Linear Programming Algorithms - Interior point methods

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

Let

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

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 has 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 the *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).

The **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 $\mathbf{x} \in S$.

Since is $\Phi(\mathbf{x})$ is strict convexity on a convex domain, it has has a unique minimizer.

4: Using calculus, find the analytic center of a square in $\mathbb{R}^2$ 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 $\mathbb{R}^2$ defined by equations

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

Notice that the analytic center depends on the functions $g_i(x)$, not just the set of feasible solutions. In particular, it does not have to coincide with geometric center. Also, we may define analytic center in the same way even if the functions $g_i(x)$ are not convex. Compare the the answer with the solution to the previous exercise.

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

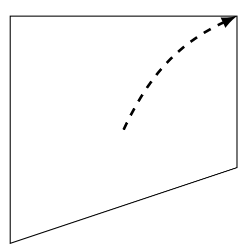
$$(P_t) = \text{ minimize } tf(\mathbf{x}) + \Phi(\mathbf{x}).$$

(assume that the optimal solution is unique)

The **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 and m number of constraints. Hence the central path converges to $\mathbf{x}^\star$ for (P).

There exists a "nice" implementation for linear programming, where the central path looks somewhat nice.

6: Compute central path for the following problem

$$
(P)\begin{cases}
\text{minimize} & -x_1 \\
\text{subject to} & x_1 \le 1 \\
 & x_2 \le 1 \\
 & x_1 \ge 0 \\
 & x_2 \ge 0
\end{cases}
$$

and find the optimal solution using the central path. Plot (sketch) the set of feasible solutions and the computed central path. Lot of calculus...

Hint: The central path is formed by points that are optimal solutions to

$$\min\ h_t(x_1, x_2) = -tx_1 + \Phi(x_1, x_2),$$

where $t \ge 0$ and $\Phi(x_1, x_2)$ is the barrier function. Take partial derivatives of $h_t(x_1, x_2)$ to obtain optimal solution $(x_1, x_2)_t$. For $t \in [0, \infty]$ these points of optimal solutions will form a curve. Plot or describe the curve. This curve is the central path.