Nonlinear Optimization Lecture 23

Garrick Aden-Buie

Thursday, April 14, 2016

Diagonalization Algorithm

Step 0 *Initialization.* Determine an initial feasible solution $x^0 \in \Omega$ to $\min 0$ s.t. $x \in \Omega$ and set $k = 0$.

Step 1 *Solve the diagonalized VI.* Compute

$$\begin{aligned} & F_i^k(x_i) && \forall i = 1, \dots, n \\ (VI^k) \quad & \sum_{i=1}^{n} F_i^k(x_i^*)(x_i - x_i^*) \geq 0 && \forall x \in \omega \end{aligned}$$

Note. The VI problem is an optimization problem: $\min \sum_{i=1}^{n} \int_0^x F_i(z_i) dz_i$ s.t. $x_i \in \Omega$. Let $k^{k+1} = x^*$.

Step 2 If $\|x^{k+1} - x^k\| < \epsilon$, stop. Otherwise, set $k = k + 1$ and go to **Step 1**.

Fixed-Point Algorithm

$$VI(F, \Omega) \Leftrightarrow FPP_{\min}(F, \Omega)$$

The fixed-point problem gives the following relationship:

$$\begin{aligned} y - F(y) \quad \Leftrightarrow & \quad y \in \Omega \\ & \quad y = P_\Omega[y - \alpha F(y)], \alpha \text{ is a constant} \\ \Rightarrow & \quad y^{k+1} = P_\Omega[y^k - \alpha F(y^k)] \end{aligned}$$

where the intuition is that because of the fixed-point nature, if you scale $F(y)$ by α and project onto Ω, you remain at the point y (hence, *fixed-point*).

Step 0 *Initialization.* Determine an initial feasible solution $x^0 \in \Omega$ to $\min 0$ s.t. $x \in \Omega$ and set $k = 0$.

Step 1 Given x^k, perform the projection

$$\begin{aligned} y^{k+1} &= P_\Omega[y^k - \alpha F(y^k)] \\ x^{k+1} = x^* &\Leftarrow \begin{cases} \min & \|x^k - \alpha F(x^k) - x\|^2 \\ \text{s.t.} & x \in \Omega \end{cases} \end{aligned}$$

Note. α needs to be sufficiently small but there are no great guidelines on how to set and manipulate α.

Step 2 If $\|x^{k+1} - x^k\| < \epsilon$, stop. Otherwise, set $k = k + 1$ and go to **Step 1**.

Gap function[1]

Using $VI(F, \Omega)$, the gap function is used to check the solution's correctness:

Gap functions must satisfy the following

- $G(x) \geq 0 \; \forall x \in \Omega$
- $G(x) = 0$ if and only if x solves $VI(F, \Omega)$

Theorem

- Ω, a closed convex set in $\mathbb{R}^n$

Assume

- that $[F(x) - F(y)]^T(x - y) \geq \mu\|x - y\|^2 > 0$ for some $\mu > 0 \; \forall x, y \in \Omega$ (strongly monotone with modulus μ).
- $\|F(x) - F(y)\| \leq L\|x - y\|$ for some $L > 0 \; \forall x, y \in \Omega$.
  - *Lipschitz continuous*: strong form of uniform continuity.
  - The idea is that given some x, y we can always have some idea of worst-case scenario of how far apart their function values are. The existence of L provides an upper bound on the distance between the function values of any two points x, y.

If $\alpha < \frac{2\mu}{L^2}$
Then the fixed-point algorithm converges to the unique solution of the VI.

Note that α is difficult to compute, but this theorem proves that with certain values convergence of the fixed-point algorithm is guaranteed. **But** even if α is *greater* than $\frac{2\mu}{L^2}$, convergence may still occur.

To prove this theorem we need the following lemma.

Lemma

$$\|P_\Omega[x] - P_\Omega[y]\| \leq \|x - y\| \;\; \forall x, y \in \mathbb{R}^n$$

where $\Omega \subset \mathbb{R}^n$ is convex and closed.

Remark. The projection operator is a *non-expansive* mapping, i.e. $\|G(x) - G(y)\| < \|x - y\|$.

[1] Sometimes called a *Merit Function*.

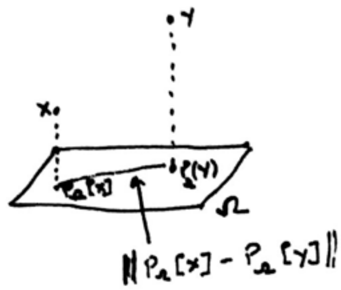

Figure 1: Illustration of lemma.

Nework User Equilibrium

Consider a set of paths through a network of nodes. Let p_i be one path through the system, where each path is a numbered connection between nodes. Let h_{p_i} be the flow rate in each path and $c_{p_i}(h)$ the path cost as a function of the flow rate.

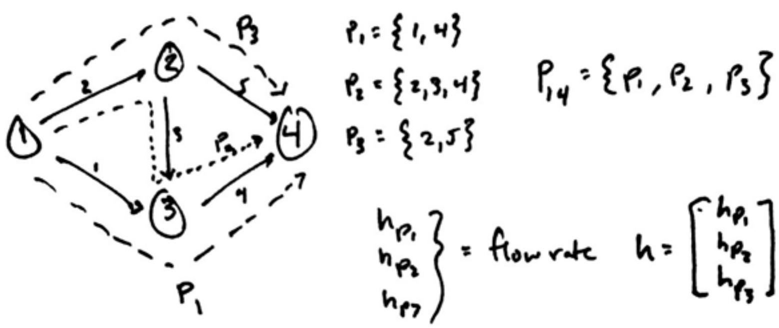

Figure 2: Network Illustration

Definition: Wardrop's First Principle. A vector h is a user equilibrium traffic flow pattern if the following are true:

$$h_p > 0 \;\; \forall p \in P_{ij} \Rightarrow c_p = \min_{p' \in P_{ij}} c_{p'}$$

for all $(i, j) \in W$, the set of OD pairs (origination, destination).

In other words, if the flow rate on over a particular path, then the cost along that path must be the minimum (if not the minimum then people would not be using the path).

The converse is also implied:

$$c_p(h) > \min_{p' \in P_{ij}} c_{p'}(h) \Rightarrow h_p = 0$$