ORIE 6300 Mathematical Programming I	October 13, 2016
Lecture 15	
Lecturer: Damek Davis	*Scribe: Huanyu Zhang*

1 Last Time

Definition 1 (Bland's Rule) *Choose the entering basic variable x_j such that j is the smallest index with $\bar{c}_j < 0$. Also choose the leaving basic variable i with the smallest index (in case of ties in the ratio test).*

Theorem 1 (Termination with Bland's Rule) *If the simplex method uses Bland's rule, it terminates in finite number of steps with optimal solution. (i.e. no cycling)*

2 Low Cost Methods for LPs

The simplex method is extremely fast when inverting A_B is easy. In general, $O(m^3)$ operations are required to solve linear equation $A_B y = c_B$, etc. For cases when $O(m^3)$ is too large, say $m \geq 10^6$ and $m^3 \geq 10^{18}$, one iteration of simplex method is too expensive. There do exist approximate variants of simplex method that are cheaper, for example, column or row generation methods. See David Williamson's notes for more details. However, for extremely big problems, we sometimes resort first-order algorithms, which only use matrix vector products and no inversions. This brings the cost down to $O(mn)$ operations per iteration. With these methods, accuracy ε can be achieved in $O(\frac{1}{\varepsilon})$ iterations, which is not great, but its sometimes the best we can do.

Such algorithms look for fixed points of a nonlinear operator. For example, we saw in HW6 that

$$x^* \in \operatorname{argmin}\{c^T x \mid Ax = b, x \geq 0\} \qquad \text{and} \qquad y^* \in \operatorname{argmax}\{b^T y \mid A^T y \leq c\},$$

if and only if

$$T \begin{bmatrix} y^* \\ x^* \end{bmatrix} := \begin{bmatrix} y^* \\ x^* \end{bmatrix},$$

where

$$T \begin{bmatrix} x \\ y \end{bmatrix} := \begin{bmatrix} y - \gamma\left(Ax - b\right) \\ \max\{x + \tau\left(A^T y - c\right), 0\} \end{bmatrix},$$

When looking for fixed-points of nonlinear operators N, there is an obvious algorithm: given $Z^0 \in \mathbb{R}^n$, $N : \mathbb{R}^n \to \mathbb{R}^n$, iterate $Z^{k+1} = NZ^k$. Why? Assume $Z^k \to \bar{Z}$ as $k \to \infty$, then $NZ^k \to N\bar{Z} \Rightarrow N\bar{Z} = \bar{Z}$. So $\bar{Z}$ is a fixed point.

If N is linear, then $\bar{Z}$ is an eigenvector of N and the algorithm is called power method. For general N, this algorithm is called Krasnosel'skii-Mann (KM) iteration. The most popular variant of this algorithm includes a relaxation parameter $\lambda \in (0,1)$ and iterates $Z^{k+1} = (1-\lambda)Z^k + \lambda N Z^k$. It is not the only method used to solve nonlinear equations, but it is the most popular in convex optimization.

Definition 2 *We denote the set of fixed points of N by $Fix(N) := \{x \in \mathbb{R}^n | N(x) = x\}$.*

Exercise 1 *Exercise: Let $N_\lambda = (1-\lambda)I + \lambda N$, prove that $Fix(N_\lambda) = Fix(N)$.*

When does the KM iteration converge to an element of $Fix(N)$?

Theorem 2 (KM Iteration converges) *Suppose $N : \mathbb{R}^n \to \mathbb{R}^n$ is 1-Lipschitz continuous, i.e. $(\forall x \in \mathbb{R}^n, \forall y \in \mathbb{R}^n)$, $\|N_x - N_y\| \leq \|x - y\|$, that $Fix(N) \neq \varnothing$, and $\lambda \in (0,1)$. Then given any $Z^0 \in \mathbb{R}^n$, the sequence $\{Z^k\}_{k\in N}$ generated by the KM iteration $Z^{k+1} = (1-\lambda)Z^k + \lambda N Z^k$ converges to an elementent of $Fix(N)$*

Proof: Let $\bar{Z} \in Fix(N)$, then the sequence $\{Z^k\}_{k\in N}$ is bounded because

$$\begin{array}{rcl} \|Z^{k+1} - \bar{Z}\| & = & \|(1-\lambda)(Z^k - \bar{Z}) + \lambda(NZ^k - N\bar{Z}))\| \\ & \leq & (1-\lambda)\|Z^k - \bar{Z}\| + \lambda\|NZ^k - \bar{Z}\| \\ & \leq & (1-\lambda)\|Z^k - \bar{Z}\| + \lambda\|NZ^k - N\bar{Z}\| \\ & \leq & \|Z^k - \bar{Z}\|, \end{array}$$

and so $\|Z^k - \bar{Z}\| \leq \|Z^{k+1} - \bar{Z}\| \leq \|Z^0 - \bar{Z}\|$ and $\|Z^k - \bar{Z}\|$ converges to some $\xi \geq 0$. Thus we have $\{Z^k\}_{k\in N} \subseteq \mathbb{B}(\bar{Z}, \|\bar{Z} - Z^0\|)$. If $\{Z^k\}_{k\in\mathbb{N}}$ converges, then its limit is an element of $Fix(N)$ and the proof is complete.

Assume that $\{Z^k\}_{k\in N}$ doesn't converge.

Exercise 2 *If $\{Z^k\}_{k\in N}$ has a unique limit point, then it converges.*

Thus $\{Z^k\}_{k\in N}$ must have at least two limit points, say Z_1 and Z_2, and assume that $Z^{j_k} \to Z_1$ and $Z^{l_k} \to Z_2$ as $k \to \infty$.

Suppose that $Z_1, Z_2 \in \mathrm{Fix}(N)$, then let $\xi_1 = \lim_{k\to\infty} \|Z^k - Z_1\|$ and $\xi_2 = \lim_{k\to\infty} \|Z^k - Z_2\|$ (the limit exists by the argument at the beginning of the proof). We have (exercise!)

$$\begin{array}{rcl} 2\langle Z^k, Z_1 - Z_2\rangle & = & \|Z^k - Z_2\|^2 - \|Z_k - Z_1\|^2 + \|Z_2\|^2 - \|Z_1\|^2 \\ & \overset{k\to\infty}{\to} & \underbrace{\xi_2 - \xi_1 + \|Z_2\|^2 - \|Z_1\|^2}_{:=l} \end{array}$$

Thus

$$\begin{array}{rcl} 2\langle Z^{j_k}, Z_1 - Z_2\rangle & \to & 2\langle Z_1, Z_1 - Z_2\rangle \\ 2\langle Z^{l_k}, Z_1 - Z_2\rangle & \to & 2\langle Z_2, Z_1 - Z_2\rangle \end{array}$$

and

$$2\langle Z_1, Z_1 - Z_2\rangle = l = 2\langle Z_1, Z_1 - Z_2\rangle$$

so $\|Z_1 - Z_2\|^2 = 0$. Thus $Z_1 = Z_2$ and we have reached a contradiction.

Therefore, to complete the proof, we must show that every limit point of $\{Z^k\}_{k\in N}$ is actually a fixed point. To do this, we prove the two claims,

Claim 3 $\{\|Z^k - N(Z^k)\|\}_{k\in\mathbb{N}}$ *is nonincreasing.*

Claim 4 $(\forall Z \in \mathbb{R}^n, \forall \bar{Z} \in Fix(N))$, $\|N_\lambda(Z) - \bar{Z}\|^2 \leq \|Z - \bar{Z}\|^2 - \lambda(1-\lambda)\|Z - N(Z)\|^2$.

How do we use these claims? First Claim 4 implies that $\forall k \in \mathbb{N}$,

$$\begin{aligned} \|Z^{k+1} - Z^k\|^2 &= \|N_\lambda(Z^k) - Z^k\|^2 \\ &\leq \|Z^k - \bar{Z}\|^2 - \lambda(1-\lambda)\|Z^k - N(Z^k)\|^2 \end{aligned}$$

Then for $\forall T \in \mathbb{N}$,

$$\begin{aligned} \sum_{k=0}^{T} \lambda(1-\lambda)\|Z^k - N(Z^k)\|^2 &\leq \sum_{k=0}^{T}\{\|Z^k - \bar{Z}\|^2 - \|Z^{k+1} - \bar{Z}\|^2\} \\ &\leq \|Z^0 - \bar{Z}\|^2 - \|Z^{T+1} - \bar{Z}\|^2 \\ &\leq \|Z^0 - \bar{Z}\|^2 \end{aligned}$$

So $\forall T \in \mathbb{N}$, Claim 3 implies

$$\begin{aligned} \|Z^T - N(Z^T)\|^2 &\leq \tfrac{1}{T}\sum_{k=0}^{T} \|Z^k - N(Z^k)\|^2 \\ &\leq \tfrac{\|Z^0 - \bar{Z}\|^2}{T(1-\lambda)\lambda} \\ &\to 0 \text{ when } T \to \infty \end{aligned}$$

Thus, $\|Z^k - N(Z^k)\| \to 0$ as $k \to \infty$.

Now suppose that $\hat{Z}$ is a limit point of ${Z^k}_{k\in\mathbb{N}}$, say $Z^{j_k} \to \hat{Z}$. Then $\|Z^{j_k} - N(Z^{j_k})\| \to 0$ as $k \to \infty$. $\|\hat{Z} - N(\hat{Z})\| = 0 \Rightarrow \hat{Z} = N(\hat{Z})$ So any limit point of $\{Z^k\}_{k\in\mathbb{N}}$ is a fixed point.

Thus, the only loose ends left are the proofs of the Claims.

Proof: [of Claim 3]

$$\begin{aligned} & \|Z^{k+1} - N(Z^{k+1})\| \\ = \; & \|(1-\lambda)Z^k + \lambda N(Z^k) - N(Z^{k+1})\| \\ = \; & \|(1-\lambda)(Z^k - N(Z^k)) + N(Z^k) - N(Z^{k+1})\| \\ \leq \; & (1-\lambda)\|Z^k - N(Z^K)\| + \|N(Z^k) - N(Z^{k+1})\| \\ \leq \; & (1-\lambda)\|Z^k - N(Z^K)\| + \|Z^k - Z^{k+1}\| \\ \leq \; & (1-\lambda)\|Z^k - N(Z^K)\| + \|Z^k - [(1-\lambda)Z^k + \lambda N(Z^k)]\| \\ \leq \; & (1-\lambda)\|Z^k - N(Z^K)\| + \lambda\|Z^k - N(Z^k)\| \\ = \; & \|Z^k - N(Z^k))\| \end{aligned}$$

□

Proof: [of Claim 4]

$$\begin{aligned} \|N_\lambda(Z) - \bar{Z}\|^2 &= \|(1-\lambda)(Z - \bar{Z}) + \lambda(N(Z) - \bar{Z})\|^2 \\ \text{Exercise !} \quad &= (1-\lambda)\|Z - \bar{Z}\|^2 + \lambda\|N(Z) - \bar{Z}\|^2 - \lambda(1-\lambda)\|Z - N(Z)\|^2 \\ &= (1-\lambda)\|Z - \bar{Z}\|^2 + \lambda\|N(Z) - N(\bar{Z})\|^2 - \lambda(1-\lambda)\|Z - N(Z)\|^2 \\ &\leq \|(Z - \bar{Z})\|^2 - \lambda(1-\lambda)\|Z - N(\bar{Z})\|^2 \end{aligned}$$

□ □