COMPUTATIONAL METHODS IN OPTIMIZATION

September 25, 2023

Today: Basics of primal-dual interior-point methods (in the context of LPs)

INTERIOR-POINT METHODS (IPM) FOR LINEAR PROGRAMMING

↳ Why IPMs ?

→ Very effective for structured problems
→ Usually better than simplex methods when the number of variables/constraints is large
→ Unlike simplex methods, they extend "easily" to nonlinear programs (QPs, SDPs, conic programs, ...)

Optimality and duality results for LP.

(1) Optimality and

Problem :
(primal)

$$(P) \quad \underset{x \in \mathbb{R}^n}{\text{minimize}} \quad c^T x \quad \text{s.t.} \quad \begin{array}{l} Ax = b \\ x \geq 0 \end{array}$$

$c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, $A \in \mathbb{R}^{m \times n}$

(for simplicity, we assume that $\text{rank}(A) = m \leq n$)

Dual of (P)

$$(D) \quad \underset{\substack{y \in \mathbb{R}^m \\ s \in \mathbb{R}^n}}{\text{maximize}} \quad b^T y \quad \text{s.t.} \quad \begin{array}{l} A^T y + s = c \\ s \geq 0 \end{array}$$

Th (Strong duality for LP)
Consider (P) and (D). There are 3 possible scenarios

i) Both problems are infeasible, i.e.

$\{x \mid Ax = b, x \geq 0\} = \emptyset$ and $\{(y,s) \mid A^T y + s = c, s \geq 0\} = \emptyset$

i:) One of (P) and (D) is infeasible and the other is unbounded

Ex) $\{x \mid Ax=b, x\geq 0\} = \emptyset$

and $\max_{\substack{y\in\mathbb{R}^m \\ s\in\mathbb{R}^n}} b^T y$ s.t. $A^Ty+s=c$, $s\geq 0$ $= +\infty$

(with $\{(y,s) \mid A^Ty+s=c, s\geq 0\} \neq \emptyset$)

ii) Both (P) and (D) are feasible and have solutions (they are not unbounded), in which case for any triplet (x^*, y^*, s^*) such that x^* solves (P) and (y^*, s^*) solves (D), then

$$c^T x^* = b^T y^*$$

↳ The previous theorem gives guarantees on the existence of solutions to (P) (and (D)).

A characterization of those solutions is given by the KKT (Karush-Kuhn-Tucker) optimality conditions.

Theorem: KKT conditions for LP

A triplet (x^*, y^*, s^*) is a primal-dual solution of (P) (i.e. x^* solves (P) and (y^*, s^*) solves (D)) if and only if it satisfies the following conditions:

x^* feasible for (P)

i) $Ax^* = b$

ii) $A^T y^* + s^* = c$

iii) $x^* \geq 0$

iv) $s^* \geq 0$

v) $x_i^* s_i^* = 0 \quad \forall\, i = 1..n$

(y^*, s^*) feasible for (D)

Complementarity condition
- Only nonlinear condition
- When iii) iv) hold, equivalent to $(x^*)^T (s^*) = 0$

NB: Sometimes (iii), (iv), (v) are combined using the notation
$$0 \leq x^* \perp s^* \geq 0$$

Def: The set of primal-dual feasible points for ((P), (D)) is defined as

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

↳ Given $(x, y, s) \in \mathcal{F}$, the only KKT condition possibly not satisfied by (x, y, s) is (v) (complementarity)

⇒ This last condition quantifies how far the feasible point is from being optimal by looking at $x^T s$

↳ If $x^T s = 0$, then (v) is satisfied and we have found a solution

↳ If $x^T s > 0$, then we want to compute $(\bar{x}, \bar{y}, \bar{s}) \in \mathcal{F}$ such that $(\bar{x})^T (\bar{s}) < x^T s$

⇒ IPMs consider the quantity

$$\mu = \mu(x, y, s) = \frac{x^T s}{n} = \frac{1}{n} \sum_{i=1}^{n} x_i s_i$$

NB

Solution: (x^*, y^*, s^*) is a primal-dual solution if

→ $x^* \in \underset{x \in \mathbb{R}^n}{\operatorname{argmin}}\ c^T x$ s.t. $Ax = b,\ x \geq 0$

($\forall x \in \mathbb{R}^n$ such that: $Ax = b, x \geq 0$, $c^T x \geq c^T x^*$)

$\cdot\ Ax^* = b,\ x^* \geq 0$

→ $(y^*, s^*) \in \underset{\substack{y \in \mathbb{R}^m \\ s \in \mathbb{R}^n}}{\operatorname{argmax}}\ b^T y$ s.t. $A^T y + s = c$, $s \geq 0$

- $\forall (y, s) \in \mathbb{R}^m \times \mathbb{R}^n$ such that $A^T y + s = c,\ s \geq 0$: $b^T y \leq b^T y^*$
- $A^T y^* + s^* = c,\ s^* \geq 0$

↳ IPMs compute a sequence of primal-dual iterates $(x^k, y^k, s^k) \in \mathcal{F}$ such that $\{\mu^k = \frac{(x^k)^T(s^k)}{n}\}$ converges to 0 (hence (x^k, y^k, s^k) will converge to a solution)

- Primal-dual IPMs control μ^k explicitly (use it directly in the algorithm)
- Barrier IPMs control μ^k implicitly (not used in practice but used to show convergence)

(Simplex methods act directly on x and s by fixing some of those coefficients to be 0 ⇒ Force (v) to be satisfied and then check whether (i)-(iv) hold)

② Primal-dual IPMs

Starting point : equalities in the KKT conditions (i), (ii), (v)

$$\begin{cases} (i) & Ax - b = 0 \\ (ii) & A^T y + s - c = 0 \\ (v) & \underbrace{XSe}_{\begin{bmatrix} x_1 s_1 \\ \vdots \\ x_n s_n \end{bmatrix}} = 0 \end{cases}$$

$X = \operatorname{diag}(x) = \begin{bmatrix} x_1 & & 0 \\ & \ddots & \\ 0 & & x_n \end{bmatrix}$ $X \in \mathbb{R}^{n \times n}$

$S = \operatorname{diag}(s) = \begin{bmatrix} s_1 & & 0 \\ & \ddots & \\ 0 & & s_n \end{bmatrix}$

$e = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \in \mathbb{R}^n$

⟶ A system of $m + 2n$ equations with $\underbrace{m}_{|y| = m} + 2n$ unknowns $|x| = |s| = n$

KKT conditions $\Leftrightarrow$ $\begin{cases} F_0(x,y,s) = 0 \\ x \geq 0, \ s \geq 0 \end{cases}$ where $F_0(x,y,s) = \begin{bmatrix} Ax - b \\ A^T y + s - c \\ XSe \end{bmatrix}$

Primal-dual IPM principle

At every iteration k,

- Compute $(x^k, y^k, s^k) \in \mathcal{F}$ such that $x^k > 0$ and $s^k > 0$
- Then apply one iteration of Newton's method to the system $F_0(x,y,s) = 0$ starting from (x^k, y^k, s^k) $\Rightarrow (\Delta x^k, \Delta y^k, \Delta s^k)$

$\forall k \geq 1$ (x^k, y^k, s^k) depends on $(x^{k-1}, y^{k-1}, s^{k-1})$ and $(\Delta x^{k-1}, \Delta y^{k-1}, \Delta s^{k-1})$

$\hookrightarrow$ Computing $(\Delta x^k, \Delta y^k, \Delta s^k)$ requires solving a linear system

$$\begin{array}{c} \\ m \\ n \\ n \end{array}\begin{array}{c} \begin{array}{ccc} n & m & n \end{array} \\ \begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I \\ S^k & 0 & X^k \end{bmatrix} \end{array} \begin{bmatrix} \Delta x^k \\ \Delta y^k \\ \Delta s^k \end{bmatrix} \begin{array}{c} n \\ m \\ n \end{array} = \begin{bmatrix} 0 \\ 0 \\ -X^k S^k e \end{bmatrix} \begin{array}{c} m \\ n \\ n \end{array}$$

$m + 2n$

$X^k = \mathrm{diag}(x^k)$

$S^k = \mathrm{diag}(s^k)$

NB: Since $x^k > 0$ and $s^k > 0$, the last n components of the right-hand side are nonzero.

$\hookrightarrow$ Computing $(x^{k+1}, y^{k+1}, s^{k+1})$ given (x^k, y^k, s^k) and $(\Delta x^k, \Delta y^k, \Delta s^k)$

$$\begin{bmatrix} x^{k+1} \\ y^{k+1} \\ s^{k+1} \end{bmatrix} = \begin{bmatrix} x^k \\ y^k \\ s^k \end{bmatrix} + \alpha^k \begin{bmatrix} \Delta x^k \\ \Delta y^k \\ \Delta s^k \end{bmatrix}$$

$\begin{bmatrix} d_x^k \\ d_y^k \\ d_s^k \end{bmatrix} \in \mathbb{R}^{m+2n}$

where $\alpha^k > 0$ is chosen so that $x^{k+1} > 0$ and $s^{k+1} > 0$

Illustration