COMPUTATIONAL METHODS IN OPTIMIZATION

October 9, 2023

Today: End of "crash course" on interior-point methods

Next session: Monday, October 23

1 Path-following IPMs

Setup: linear programming

$$(P) \quad \underset{x \in \mathbb{R}^n}{\text{minimize}} \; c^T x \quad \text{s.t. } Ax = b, \; x \geq 0 \qquad (D) \quad \underset{\substack{y \in \mathbb{R}^m \\ s \in \mathbb{R}^n}}{\text{maximize}} \; b^T y \quad \text{s.t. } A^T y + s = c, \; s \geq 0$$

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

↳ Primal-dual IPMs generate a sequence of primal-dual iterates $\{(x^k, y^k, s^k)\}_{k \in \mathbb{N}}$ such that $\underbrace{x^k > 0, \; s^k > 0}_{\text{"Interior points"}}$ $\forall k$ and $\frac{(x^k)^T (s^k)}{n} \xrightarrow[k \to \infty]{} 0$

↳ Generic algorithm:

- Choose (x^0, y^0, s^0)
- Given (x^k, y^k, s^k), compute

$$\begin{bmatrix} x^{k+1} \\ y^{k+1} \\ s^{k+1} \end{bmatrix} = \begin{bmatrix} x^k \\ y^k \\ s^k \end{bmatrix} + \underbrace{\alpha^k}_{\text{stepsize} > 0} \overbrace{\begin{bmatrix} \Delta x^k \\ \Delta y^k \\ \Delta s^k \end{bmatrix}}^{\text{"Search direction"}}$$

↳ Assumption (for today): (x^0, y^0, s^0) is primal-dual strictly feasible

i.e. $Ax^0 = b$ $\qquad x^0 > 0$

$A^T y^0 + s^0 = c$ $\qquad s^0 > 0$

- Infeasible IPMs do not require that assumption (and typically work better than methods that do!)

- Homogeneous self-dual IPMs satisfy the assumption automatically by reformulating the problem (The solver MOSEK uses this)

Computing the search direction

Approach: Use the KKT conditions for (P) and (D)

(x^*, y^*, s^*) is a primal-dual solution $\iff$
$$\begin{cases} Ax^* - b = 0 \\ A^T y^* + s^* - c = 0 \\ x^* \geq 0 \\ s^* \geq 0 \\ x_i^* s_i^* = 0 \quad \forall i = 1..n \end{cases}$$

- In a simplex method, the exact solution is obtained by looking at (x, y, s) that satisfy $x_i s_i = 0 \ \forall i = 1..n$ ($\iff$ fix some x_i and s_i to be 0)

- In an IPM, one considers only approximate solutions (x^k, y^k, s^k) such that $x^k > 0$, $s^k > 0$ ($\Rightarrow x_i^k s_i^k > 0 \ \forall i = 1..n$) that get closer and closer to an exact solution

↳ An exact solution (x^*, y^*, s^*) satisfies (in particular)

Nonlinear system of equations →
$$\boxed{\begin{aligned} Ax^* - b &= 0 \\ A^T y^* + s^* - c &= 0 \\ X^* S^* e &= 0 \end{aligned}}$$

$$X^* = \begin{bmatrix} x_1^* & & 0 \\ & \ddots & \\ 0 & & x_n^* \end{bmatrix} \qquad S^* = \begin{bmatrix} s_1^* & & 0 \\ & \ddots & \\ 0 & & s_n^* \end{bmatrix}$$

$2n+m$ equations
$2n+m$ variables (x^*, y^*, s^*)

Given (x^k, y^k, s^k), one can apply Newton's method for nonlinear equations starting from (x^k, y^k, s^k) to obtain a direction towards a better point in the sense of the nonlinear system

$$\begin{aligned} Ax - b &= 0 \\ A^T y + s - c &= 0 \\ XSe &= 0 \end{aligned}$$

$\Rightarrow$ In practice, we apply 1 iteration of Newton's method, which corresponds to solving the linear system

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I_n \\ S^k & 0 & X^k \end{bmatrix} \underbrace{\begin{bmatrix} \Delta x^k \\ \Delta y^k \\ \Delta s^k \end{bmatrix}}_{\text{unknowns}} = - \begin{bmatrix} Ax^k - b \\ A^T y^k + s^k - c \\ X^k S^k e \end{bmatrix}$$

Nonlinear system at (x^k, y^k, s^k)

$\rightarrow$ If (x^k, y^k, s^k) is primal-dual strictly feasible, the right-hand side is $\begin{bmatrix} 0 \\ 0 \\ -X^k S^k e \end{bmatrix}$ with the last n coefficients being nonzero.

Two issues with this direction

- Moving in the direction might lead to negative coefficients for x^{k+1} and s^{k+1}
- Newton's method does not always converge

$\Rightarrow$ Because of these issues, one cannot simply do

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

$\Rightarrow$ In practice, we move in the direction for a certain stepsize $\alpha^k > 0$

Ex) $\alpha^k = \max_{\alpha > 0} \{ x^k + \alpha \Delta x^k > 0$ and $s^k + \alpha \Delta s^k > 0 \}$

↳ The approach described above can produce iterates that get close to the boundary of the constraints ((x,y,s) primal-dual feasible and $x_i s_i = 0$ for some i)

For this reason, researchers have developed path-following methods in which:

a) the linear system is perturbed

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I_n \\ S^k & 0 & X^k \end{bmatrix} \begin{bmatrix} \Delta x^k \\ \Delta y^k \\ \Delta s^k \end{bmatrix} = - \begin{bmatrix} Ax^k - b \\ A^T y^k + s^k - c \\ X^k S^k e - \sigma \mu^k e \end{bmatrix}$$

where $\sigma \in (0,1)$ and $\mu^k = \frac{(x^k)^T (s^k)}{n}$

b) α^k is computed under more restrictive conditions

Short-step path-following: Compute $\alpha^k > 0$ so that $x^{k+1} > 0$, $s^{k+1} > 0$

and $\underbrace{\| X^{k+1} S^{k+1} e - \mu^{k+1} e \| \leq \gamma \mu^{k+1}}$ for $\gamma \in (0,1)$

"The coefficients $x_i^k s_i^k$ are decreasing in an uniform fashion"

Long-step path-following: Compute $\alpha^k > 0$ so that $x^{k+1} > 0$, $s^{k+1} > 0$

and $\underbrace{x_i^{k+1} s_i^{k+1} \geq \gamma \mu^{k+1}}$ for $\gamma \in (0,1)$ $\forall i = 1..m$

Every component of $[x_i^{k+1} s_i^{k+1}]$ represents at least a fraction of μ^{k+1}

→ Avoids very small values of $x_i^{k+1} s_i^{k+1}$ compared to the others

2 Barrier functions and IPMs

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

Idea: • Want to get to a solution from "the interior"

$\{x^k\}$ $x^k > 0 \Rightarrow x^*$ solution

• Approach: Replace the constraints $x \geq 0$ by a barrier term in the objective to guarantee positivity

At every iteration k, we consider

$$(P_{\mu^k}) \quad \underset{x \in \mathbb{R}^n}{\text{minimize}} \ c^T x - \mu^k \sum_{i=1}^n \ln(x_i) \quad \text{s.t.} \ Ax = b, \quad x \geq 0$$

for some $\mu^k > 0$

↳ $x \longmapsto c^T x - \mu^k \sum_{i=1}^n \ln(x_i)$ is nonlinear

equal to $+\infty$ at every x with a zero coordinate.
and it excludes de facto negative coordinates

Barrier IPMs compute a sequence of solutions of subproblems (P_{μ^k}) for decreasing values of μ^k

As $\mu^k \to 0$, the solution gets closer to that of the original problem (but it remains in the interior!)

↳ there exist primal, dual, and primal-dual barrier methods depending on which problem(s) are solved