set $\mathcal{F}^o$:

$$\mathcal{F}^o = \{(x, \lambda, s) \mid Ax = b,\ A^T\lambda + s = c,\ (x, s) > 0\}$$

The central path $\mathcal{C}$ is an arc of strictly feasible points parametrized by $\tau > 0$ generating $(x_\tau, \lambda_\tau, s_\tau) \in \mathcal{C}$ using:

$$\begin{aligned} A^T\lambda + s &= c \\ Ax &= b \\ x_i s_i &= \tau, \quad i = 1, 2, \ldots, n \\ (x, s) &> 0 \end{aligned}$$

Then:

$$\mathcal{C} = \{(x_\tau, \lambda_\tau, s_\tau) \mid \tau > 0\}$$

The so-called log-barrier method then becomes:

$$\min\ c^T x - \tau \sum_{i=1}^{n} \ln x_i \quad \sim (*)$$

$$\text{subject to} \quad Ax = b$$

with KKT conditions:

$$i=1,\dots,n:\ \underbrace{c_i - \frac{\tau}{x_i}}_{\nabla_{x_i}(*)} - \underbrace{A_{\bullet i}^T}_{x_i\text{-th column of } A}\ \lambda = 0 \quad \text{and} \quad Ax=b$$

Lagrange multiplier

Note that (*) is strictly convex, giving sufficient and necessary conditions for optimality.

As $\tau \downarrow 0$, we approximate the primal dual solution, maintaining positivity for x_i, s_i with $x_i s_i \to 0$.

$\Rightarrow$ Can take longer steps in the interior.

<u>Algorithm 14.3</u> gives a practical method for solving the primal dual equations.