Lecture 8

Plus properties, merit functions and gap functions

September 28, 2008

Outline

- Plus-properties and F-uniqueness
- Equation reformulations of VI/CPs
- Merit functions
- Gap merit functions

FP-I book: Sections 1.5, 2.3.1 and 2.3.2

Plus properties and F-uniqueness

- Last time we provided a result regarding the singleton nature of F(SOL(K,F)).

- The solution set SOL(K,F) is said to be **F-unique** if F(SOL(K,F)) is at most a singleton.

- What classes of functions yield this property?

Plus and co-coercive mappings

Definition 1 *A mapping* $F : K \to \mathbb{R}^n$ *is said to be*

(a) ***pseudo-monotone-plus*** *on* K *if it is pseudo-monotone on* K *and for all* $x, y \in K$,

$$[(x-y)^T F(y) \geq 0 \text{ and } (x-y)^T F(x) = 0] \quad \Longrightarrow \quad F(x) = F(y).$$

(b) ***monotone-plus*** *on* K *if it is monotone on* K *and for all* $x, y \in K$,

$$(x-y)^T (F(x) - F(y)) = 0 \quad \Longrightarrow \quad F(x) = F(y).$$

(c) ***co-coercive*** *on* K *if there exists a constant* $c > 0$ *such that*

$$(F(x) - F(y))^T (x - y) \geq c \|F(x) - F(y)\|^2, \quad \forall x, y \in K.$$

Proposition 1 *Let $F : K \rightarrow \mathbb{R}^n$ be pseudo-monotone plus on the convex set K. Then the solution set SOL(K,F) is F-unique.*

Proof: By Prop. 2.3.6, for a pseudo-monotone F and any two solutions $x^1, x^2 \in SOL(K, F)$, we have

$$(x^1 - x^2)^T(F(x^1) - F(x^2)) = 0.$$

By "plus" property, it follows $F(x^1) = F(x^2)$.

Implications

- If F is monotone plus it is pseudo-monotone plus.
- If F is co-coercive on K, then it is monotone and Lipschitz continuous on K (also, nonexpansive when $c \geq 1$).
- Converse holds when F is a gradient map $(F = \nabla\theta)$:
  a monotone Lipschitz continuous gradient map over an open convex set C is co-coercive over C.

Composite Monotone Mappings

Definition 2 *Consider a mapping $F : K \to \mathbb{R}^n$ given by*

$$F(x) = A^T G(Ax) + b \qquad \textit{for all} x \in K,$$

where $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^n$, and $G : R_A \to \mathbb{R}^m$ with R_A being the range of matrix A. The mapping F is said to be

(a) ***monotone composite*** *if the mapping G is monotone on the range R_A.**

(b) ***L-monotone composite*** *if the mapping G is Lipschitz continuous and monotone on the range R_A.*†

*strictly (strongly) monotone composite if G is strictly (strongly) monotone on R_A

†L-strictly (strongly) monotone composite if the mapping G is Lipschitz continuous and strictly (strongly) monotone on R_A

Implications

- If F is strictly monotone composite on K, then it is monotone plus on K.
- If F is L-strongly monotone composite on K, then it is co-coercive there.
  Proof:

$$\begin{aligned}(x-y)^T(F(x)-F(y)) &= (Ax-Ay)^T(G(Ax)-G(Ay)) \\ &\geq c\|Ax-Ay\|^2;\end{aligned}$$

$$\|F(x)-F(y)\|^2 = \|A^T(G(Ax)-G(Ay))\|^2 \leq \|A\|^2L^2\|Ax-Ay\|^2$$

$$\|Ax-Ay\|^2 \geq \frac{1}{L^2\|A^T\|^2}\|F(x)-F(y)\|^2$$

$$(F(x)-F(y))^T(x-y) \geq \frac{c}{L^2\|A^T\|^2}\|F(x)-F(y)\|^2.$$

Description of SOL(K,F)

If SOL(K,F) is F-unique, it may be described simply based on the following result.

Proposition 2 *[Prop. 2.3.12 FP-I] Suppose that $F(SOL(K,F)) = \{w\}$. Then*

$$SOL(K,F) = F^{-1}(w) \cap \arg\min\{x^T w : x \in K\}$$

In addition, if K is polyhedral, then the KKT multiplier set $\mathcal{M}(x)$ is a polyhedron that is independent of the optimal point $x \in SOL(K,F)$.

Proof:

(i) Let $x \in SOL(K,F)$ and define $S = F^{-1}(w) \cap \arg\min\{x^T w : x \in K\}$. Therefore, we have $x \in F^{-1}(w)$. Moreover, since x is a solution to $VI(K,F)$, we have

$$(y-x)^T F(x) \geq 0 \quad \forall y \in K \qquad \Longleftrightarrow \qquad y^T w \geq x^T w \quad \forall y \in K.$$

Therefore x is a solution **iff** it solves the problem $\min_{y \in K} y^T w$.
Let $x \in S$. Then, $F(x) = w$ and $x \in SOL(K, F)$. Therefore SOL(K,F) = S.

(ii) Let $K \equiv \{x : Ax \leq b\}$ for some A, b. Then, we claim that for every $x \in SOL(K, F)$, the KKT multiplier set $\mathcal{M}(x)$ is the optimal solution set of the following linear problem:

$$(DP) \qquad \begin{array}{ll} \text{minimize} & b^T \lambda \\ \text{subject to} & w + A^T \lambda = 0 \\ & \lambda \geq 0. \end{array}$$

This set of multipliers is polyhedral and independent of x. The dual of this linear problem is given by

$$(LP) \qquad \begin{array}{ll} \text{minimize} & y^T w \\ \text{subject to} & Ay \leq b. \end{array}$$

We need to show that $\lambda \in \mathcal{M}(x)$ if and only if λ is an optimal solution to (DP).

- Let $\lambda \in \mathcal{M}(x)$ for some $x \in SOL(K, F)$. Then x is a solution to (LP) and λ is the associated dual variable satisfying $w + A^T\lambda = 0$. Moreover, by complementary slackness, we have $\lambda \perp Ax - b$. From LP duality, it follows that λ is an optimal solution to (DP).
- Let $\lambda \in SOL(DP)$. Then, we have $x \in SOL(K, F)$. But $\lambda \perp Ax - b$, establishing that $\lambda \in \mathcal{M}(x)$.

Polyhedrality of SOL(K,F)

- Note that we only assert the polyhedrality of the KKT multiplier set $\mathcal{M}(x)$ when K is polyhedral.
- We do not know whether the solution set $SOL(K, F)$ is polyhedral
- When F is a strictly monotone composite mapping on K, then $SOL(K, F)$ is polyhedral provided that K is polyhedral.
- Interesting - we do not need F to be affine nor G Lipschitz
- This allows us to construct error bounds for VIs of this class

Corollary 3 *[Cor. 2.3.13 of FP-I] Let F be a strictly monotone composite mapping on K. If $SOL(K, F) \neq \emptyset$, then $A(SOL(K, F))$ is a singleton given by some $\{v\}$. Furthermore, it holds that*

$$SOL(K, F) = A^{-1}v \cap \arg\min\{x^T w : x \in K\},$$

where w is the single element of $F(SOL(K, F))$.
If in addition, K is polyhedral so is $SOL(K, F)$.

Equation reformulations of NCP

- VI/CP does not always arise from optimization problems
- Equivalent formulations of VI/CPs as systems of nonlinear equations
- Beneficial for analytical and computational purposes
- Consider NCP(F) which is to find a an x such that $0 \leq x \perp F(x) \geq 0$.
- An equation reformulation is given by

$$x \in SOL(F) \quad \Leftrightarrow \quad H(x,w) \equiv \begin{pmatrix} w - F(x) \\ w \circ x \end{pmatrix} = 0,$$

  with additional restriction that $x \geq 0$ and $w \geq 0$. Here $\circ$ represents the *Hadamard* product; $a \circ b = \begin{pmatrix} a_1 b_1 \\ \vdots \\ a_n b_n \end{pmatrix}$.
- Square set of equations that inherit differentiability properties of F
- Suitable for iterative methods for solving the system $H(x,w) = 0$.

Complementarity or C-Functions

Definition 3 *A function $\psi : \mathbb{R}^2 \to \mathbb{R}$ is called a* ***C-function****, if for any pair $(a,b) \in \mathbb{R}^2$,*

$$\psi(a,b) = 0 \Leftrightarrow [(a,b) \geq 0 \text{ and } ab = 0];$$

equivalently,

ψ is a C-function if the set of zeros are the nonnegative axes in $\mathbb{R}^2$.

Examples:

- min function: $\psi_{\min}(a,b) = \min\{a,b\}$.
- Fischer-Burmeister function
$$\psi_{\mathrm{FB}}(a,b) = \sqrt{a^2+b^2} - (a+b), \qquad \forall (a,b) \in \mathbb{R}^2$$
- For a scalar $\tau > 0$, a variant of ψ_{FB} is given by for all $(a,b) \in \mathbb{R}^2$,
$$\psi_{\mathrm{CCK}}(a,b;\tau) = \sqrt{a^2+b^2} - (a+b) - \tau \max\{0,a\} \max\{0,b\}.$$

Equation reformulations of VI

- The min function

$$\mathbf{F}_{\min}(x) = \mathrm{min}(x, F(x)) = x + \mathrm{min}(0, F(x) - x) = x - \mathrm{max}(0, x - F(x)).$$

- The max function $\mathrm{max}(0, x)$ is a Euclidean projector onto the nonnegative orthant and is denoted by $\Pi_{\mathbb{R}^n_+}$

$$\mathbf{F}_{\min}(x) = x - \Pi_{\mathbb{R}^n_+}(x - F(x)).$$

- Euclidean projection of x on K is denoted by $\Pi_K(x)$, the solution to the problem

$$\begin{aligned} \min \quad & \tfrac{1}{2}(y - x)^T(y - x) \\ \text{subject to} \quad & y \in K. \end{aligned}$$

- If K is polyhedral, this is a convex QP; K is not polyhedral - generally nontrivial to compute $\Pi_K(x)$

- The Euclidean projector has several important properties for closed convex nonempty K: **Proof - homework**
  - For each $x \in \mathbb{R}^n$, the projection $\Pi_K(x)$ exists and it is unique
  - For each $x \in \mathbb{R}^n$, the projection $\Pi_K(x)$ is the unique vector $\bar{x}$ satisfying

$$(y - \bar{x})^T(\bar{x} - x) \geq 0 \qquad \forall\, y \in K.$$

  - As a function of x, the projection mapping $\Pi_K(x)$ is co-coercive and nonexpansive in that for all $x, y \in \mathbb{R}^n$,

$$(\Pi_K(x) - \Pi_K(y))^T(x - y) \geq \|\Pi_K(x) - \Pi_K(y)\|^2,$$
$$\|\Pi_K(x) - \Pi_K(y)\|_2 \leq \|x - y\|_2.$$

  (Π_K is a globally Lipschitz continuous function)

Natural maps and min-functions

- Recall that $x \in SOL(NCP(F)) \iff \min(x, F(x)) = 0$ can be written as

$$x - \Pi_{\mathbb{R}^n_+}(x - F(x)) = 0.$$

- Leads to a generalization

- Consider a more general closed convex set $K \subseteq \mathbb{R}^n$

Proposition 4 *Let $K \subseteq \mathbb{R}^n$ be closed convex and $F : K \to \mathbb{R}^n$ be arbitrary. It holds that*

$$[x \in SOL(K, F)] \Leftrightarrow [\mathbf{F}^{nat}_K(x) = 0],$$

where

$$\mathbf{F}^{nat}_K(v) \equiv v - \Pi_K(v - F(v)).$$

Proof: (See Lecture 5)

Proposition 5 *Let $K \subseteq \mathbb{R}^n$ be a closed convex set and $F : K \to \mathbb{R}^n$ be arbitrary. It holds that $x \in SOL(K, F)$ if and only if there exists a vector z such that*

$$x = \Pi_K(z) \qquad \textit{and} \qquad \mathbf{F}_K^{nor}(z) = 0,$$

where

$$\mathbf{F}_K^{nor}(v) \equiv F(\Pi_K(v)) + v - \Pi_K(v).$$

Proof:

- Suppose that $x \in SOL(K, F)$. Then $x = \Pi_K(x - F(x))$ by Proposition 4. Let $z = x - F(x) \implies x = \Pi_K(z)$.

$$\begin{aligned} \mathbf{F}_K^{\rm nor}(z) &\equiv F(\Pi_K(z)) + z - \Pi_K(z) \\ &= F(x) + x - F(x) - x = 0. \end{aligned}$$

- Given a vector z such that $x = \Pi_K(z)$ and $\mathbf{F}_K^{\rm nor}(z) = 0$.

$$\begin{aligned} 0 &= \mathbf{F}^{\rm nor}(z) \\ &= F(\Pi_K(z)) + z - \Pi_K(z) \\ &= F(x) + z - x \\ \implies \quad z &= x - F(x). \end{aligned}$$

Thus, $x = \Pi_K(z)$ and $z = x - F(x)$, implying

$$x = \Pi_K(x - F(x)) \qquad \Longleftrightarrow \qquad \mathbf{F}_K^{\rm nat}(x) = 0.$$

By Proposition 4, it follows $x \in SOL(K, F)$.

- The two equation reformulations of VI's are

$$\mathbf{F}_K^{\mathrm{nat}}(x) = 0$$

and

$$\mathbf{F}_K^{\mathrm{nor}}(z) = 0.$$

- The natural map is on the original space while the normal map has a change of variable and is on the whole space, viz. $\mathbb{R}^n$.

- The latter, given its definition over the whole space, is useful for computations

Merit functions

- Just presented several reformulations of VIs and CPs using equation based reformulations

- A contrasting approach casts the problem as a minimization problem through the use of a merit function

- Consider $NCP(F)$: $x \in SOL(F)$ if and only if x is a solution to the following problem with optimal value zero:

$$\begin{aligned} \text{min} \quad & y^T F(y) \\ \text{subject to} \quad & y \geq 0 \\ & F(y) \geq 0. \end{aligned}$$

- The function $y^T F(y)$ is a merit function for $NCP(F)$

Definition 4 *A merit function for the $VI(K, F)$ on a closed convex set $X \supseteq K$ is a nonnegative function $\theta : X \to \mathbb{R}_+$ such that $x \in SOL(K, F)$ if and only if $x \in X$ and $\theta(x) = 0$; that is solutions of $VI(K, F)$ coincide with the global solutions of*

$$\begin{aligned} \textit{minimize} \quad & \theta(y) \\ \textit{subject to} \quad & y \in X, \end{aligned}$$

whose optimal value is $\theta^* = 0$.

- If $SOL(K, F)$ is empty, then the optimal value $\theta^* > 0$ or θ has no global minimizer over X
- Other merit functions:
  - If $H(x) \equiv 0$ is an equation reformulation, then as a merit function we can use $\theta(x) = \|H(x)\|^r$ for $x \in X$, where r is a positive integer
  - The min-merit function $[NCP(F)]$: $\theta_{\min}(x) = \sum_{i=1}^n (\min(x_i, F_i(x)))^2$.
  - If ψ is a C-function, then $\theta_\psi(x) \equiv \|\mathbf{F}_\psi(x)\|_2^2 = \sum_{i=1}^n \psi(x_i, F_i(x))^2$.

The Gap merit function

- Consider a merit function that is not a consequence of equation-based reformulations
- We term it as **gap** function and define it on the domain $\mathcal{D}$ of F and is given by

$$\theta_{\rm gap}(x) \equiv \sup_{y \in K} F(x)^T(x-y), \quad x \in \mathcal{D} \supseteq K.$$

- Specifically if K is a cone, we have

$$\theta_{\rm gap}(x) = \begin{cases} F(x)^T x & \text{if } F(x) \in K^* \\ +\infty & \text{otherwise }. \end{cases},$$

where K^* is a dual cone for K.

- In general, $x \in SOL(K, F)$ if and only if x is a global solution of the constrained gap minimization problem with $\theta_{\rm gap}(x) = 0$.:

$$\begin{aligned} \min \quad & \theta_{\rm gap}(z) \\ \text{subject to} \quad & z \in K. \end{aligned}$$

- When K is a cone, then $VI(K, F) \equiv CP(K, F)$ and the gap function reduces to

$$\begin{aligned} \min \quad & x^T F(x) \\ \text{subject to} \quad & x \in K \\ & F(x) \in K^*. \end{aligned}$$

- Note that this problem is tractable if F is smooth and K is polyhedral for instance

A Dual Gap function

- We may also define a dual gap function as follows:

$$\theta_{\rm dual}(x) \equiv \inf_{y \in K} F(y)^T(y - x), \quad x \in \mathbb{R}^n.$$

- Refer to the earlier one as the primal gap function

- Following inequalities hold:

$$-\infty \le \theta_{\rm dual}(x) \le 0 \le \theta_{\rm gap}(x) \le \infty \quad \forall x \in K.$$

- The primal gap function is defined for $x \in \mathcal{D}$ while the dual gap function is defined for all x

- Recall that for a function f,

$$epi(f) = \{(x, w) : f(x) \leq w\}$$

and f is a convex function if and only if $epi(f)$ is a convex set.

- The hypograph of $\theta_{\rm dual}$ is given by

$$\cap_{y \in K}\{(x, \eta) : F(y)^T(y - x) \geq \eta\}.$$

This is an intersection of a family of halfspaces, therefore it is convex implying that the $\theta_{\rm dual}$ is concave function

- The evaluation of $\theta_{\rm dual}(x)$ requires solving

$$\begin{aligned} \text{max} \quad & (y - x)^T F(y) \\ \text{subject to} \quad & y \in K. \end{aligned}$$

- generally non-convex (unless F is an affine monotone function)

- We may introduce a dual-gap program given by

$$\begin{aligned} \min \quad & \theta_{\rm dual}(x) \\ \text{subject to} \quad & x \in K. \end{aligned}$$

  - Since $\theta_{\rm dual}$ is a concave function, this is a concave program
  - Hidden difficulty is in the evaluation of $\theta_{\rm dual}$

- Conclude with the following result that follows from some earlier propositions:

Proposition 6 *Let K be closed and convex and $F : K \to \mathbb{R}^n$ be continuous. If F is pseudo-monotone on K, then $x \in SOL(K, F)$ if and only if x is a global maximizer of*

$$\begin{aligned} \max \quad & \theta_{dual}(x) \\ \textit{subject to} \quad & x \in K. \end{aligned}$$

and $\theta_{dual}(x) = 0$.

Proof: Recall the relation

$$SOL(K, F) = \cap_{y \in K} \{x \in K \mid F(y)^T(y - x) \geq 0\}.$$