ORIE 3310/5310
Lecture 27 (Draft)

~~26~~ 30 April 2013

Today: Interior Point Methods

- This week, we will study an algorithm called interior-point method, for solving constrained optimization problems.
  - Although int. pt. methods can be used to solve nonlinear programming problems, we are only going to develop this method for solving linear programming problems.
  - That is, we will consider solving LP in standard form:

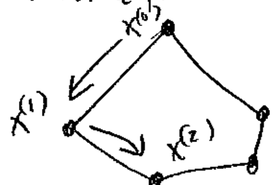
$$\begin{aligned} \text{Max} \quad & c^T x \\ \text{s.t.} \quad & Ax \le b \\ & x \geqslant 0. \end{aligned}$$

→ The version of int. pt. methods for (LP) captures the big ideas that also applies for NLP (without the technical difficulties).

- Big-picture idea of interior pt methods (in contrast to the simplex method).

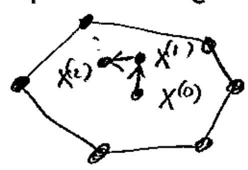
Simplex Method

- Initial point is an extreme point / a basic feasible solution.
- At each iteration,
  - our current iterate $x^{(i)}$ is an extreme point.
  - We perform simplex pivot, an obtain a new extreme point as new iterate



Int. Pt. Methods

- Initial point is an interior pt (inside the feasible region; not on boundary).
- At each iteration,
  - our current iterate $x^{(i)}$ is an int. pt.
  - We use Newton's method to obtain the next iterate, $x^{(i+1)}$, which is also a point in the interior of the feasible region



- Before writing the formal algorithm, we will develop the algorithm together via an example:

<u>Ex 1</u>.

$$\begin{aligned} \text{Max} \quad & 3x_1 \\ \text{s.t.} \quad & x_1 \leq 2 \\ & x_1 \geq 0 \end{aligned}$$

Rewriting the first constraint:

$$\left.\begin{aligned} \text{Max} \quad & 3x_1 \\ \text{s.t.} \quad & 2-x_1 \geq 0 \\ & x_1 \geq 0 \end{aligned}\right\} (*).$$

Then, consider the following <u>barrier function</u>, where μ is a fixed positive number

$$f_\mu(x_1) = 3x_1 + \mu\left[\log(2-x_1) + \log(x_1)\right]$$

<u>Remarks</u> ⓪ The function $\log(x)$ is defined when $x>0$. log = natural log ($\ln$)

① This function is defined when the quantities that we take log of are positive

i.e. for x_1 such that

$2-x_1 > 0$ and

$x_1 > 0$.

$\therefore f_\mu(x_1)$ is defined in the interior of the feasible region of $(*)$

② Fix $\mu>0$ (say $\mu=1$)

As x_1 approaches 0, $f_\mu(x_1) \downarrow -\infty$.

As x_1 approaches 2, $f_\mu(x_1) \downarrow -\infty$.

$\therefore f_\mu(x_1) \uparrow +\infty$ as x_1 approaches the boundaries of the feasible region.

because $-\log x$ is convex

③ $-f_\mu(x_1)$ is a convex function, so to minimize $(-f_\mu(x))$, we can solve: $\frac{d}{dx_1} f_\mu(x_1) = 0$, then use Newton's method:

$$f_\mu'(x_1) = \frac{d}{dx_1} f_\mu'(x_1)$$
$$= 3 + \frac{\mu}{2-x_1} - \frac{\mu}{x_1}$$

Solving for x_1:
$$-3 + \frac{\mu}{2-x_1} - \frac{\mu}{x_1} = 0$$
$$\frac{x_1\mu - 2\mu + x_1\mu}{(2-x_1)x_1} = 3$$
$$2x_1\mu - 2\mu = 6x_1 - 3x_1^2$$
$$\boxed{3x_1^2 + (2\mu-6)x_1 - 2\mu = 0}$$ Solve for x_1

Option 1: Newton's method (several iterations).
Option 2: Solve directly
Since only a quadratic, then let's solve this directly
(but if more complicated, might want to resort to Newton's Method).
∴ Use the quadratic formula:
$$x_1 = \frac{-(2\mu-6) \pm \sqrt{(2\mu-6)^2 - 4(-2\mu)(3)}}{2\cdot 3}$$
$$= \frac{6-2\mu}{6} \pm \frac{1}{6}\sqrt{4\mu^2 - \cancel{24\mu} + 36 + \cancel{24\mu}}$$
$$= 1 - \frac{\mu}{3} \pm \frac{1}{6}\, 2\sqrt{\mu^2+9}$$
$$= 1 - \frac{\mu}{3} \pm \frac{1}{3}\sqrt{\mu^2+9}$$

$1 - \frac{4}{3} + \frac{1}{3}\sqrt{25}$

$-\frac{1}{3} + \frac{5}{3}$

$$x_1 = 1 - \frac{\mu}{3} - \underbrace{\frac{1}{3}\sqrt{\mu^2+9}}_{>1} \quad \text{or} \quad x_1 = 1 - \frac{\mu}{3} + \frac{\sqrt{\mu^2+9}}{3}$$

∴ $x_1 < 0$ (bad) ∴ $\boxed{x_1 = 1 - \frac{\mu}{3} + \frac{1}{3}\sqrt{\mu^2+9}}$

- For each μ, the opt sol'n to $\max_{x} f_\mu(x)$ is $x(\mu) := 1 - \frac{\mu}{3} + \frac{1}{3}\sqrt{\mu+}$
- $x(\mu)$ is always in the interior of (*)'s feas. region.
- As $\mu \downarrow 0$, $x(\mu) \to 1 + \frac{1}{3}\sqrt{9} = 1 + 1 = 2$, which is indeed optimal for the LP (*).

The same idea hold when we have more than one variables.

Ex 2:

$$\begin{aligned} \text{Max} \quad & 192 x_1 + 97 x_2 . \\ \text{s.t.} \quad & x_1 + x_2 \leq 100 . \\ & x_1, x_2 \geq 0 . \end{aligned}$$

$$\therefore \left.\begin{aligned} \text{Max} \quad & 192 x_1 + 97 x_2 \\ & 100 - x_1 - x_2 \geq 0 \\ & x_1 \geq 0 \\ & x_2 \geq 0 . \end{aligned}\right\} (**)$$

The barrier function:

$$f_\mu(x_1, x_2) = 192 x_1 + 97 x_2 + \mu\left[\log(100 - x_1 - x_2) + \log x_1 + \log x_2\right]$$

Remarks: as before

① $f_\mu(x_1, x_2)$ is defined only in the interior of the feasible region of (**).

③ $-f_\mu(x_1, x_2)$ is convex.

Can use Newton's method to find $x(\mu)$ that is optimal for: $\max_{x \in \mathbb{R}^n} f_\mu(x)$

② as $x \to$ boundary of feas. region, $f_\mu(x) \uparrow +\infty$.

④ $x(\mu)$ is feasible for (**), in particular, it is in the interior of the feasible region.

⑤ As $\mu \downarrow 0$, $x(\mu)$ approaches the optimal solution for the LP (**).

To solve for $x(\mu)$: find x such that $\nabla f_\mu(x) = 0$.

$$\left.\begin{aligned} &(A) \text{ ---- } \frac{\partial}{\partial x_1} f_\mu(x) = 192 - \frac{\mu}{100 - x_1 - x_2} + \frac{\mu}{x_1} = 0 . \\ &(B) \text{ ---- } \frac{\partial}{\partial x_2} f_\mu(x) = 97 - \frac{\mu}{100 - x_1 - x_2} + \frac{\mu}{x_2} = 0 \end{aligned}\right\} \text{keep on board}$$

For each fixed μ, solve for x_1, x_2.

$$\frac{x_1\mu - 100\mu + x_1\mu + x_2\mu}{(100-x_1-x_2)x_1} = 192.$$

$$\frac{x_2\mu - 100\mu + x_1\mu + x_2\mu}{(100-x_1-x_2)x_2} = 97.$$

$g_1 \rightarrow$ $192x_1^2 + (2\mu - 19200)x_1 - 100\mu + x_2\mu + 192x_1x_2 = 0.$

$g_2 \rightarrow$ $97x_2^2 + (2\mu - 9700)x_2 - 100\mu + x_1\mu + 97x_1x_2 = 0.$

(May not want to do on board:)

$$\nabla g_1(x_1,x_2) = \begin{pmatrix} 384x_1 + (2\mu - 19200) + 192x_2 \\ \mu + 192x_1 \end{pmatrix}$$

$$\nabla g_2(x_1,x_2) = \begin{pmatrix} \mu + 97x_2 \\ 194x_2 + (2\mu - 9700) + 97x_1 \end{pmatrix}$$

etc.

Summarize the algo:
Fix $\beta > 1$
Step 1: Choose initial $\mu^{(0)}$
~~Step 2: Find $x^{(0)} = x(\mu^{(0)})$ Decrease $\mu^{(0)}$: μ~~
Step 2: At ith iteration we have $\mu^{(i)}$
Solve: Max $f_{\mu^{(i)}}(x)$.
opt. soln: $x(\mu^{(i)})$
Step 3: Decrease $\mu^{(i)}$:
$\mu^{(i+1)} = \mu^{(i)}/\beta$.
Repeat step 2, 3

Note: Step 2 can be done via Newton's method.

Rewrite Algo w/ NM incorporated.

Check: for $\mu = 194$,
$x(\mu) = (x_1(\mu), x_2(\mu)) = (97, 2)$

Analysis & Convergence

$\rightarrow$ The objective value: $192 \cdot 97 + 97 \cdot 2 = 18818$.

$\rightarrow$ Actual optimal: $192 \cdot 100 + 97 \cdot 0 = 19200$
$x^* = (100, 0)$

- $x(194) = (97, 2)$ is quite close to $(100, 0)$
- $x(194)^T c = 18818$ is quite close to 19200
  the difference is $19200 - 18818 = 382$, small compared to 19200.

- What we will do next is a little bit of analysis to quantify what we mean by "$x(\mu)$ is close to x^*," the actual LP solution, as μ decreases to 0.

Consider the dual LP of (**):

$$\begin{aligned} \text{Min} \quad & 100y \\ \text{s.t.} \quad & y \geq 192 \\ & y \geq 97 \\ & y \geq 0 \end{aligned}$$

→ Rewriting w/ slack variables:

$$\left.\begin{aligned} \text{Min} \quad & 100y \\ \text{s.t.} \quad & y - t_1 = 192 \\ & y - t_2 = 97 \\ & y, t_1, t_2 \geq 0 \end{aligned}\right\} \quad (***)$$

<u>Recall from LP duality theory</u>

weak duality: $c^T x \leq b^T y \quad \forall\, x, y$ feasible (P), (D) soln.

So: for any x_1, x_2 feasible for (**)

y, t_1, t_2 feasible for (***),

$$192x_1 + 97x_2 \leq 100y$$

- Given (x_1, x_2), (y, t_1, t_2) <u>feasible</u> (but not necessarily complementary), the duality gap is: $100y - 192x_1 - 97x_2 \geq 0$.

- Coming back to our example,
  given μ, and $X(\mu)$ optimal for
  $\max f_\mu(x)$

① $X(\mu)$ is feasible for (**), the primal LP**

$X(\mu) = (\bar{X}_1, \bar{X}_2) = \bar{X}$

② We can easily obtain a feasible dual solution (y, t_1, t_2) : let

$$\bar{y} = \frac{\mu}{100 - \bar{x}_1 - \bar{x}_2}$$

$$\bar{t}_1 = \frac{\mu}{\bar{x}_1}$$

$$\bar{t}_2 = \frac{\mu}{\bar{x}_2}$$

From Ⓐ and Ⓑ,

$$192 - \bar{y} + \bar{t}_1 = 0$$
$$97 - \bar{y} + \bar{t}_2 = 0.$$

$$\therefore \quad \bar{y} - \bar{t}_1 = 192$$
$$\bar{y} - \bar{t}_2 = 97$$

so, $(\bar{y}, \bar{t}_1, \bar{t}_2)$ is feasible for (***).

③ The duality gap between $\bar{x}$ and $(\bar{y}, \bar{t}_1, \bar{t}_2)$:

$$100\bar{y} - 192\bar{x}_1 - 97\bar{x}_2$$

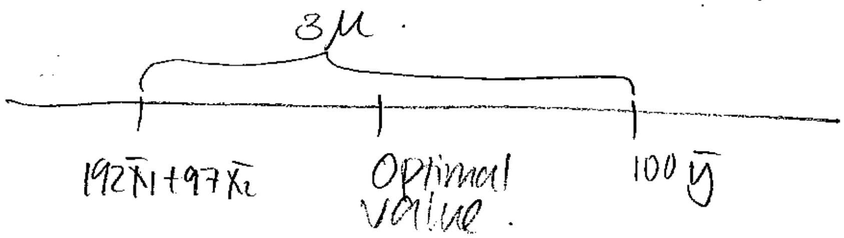
$$= 100\bar{y} - (\bar{y} - \bar{t}_1)\bar{x}_1 - (\bar{y} - \bar{t}_2)\bar{x}_2$$

$$= (100 - \bar{x}_1 - \bar{x}_2)\bar{y} + \bar{x}_1\bar{t}_1 + \bar{x}_2\bar{t}_2$$

$$= \mu + \mu + \mu = 3\mu.$$

Note that: for each μ:

$$192\bar{x}_1 + 97\bar{x}_2 \le \text{opt primal} = \text{opt dual} \le 100\bar{y}.$$



So, as $\mu \to 0$: