Fall 2016, MATH-566

Linear Programming Algorithms - Interior point methods

Source: Chapter 11 of Convex Optimization, Stephen Boyd and Lieven Vandenberghe

Let

$$(P)\begin{cases}\text{minimize} & f(\mathbf{x})\\ \text{subject to} & g_i(\mathbf{x}) \le 0 \text{ for } i = 1,\ldots,m\end{cases}$$

where f, g are convex, twice continuously differentiable and optimal solution $\mathbf{x}^\star$ exists.
Moreover, let (P) be *superconsistent*, that is $\exists\mathbf{x}, \forall i, g_i(\mathbf{x}) < 0$. In other words, the set of feasible solutions have full dimension.

(the setup covers linear, quadratic, geometric, semidefinite, ... programming).

Idea: Change the (P) to a problem without constraints but difficult objective function.

Let

$$(P') = \text{minimize } f(\mathbf{x}) + \sum_{i=1}^{m} I(g_i(\mathbf{x})),$$

where I is an indicator function

$$I(u) = \begin{cases}0 & \text{if } u \le 0\\ +\infty & \text{if } u > 0.\end{cases}$$

1: What is the optimal solution to (P')?

2: Can you solve (P') by methods from calculus?

Use approximation of $I(u) \approx -c\log(-u)$, where $c > 0$.

3: Sketch $I(u)$ and its approximations. Is the approximation better when c is large or small?

For $t > 0$, we consider a smooth unconstrained approximation of (P')

$$\text{minimize } f(\mathbf{x}) - \frac{1}{t}\sum_{i=1}^{m} \log(-g_i(\mathbf{x})).$$

Define *logarithmic barrier function*

$$\Phi(\mathbf{x}) = -\sum_{i=1}^{m} \log(-g_i(\mathbf{x})),$$

for all $\mathbf{x}$ where $g_i(\mathbf{x}) < 0$ (interior of feasible solutions)
Analytic center of the set $S = \{\mathbf{x} : g_i(\mathbf{x}) \leq 0\} \subseteq \mathbb{R}^n$ is $\mathbf{x}^\star$ minimizing $\Phi(\mathbf{x})$ over all $\mathbf{x} \in S$.

4: Find the analytic center of a square in $\mathbf{R}^2$ defined by equations

$$x_1 \geq 0, x_2 \geq 0, x_1 \leq 1, x_2 \leq 1$$

5: Find the analytic center of a square in $\mathbf{R}^2$ defined by equations

$$x_1 \geq 0, x_2 \geq 0, (1 - x_1)^3 \geq 0, (1 - x_2)^3 \geq 0.$$

Notice it is possible to define center even if functions are not convex everywhere and the center depends on the functions.

For $t > 0$ define $\mathbf{x}^\star(t)$ as the optimal solution of

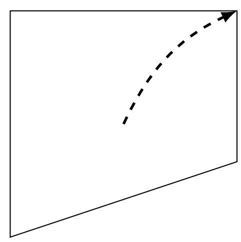
$$(P_t) = \text{ minimize } tf(\mathbf{x}) + \Phi(\mathbf{x}).$$

(assume that the optimal solution is unique)

Central path is $\{\mathbf{x}^\star(t) : t \geq 0\}$.

Interior point method idea: Start in the analytical center and follow the central path.

In iterations increase t and recompute the new optimum using Newton's method. Recall that Newton's method works well if the initial point of Newton's method is close to the optimal solution. With small increases of t, the starting point is close to the optimum.



There exists a notion of dual program (D) for (P), (based on Karush-Kuhn-Tucker theorem). It gives solutions to the dual $\mathbf{y}^\star(t)$ such that

$$f(\mathbf{x}^\star(t)) - h(\mathbf{y}^\star(t)) \leq \frac{m}{t},$$

where h is the objective function of the dual program. Hence the central path converges to $\mathbf{x}^\star$ for (P).