Numerical Methods for Complementarity Problems, Variational Inequalities, and Extended Systems

Jong-Shi Pang

Department of Industrial and Enterprise Systems Engineering

University of Illinois at Urbana-Champaign

presented at

CIMPA-UNESCO-VIETNAM School
Variational Inequalities and Related problems
Monday May 10, 2010, 2:00–3:45 PM

Lecture I: Linear Complementarity Problems

Contents of Presentation

- Classes of Methods

— Quadratic programming based methods

— Direct (pivoting) methods

— Matrix splitting methods

— Nonsmooth-equation based methods

— Smoothing methods

- This Lecture : Lemke's Pivoting Method

— Description of the method

— A new termination result

— Special cases

Classes of Methods: A Synopsis

- Quadratic Programming Based Methods

— for symmetric problems

— hybrid between simple iterations and active-set methods

- Pivoting methods

— solving linear equations under various pivot rules

— principal pivoting, Lemke, and many variations

— (Lemke) besides broad applicability, motivated study of matrix classes

— relying on numerical linear algebraic methods for practical implementation

- Matrix splitting methods

— based on a simple fixed-point formulation

— projected Jacobi, Gauss-Seidel, and successive overrelaxation

— simple projection, Tikhonov, and proximal regularization

— restricted classes of problems, but amenable to distributed implementation

Classes of Methods (cont.)

- Nonsmooth-equation based methods

— based on the theory of non-differentiable C(omplementarity)-functions

— extending classical Newton methods for differentiable equations

— globalization via line searches or trust-region methods

— extendable to nonlinear complementarity problems

- Smoothing methods

— smoothing the C-functions and apply methods for differentiable equations

— interior-point methods

— applicable to solvable problems of the P_0-type

— extendable to nonlinear complementarity problems

Lemke's method (1965) in revised form

Given a vector $q \in \mathbb{R}^n$ and matrix $M \in \mathbb{R}^{n\times n}$, the linear complementarity problem, denoted LCP (q, M), is to find $z \in \mathbf{R}^n$ such that

$$0 \leq z \perp w \triangleq q + Mz \geq 0,$$

where $\perp$ is the perpendicularity notation, which in this case denotes the complementarity condition between the two vectors z and w.

Initialization. The method works with the augmented $n \times (2n + 1)$ linear system:

$$Ax \triangleq \left[\begin{array}{ccc} I & -M & -d \end{array} \right] \begin{pmatrix} w \\ z \\ \tau \end{pmatrix} = q,$$

where d is a nonnegative vector such that a scalar $\bar{\tau} > 0$ exists such that $q + \tau d \geq 0$ for all $\tau \geq \bar{\tau}$. The method starts at a basic feasible solution (bfs) of the system: $Ax = q$ and $x \geq 0$ such that $z^T w = 0$ and $\tau > 0$ is basic. Such a bfs is called almost complementary (a.c.).

The initial a.c. bfs has $(w, z, \tau) = (q + \bar{\tau}d, 0, \bar{\tau})$.

For each $i = 1, \cdots, n$, (w_i, z_i) is called a complementary pair.

Lemke's method (1965) in revised form (cont.)

The initial ratio test:

$$\max_{1 \le i \le n} \left\{ -\frac{q_i}{d_i} \right\} = -\frac{q_{i_0}}{d_{i_0}}.$$

The initial basis B is the identity matrix with w is the basic variables; τ as the first entering (nonbasic) variable; and w_{i_0} is the first outgoing (basic) variable.

Pivot rule. Choose as the next entering variable the nonbasic variable whose (basic) complement has just become nonbasic.

General iteration. Geometrically, the method moves from one a.c. bfs to an adjacent a.c. bfs. Algebraically, this is accomplished by a minimum ratio test followed by a pivot step, which is carried out by solving 2 systems of linear equations with the same basis matrix.

Let B be the current basis such that $\boxed{\bar{q} \triangleq B^{-1}q \ge 0}$; let $A_{\bullet j}$ be the entering column. Let $\boxed{\bar{A}_{\bullet j} \triangleq B^{-1}A_{\bullet j}}$.

Minimum ratio test: $\min \left\{ \frac{\bar{q}_i}{\bar{A}_{ij}} \mid \bar{A}_{ij} > 0 \right\} = \frac{\bar{q}_{i_0}}{\bar{A}_{i_0 j}}$ if there is at least one positive $\bar{A}_{i_0 j}$.

Termination criteria.

1. If $\bar{A}_{ij} \leq 0$ for all i, method terminates on a ray.

2. Otherwise, a new basis B^{new} is obtained with one exchange of one basic column with $A_{\bullet j}$.

3. Method continues until τ becomes nonbasic; at that time, a complementarity solution to the LCP (q, M) is successfully obtained.

Analysis.

1. Does method terminate in finite number of iterations?

2. What does ray termination mean for the LCP?

3. What classes of LCPs ensure successful termination?

A new termination results

Semi-copositive matrices, called semi-monotone in CPS (1992,2009).

A matrix $M \in \mathbb{R}^{n \times n}$ is **semi-copositive** if for every $x \in \mathbb{R}^n_+ \setminus \{0\}$, an index i exists such that $x_i > 0$ and $(Mx)_i \geq 0$.

Theorem. If M is semi-copositive and $\bigcup_{\tau > 0} \text{SOL}(q + \tau d, M)$ is bounded, then Lemke's method will successfully compute a solution of the LCP (q, M) in a finite number of pivots.

Numerical Methods for Complementarity Problems, Variational Inequalities, and Extended Systems

Jong-Shi Pang

Department of Industrial and Enterprise Systems Engineering

University of Illinois at Urbana-Champaign

presented at

CIMPA-UNESCO-VIETNAM School
Variational Inequalities and Related problems
Wednesday May 12, 2010, 2:00–3:45 PM

Lecture II: Newton Methods and Globalization

Contents of Presentation

- Affine variational inequalities

— conversion to linear complementarity problems

- The Josephy-Newton method

— sequential linearization

— convergence under solution (strong) stability

— conditions for (strong) stability

— globalization for monotone problems

- Semismooth Equations

— motivation

— semismooth functions

— semismooth Newton methods

— application to the NCP

Affine variational inequalities

Given a polyhedron $K \triangleq \{x \in \mathbb{R}^n \mid Ax \leq b\}$, a vector $q \in \mathbb{R}^n$, and a matrix $M \in \mathbb{R}^{n \times n}$, the affine variational inequality, denoted AVI (K, q, M), is to find $z \in \mathbf{R}^n$ such that

$$z \in K \quad \text{and} \quad (z' - z)^T(q + Mz) \geq 0 \; \forall z' \in K.$$

- The LCP (q, M) is a special case with $K = \mathbb{R}^n_+$.
- $z \in \text{SOL}(K, q, M)$ if and only if $z \in \underset{z' \in K}{\text{argmin}}\,(q + Mz)^T z'$.
- $z \in \text{SOL}(K, q, M)$ if and only if $\exists\ \lambda$ such that the Karush-Kuhn-Tucker (KKT) conditions hold:

$$\left.\begin{array}{rcl} 0 & = & q + Mz + A^T\lambda \\ 0 & \leq & \lambda \perp b - Az \geq 0 \end{array}\right\}, \quad \text{which is a mixed LCP.}$$

- Thus, **AVI $\Leftrightarrow$ mixed LCP.**

- The equivalent mixed LCP is beneficial for many purposes; but, it sometimes imposes un-necessarily restrictive conditions on the AVI.

Conversion of a mixed LCP to a standard LCP

- M is nonsingular: solving for $z = -M^{-1}(q + A^T\lambda)$ and substituting, yields

$$0 \leq \lambda \perp \bar{b} + \bar{M}\lambda \geq 0, \quad \text{where} \quad \bar{M} \triangleq AM^{-1}A^T.$$

- The matrix $\bar{M}$ is (symmetric) positive semidefinite if M is so.

- More generally, if there exists a set B of rows of A such that the matrix

$$\begin{bmatrix} M & B^T \\ -B & 0 \end{bmatrix}$$

is nonsingular, a similar conversion can be carried out, preserving the positive semidefiniteness of M if applicable.

- As a result, Lemke's method is applicable.

- A version of this method exists for solving the AVI that does not require the conversion.

The Josephy-Newton Method for Variational inequalities

Given a closed convex $K \subseteq \mathbb{R}^n$ and a differentiable self-mapping $\Phi : \mathbb{R}^n \to \mathbb{R}^n$,

$$\text{VI}(K, \Phi) : z \in K \;\text{ and }\; (z' - z)^T \Phi(z) \geq 0 \;\; \forall z' \in K.$$

Given $z^\nu \in K$, let $z^{\nu+1}$ be an appropriate solution of the the semi-linearized VI (K, Φ^ν), where, with $J\Phi(z^\nu)$ denoting the Jacobian matrix of Φ at z^ν,

$$\Phi^\nu(z) \triangleq \Phi(z^\nu) + J\Phi(z^\nu)(z - z^\nu),$$

is the linearization of Φ at z^ν.

- A further approximation is needed when K is not polyhedral.
- The linearization idea can be extended to a point-based approximation (PBA); i.e., $\Phi \approx \Psi$ near z^*, meaning

$$\lim_{z^* \neq z \to z^*} \frac{\Phi(z) - \Psi(z)}{\| z - z^* \|} = 0.$$

Example. A composite nonsmooth function: $\Phi(z) \triangleq \Xi \circ F(z)$, where Ξ is Lipschitz and F is differentiable; then Ψ is a PBA of Φ at z^*, where

$$\Psi(z) \triangleq \Xi \circ [\, F(z^*) + JF(z^*)(z - z^*) \,]$$

Convergence under Solution Stability

A solution $x^* \in \text{SOL}(K, \Phi)$ is said to be stable if

(a) an open neighborhood $\mathcal{N}$ of x^* exists satisfying $\mathcal{N} \cap \text{SOL}(K, \Phi) = \{x^*\}$,

(b) for any such neighborhood $\mathcal{N}$, there exist two positive scalars c and ε such that, for every continuous Ψ satisfying

$$\sup \{ \| \Phi(x) - \Psi(x) \| \mid x \in K \cap \mathcal{N} \} \leq \varepsilon, \tag{1}$$

$\text{SOL}(K, \Psi) \cap \mathcal{N} \neq \emptyset$; moreover,

$$\| x - x^* \| \leq c \| \Phi(x) - \Psi(x) \| \quad \forall x \in \text{SOL}(K, \Psi) \cap \mathcal{N}.$$

The solution x^* is said to be strongly stable if

(a) x^* is stable, and

(b) for every neighborhood $\mathcal{N}$ with corresponding scalars c and ε as above, and for any two continuous functions $\widehat{\Psi}$ and $\widetilde{\Psi}$ satisfying (1),

$$\| x - x' \| \leq c \| \widehat{e}(x) - \widetilde{e}(x') \|,$$

for all $x \in \text{SOL}(K, \widehat{\Psi}) \cap \mathcal{N}$ and $x' \in \text{SOL}(K, \widetilde{\Psi}) \cap \mathcal{N}$, where $\widehat{e}(x) \triangleq \Phi(x) - \widehat{\Psi}(x)$ and $\widetilde{e}(x') \triangleq \Phi(x') - \widetilde{\Psi}(x')$.

Remarks.

• stability involves 3 requirements:

— isolatedness of the unperturbed solution

— local solvability of nearby VIs near the base solution

— pointwise proximity of such perturbed solutions to the base solution;

• strong stability involves in addition:

— locally unique solvability of nearby VIs

— Lipschitz property of such perturbed solutions;

• regularity and strong regularity can be similarly defined by restricting the perturbation to be: $\Psi(x) \triangleq \Phi(x) + e$, where $\|e\|$ is sufficiently small;

$$
\begin{array}{ccc}
\text{strong stability} & \Rightarrow & \text{stability} \\
\Updownarrow & & \Downarrow \\
\text{strong regularity} & \Rightarrow & \text{regularity}
\end{array}
$$

• if Φ is pseudo-monotone on K, then regularity = uniqueness + stability.

Theorem. Let K be closed and convex, and Φ be continuously differentiable near a solution $z^* \in \mathrm{SOL}(K, \Phi)$. Suppose that $z^* \in \mathrm{SOL}(K, \Phi^*)$ is stable, where

$$\Phi^*(z) \triangleq \Phi(z^*) + J\Phi(z^*)(\, z - z^* \,).$$

An open neighborhood $\mathcal{N}$ of z^* exists such that

(a) for all $z^0 \in \mathcal{N} \cap K$, a sequence $\{z^{\nu+1}\}$, where each $z^{\nu+1} \in \mathcal{N} \cap \mathrm{SOL}(K, \Phi^\nu)$, exists and converges Q-superlinearly to z^*; i.e,.

$$\lim_{\nu \to \infty} \frac{\| \, z^{\nu+1} - z^* \, \|}{\| \, z^\nu - z^* \, \|} = 0;$$

(b) if in addition $J\Phi$ is Lipschitz continuous near z^*, then the convergence is Q-quadratic; i.e.,

$$\limsup_{\nu \to \infty} \frac{\| \, z^{\nu+1} - z^* \, \|}{\| \, z^\nu - z^* \, \|^2} < \infty.$$

Conditions for (Strong) Stability

The critical cone of the VI at $x^* \in \text{SOL}(K, \Phi)$:

$$\mathcal{C}(K, \Phi; x^*) \triangleq \mathcal{T}(K; x^*) \cap \Phi(x^*)^\perp,$$

where $\mathcal{T}(K; x^*)$ is the tangent cone of K at $x^* \in K$; i.e.,

$$\mathcal{T}(K; x^*) \triangleq \left\{ v = \lim_{k\to\infty} \frac{x^k - x^*}{\tau_k} \mid \begin{array}{l} \{x^k\} \subset K \text{ and } \{\tau_k\} \subset \mathbb{R}_{++} \\ \lim_{k\to\infty} x^k = x^*, \quad \lim_{k\to\infty} \tau_k = 0 \end{array} \right\}.$$

If $K \triangleq \prod_{\alpha=1}^{Q} K^\alpha$, with $K^\alpha \subseteq \mathbb{R}^{n_\alpha}$, then with $x = (x^\alpha)_{\alpha=1}^Q$ and $\Phi(x) = (\Phi^\alpha(x))_{\alpha=1}^Q$,

$$\mathcal{C}(K, \Phi; x^*) = \prod_{\alpha=1}^{Q} \left[\mathcal{T}(K^\alpha; x^{*,\alpha}) \cap \Phi^\alpha(x^*)^\perp \right].$$

Theorem. Let each K^α be polyhedral. If $J\Phi(x^*)$ is semi-copositive on $\mathcal{C}(K, \Phi; z^*)$, then z^* is a stable solution of the (linearly constrained) VI (K, Φ) if and only if $\text{SOL}(\mathcal{C}(K, \Phi; z^*), 0, J\Phi(z^*)) = \{0\}$.

Conditions for (Strong) Stability : The finitely representable case

Let $K \triangleq \{x \mid g(x) \leq 0\}$, where $g : \mathbb{R}^n \to \mathbb{R}^m$ is twice continuously differentiable and each component function g_i is convex.

Under a constraint qualification, $z^* \in \text{SOL}(K, \Phi)$ if and only if λ^* exists

$$\left.\begin{array}{l} 0 = L(z^*, \lambda^*) \triangleq \Phi(z^*) + \sum_{i=1}^{m} \lambda_i^* \nabla g_i(z^*) \\ 0 \leq \lambda^* \perp -g(z^*) \geq 0 \end{array}\right\} \text{ the KKT conditions.}$$

Theorem. The pair (z^*, λ^*) is a strongly stable KKT pair if and only if

$$\left[\begin{array}{cc} J_z L(z^*, \lambda^*) & J g_\alpha(z^*)^T \\ -J g_\alpha(z^*) & 0 \end{array}\right]$$

is nonsingular, and the following Schur complement is a P-matrix:

$$\left[\begin{array}{cc} J g_\beta(z^*) & 0 \end{array}\right] \left[\begin{array}{cc} J_z L(z^*, \lambda^*) & J g_\alpha(z^*)^T \\ -J g_\alpha(z^*) & 0 \end{array}\right]^{-1} \left[\begin{array}{c} J g_\beta(z^*)^T \\ 0 \end{array}\right],$$

where $\alpha \triangleq \{i \mid g_i(z^*) = 0 < \lambda^*\}$ and $\alpha \triangleq \{i \mid g_i(z^*) = 0 = \lambda^*\}$.

Globalization for monotone problems

The gap function: for a compact set K,

$$\theta_{\text{gap}}(x) \triangleq \max_{y \in K} \Phi(x)^T (x - y), \quad x \in K.$$

For all $x \in K$,

(a) $\theta_{\text{gap}}(x) \geq 0$;

(b) $\theta_{\text{gap}}(x) = 0$ if and only if $x \in \text{SOL}(K, \Phi)$;

(c) θ_{gap} is directionally differentiable with the directional derivative given by

$$\theta'_{\text{gap}}(x; d) = F(x)^T d + \max \{ (x - y)^T JF(x) d \mid y \in \text{argmax}(\theta_{\text{gap}}(x)) \};$$

(d) for $x \notin \text{SOL}(K, \Phi)$, if $J\Phi(x)$ is copositive on $\mathcal{T}(K; x)$, then $\theta'_{\text{gap}}(x; dx) < 0$, where $dx \triangleq \bar{x} - x$ with $\bar{x} \in \text{SOL}(K, \Phi^x)$; i.e., the Newton direction is a descent direction of θ_{gap} at x;
motivating a descent method to globalize the JN method;

(e) if $J\Phi(x)$ is copositive on $\mathcal{T}(K; x)$, then (i) $x \in \text{SOL}(K, \Phi)$ if and only if $\text{SOL}(K, \Phi^x) \neq \emptyset$ and (ii) x is a constrained stationary point of θ_{gap} on K.

Semismooth Newton Methods for the NCP

The Nonlinear Complementarity: $0 \leq x \perp F(x) \geq 0$.

0. While locally fast convergent, the JN method suffers from the computational burden of solving an LCP at each iteration.

1. It is preferable to use nonsmooth equation reformulations that are (strongly) semismooth and whose associated merit functions are smooth.

2. It is desirable to solve a system of linear equation at each iteration.

3. In addition to being globally convergent, the resulting method should generate iterates all of whose limit points are solutions of the complementarity problem.

4. The merit functions should have bounded level sets so that at least one limit point of the iterates exists.

5. The methods should achieve a Q-superlinear convergence rate under mild conditions.

What is semismoothness and Why?

Definition. A locally Lipschitz continuous and directionally differentiable $G : \Omega \subseteq \Re^n \to \Re^m$, with Ω open, is semismooth at $\bar{x} \in \Omega$ if

$$\lim_{\substack{\bar{x} \neq x \to \bar{x} \\ H \in \partial G(x)}} \frac{G'(\bar{x}; x - \bar{x}) - H(x - \bar{x})}{\| x - \bar{x} \|} = 0,$$

where $\partial G(x)$ = the Clarke generalized Jacobian of G at x; equivalently,

$$\lim_{\substack{\bar{x} \neq x \to \bar{x} \\ H \in \partial G(x)}} \frac{\| G'(x; x - \bar{x}) - G'(\bar{x}; x - \bar{x}) \|}{\| x - \bar{x} \|} = 0;$$

G is strongly semismooth if

$$\limsup_{\bar{x} \neq x \to \bar{x}} \frac{\| G'(x; x - \bar{x}) - G'(\bar{x}; x - \bar{x}) \|}{\| x - \bar{x} \|^2} < \infty.$$

- Semismooth functions are nonsmooth functions that will yield the local fast convergence of Newton methods.

- Implicit-function theorems can be proved for parametric semismooth equations, with broad applications.

• Convex, thus norm, functions are semismooth.

• Piecewise differentiable functions are semismooth.

• (Strongly) semismooth functions are closed under addition, multiplication, and composition.

• Differentiable functions with Lipschitz gradients are strongly semismooth.

• For every $p \in [1, \infty]$, $\| \bullet \|_p$ is strongly semismooth.

• Piecewise affine functions are strongly semismooth, such as $(a, b) \mapsto \min(a, b)$.

• For every $\mu \in (0, 4)$, the function of 2 arguments:

$$(a, b) \mapsto \sqrt{(a-b)^2 + \mu ab} - (a + b)$$

is strongly semismooth.

• So is

$$\Phi(r, x_1, x_2) \equiv \min\left(\frac{r_+}{\sqrt{x_1^2 + x_2^2}}, 1 \right) x, \quad (r, x_1, x_2) \in \mathbb{R}^3.$$

A Semismooth Newton Method for Solving $G(x) = 0$

Given x^ν such that $G(x^\nu) \neq 0$, select an arbitrary element $H^\nu \in \partial G(x^\nu)$. Solve the system of linear equations $$G(x^\nu) + H^\nu d^\nu = 0$$ for d^ν; set $x^{\nu+1} \triangleq x^\nu + d^\nu$.

Local convergence. Let $G : \Omega \subseteq \Re^n \to \Re^n$, with Ω open, be semismooth at x^* in Ω satisfying $G(x^*) = 0$. If $\partial G(x^*)$ is nonsingular, then $\exists$ an open neighborhood $\mathcal{N}$ of x^* such that for all $x^0 \in \mathcal{N}$, the sequence $\{x^\nu\}$ is well defined and converges superlinearly to x^*. Furthermore, if G is strongly semismooth at x^*, then the convergence rate is Q-quadratic.

Globalization via a Smooth Merit Function

Suppose $\theta(x) \triangleq \frac{1}{2}G(x)^T G(x)$ is C^1. For any $H \in \partial G(x)$,

$$\nabla\theta(x) = H^T G(x).$$

Thus if the Newton direction $d \triangleq H^{-1}G(x)$ is well defined, it is a descent direction for θ at x.

> Given are x^ν such that $G(x^\nu) \neq 0$ and scalars $\rho > 0$, $p > 1$, and $\gamma \in (0,1)$. Select H^ν in $\partial G(x^\nu)$. Solve the system of linear equations
>
> $$G(x^\nu) + H^\nu d^\nu = 0$$
>
> for d^ν. If this system is not solvable or if the condition
>
> $$\nabla\theta(x^\nu)^T d^\nu \leq -\rho \, \| \, d^\nu \, \|^p$$
>
> is not satisfied, (re)set $d^\nu \triangleq -\nabla\theta(x^\nu)$ (steepest descent direction).
>
> Find the smallest nonnegative integer i_ν such that, with $i = i_\nu$,
>
> $$\theta(x^\nu + 2^{-i}d^\nu) \leq \theta(x^\nu) + \gamma \, 2^{-i} \, \nabla\theta(x^\nu)^T d^\nu;$$
>
> set $\tau_\nu \triangleq 2^{-i_\nu}$ and $x^{\nu+1} \triangleq x^\nu + \tau_\nu d^\nu$.

Application to the NCP

A C(omplementarity)-function is a function $\psi(a,b)$ of 2 arguments such that $\psi(a,b) = 0$ if and only if $0 \leq a \perp b \geq 0$.

- $\psi_{\mathsf{min}}(a,b) \triangleq \min(a,b)$ (the min function)
- $\psi_{\mathsf{FB}}(a,b) \triangleq \sqrt{a^2+b^2} - (\,a+b\,)$ (the F(ischer)-B(urmeister) function)
- $\psi(a,b) \triangleq \psi_{\mathsf{FB}}(a,b) - \tau \max(0,a)\max(0,b)$.

Key formulation

$$[\,0 \leq x \perp F(x) \geq 0\,] \Leftrightarrow \mathbf{F}_{\mathsf{FB}}(x) = 0,$$

where $\mathbf{F}_{\mathsf{FB}}(x) \triangleq (\,\psi_{\mathsf{FB}}(x_i, F_i(x))\,)_{i=1}^N$. Let

$$\theta_{\mathsf{FB}}(x) \triangleq \tfrac{1}{2}\mathbf{F}_{\mathsf{FB}}(x)^T\mathbf{F}_{\mathsf{FB}}(x)$$

be the merit function, which turns out to be a C^1-function.

Main Convergence

• (Existence of accumulation points) If $\{x \mid \theta_{\text{FB}}(x) \leq \theta_{\text{FB}}(x^0)\}$ is bounded, then $\{x^\nu\}$ is bounded; thus an accumulation point must exist.

• (Subsequential convergence) Every accumulation point of $\{x^\nu\}$ is a stationary point of θ_{FB}.

• (NCP solution) If x^* is an accumulation point of $\{x^\nu\}$ such that x^* is FB-regular, then x^* is a solution of the NCP (F).

• (Sequential convergence) If $\{x^\nu\}$ has an isolated limit point, then the whole sequence $\{x^\nu\}$ converges to that point.

• (Convergence rates). Suppose that x^* is a limit point of $\{x^\nu\}$ and a solution of NCP (F). Assume that $\partial \mathbf{F}_{\text{FB}}(x^*)$ is nonsingular. If $p > 2$ and $\gamma < 1/2$, then,

— the whole sequence $\{x^\nu\}$ converges to x^*,

— eventually, d^ν is always the Newton direction, and a unit step size is accepted so that $x^{\nu+1} = x^\nu + d^\nu$

— the convergence rate is Q-superlinear,

— if the Jacobian $JF(x)$ is Lipschitz continuous in a neighborhood of x^*, the convergence rate is Q-quadratic.

Numerical Methods for Complementarity Problems, Variational Inequalities, and Extended Systems

Jong-Shi Pang

Department of Industrial and Enterprise Systems Engineering

University of Illinois at Urbana-Champaign

presented at

CIMPA-UNESCO-VIETNAM School
Variational Inequalities and Related problems
Friday May 14, 2010, 2:00–3:45 PM

Lecture III: Monotone Problems and Beyond

Contents of Presentation

- Fixed-point formulations

— properties of the Euclidean projector

— single and double projection

— operator splitting

- Projection methods

— basic version for strongly monotone problems

— co-coercive problems

— pseudo-monotone problems

— without Lipschitz continuity

- Regularization methods

— Tikhonov for monotone problems

— for partitioned problems of the P_0-type

— proximal-point

- Partitioned VIs

— matrix splitting methods for LCPs

— Jacobi and Gauss-Seidel methods for distributed computation

The Euclidean Projector Π_K

Given a closed convex subset K of $\mathbb{R}^n$,

$$\Pi_K(x) \triangleq \operatorname*{argmin}_{y \in K} \ \tfrac{1}{2}(\, y - x\,)^T(\, y - x\,), \quad x \in \mathbb{R}^n.$$

- **Well-definedness** — $\Pi_K(x)$ exists and is unique.
- **Characterization** — $(y - \Pi_K(x))^T(\Pi_K(x) - x) \geq 0$ for all $y \in K$.
- **Co-coercivity** — $(x' - x)^T(\Pi_K(x') - \Pi_K(x)) \geq \|\Pi_K(x') - \Pi_K(x)\|_2^2$
- **implying Lipschitz** — $\|\Pi_K(x') - \Pi_K(x)\|_2 \leq \|x' - x\|_2$ for all x, x' in K.
- **Differentiability** — The squared distance function $\rho_K(x) \triangleq \|x - \Pi_K(x)\|_2^2$ is differentiable with gradient $\nabla\rho_K(x) = \Pi_K(x) - x$.
- **Directionally diff** — When K is polyhedral, $\Pi_K'(x; d) = \Pi_{\mathcal{C}}(d)$, where $\mathcal{C}$ is the critical cone: $\mathcal{T}(K; \Pi_K(x)) \cap (\Pi_K(x) - x)^\perp$.

Fixed-Point Formulations

Consider the VI (K, Φ): $x \in K$ and $(x' - x)^T \Phi(x) \geq 0 \;\; \forall x' \in K$.

A fixed-point formulation: $x = \Pi_K(x - \tau\Phi(x))$ for any scalar $\tau > 0$.

A double-projection formulation: $x = \Pi_K(x - \tau\Phi_K(\Pi_K(x - \tau\Phi(x))))$ for any scalar $\tau > 0$.

For the NCP: $0 \leq x \perp F(x) \geq 0$, reduce to $0 = \min(x, F(x))$.

Projection as an operator splitting: $x = \Pi_K(x - \Phi(x))$ is equivalent to

$$x \in K \;\; \text{and} \;\; (y - x)^T [\Psi(x) + \Xi(x)] \geq 0, \;\; \forall y \in K,$$

where $\Psi(x) \triangleq x$ and $\Xi(x) \triangleq \Phi(x) - x$ so that $\Phi = \Psi + \Xi$, which splits Φ as the sum of two maps.

A general splitting method: Given $x^\nu \in K$, let $x^{\nu+1} \in \text{SOL}(K, \Xi(x^\nu) + \Psi)$.

Projection Methods

$$\boxed{\text{Given } x^\nu \in K, \text{ let } x^{\nu+1} \triangleq \Pi_K(x^\nu - \tau\Phi(x^\nu))}, \quad \tau > 0$$

Theorem. Let Φ be Lipschitz with constant $L > 0$ and strongly monotone with constant $\mu > 0$ so that for all x and y in K,

$$\| \Phi(x) - \Phi(y) \|_2 \leq L \| x - y \|_2$$

and

$$(x - y)^T (\Phi(x) - \Phi(y)) \geq \mu \| x - y \|_2^2 .$$

If $\boxed{\tau \in \left(0, \frac{2\mu}{L^2}\right)}$, then the projection sequence $\{x^\nu\}_{\nu=0}^\infty$ converges linearly to the unique solution of the VI (K, Φ). □

Example. For $\Phi(x) \triangleq q + Mx$ for a positive definite M with $\lambda_{\min}$ being the least eigenvalue of the symmetric part $\frac{1}{2}(M + M^T)$ of M, it suffices to choose $\tau \in \left(0, \frac{2\lambda_{\min}}{\|M\|_2^2}\right)$.

Drawbacks in general:

- applicable only to simple sets K
- requires knowledge of both constants L and μ
- strong monotonicity is essential.

Steps for improvement:

$$\begin{array}{ccccccc} \text{strong M.} + \text{Lip.} & \Rightarrow & \text{co-coercivity} & \Rightarrow & \text{pseudo-M.} + \text{Lip.} & \Rightarrow & \text{pseudo-M.} \\ & & \Downarrow & & & & \\ & & \text{M.} + \text{Lip.} & & & & \end{array}$$

Convergence under co-coercivity: $\exists\, c > 0$ such that for all x and y in K,

$$(x - y)^T (\Phi(x) - \Phi(y)) \geq c \| \Phi(x) - \Phi(y) \|_2^2.$$

If $\text{SOL}(K, \Phi) \neq \emptyset$ and $\tau < \dfrac{1}{2c}$, then the projection sequence converges to a solution of the VI. $\square$

An application. Consider the partitioned VI where

$$K \triangleq \left[\prod_{\alpha=1}^{Q} X^{\alpha} \right] \times \mathbb{R}^m_+ \quad \text{and} \quad \Phi(x, \pi) \triangleq \begin{pmatrix} F(x) + A^T \pi \\ b - Ax \end{pmatrix},$$

with F strongly monotone on $X \triangleq \prod_{\alpha=1}^{Q} X^{\alpha}$. Even so, Φ is not strongly monotone or co-coercive on K. For polyhedral X^{α},

$$\boxed{\text{VI } (K, \Phi) \overset{\text{KKT}}{=\!=\!=} \text{VI } (\widehat{K}, F) \overset{\text{eliminate } x}{=\!=\!=\!=\!=\!=} \text{NCP}(\widehat{F})}$$

where $\widehat{K} \triangleq \{x \in K \mid Ax \leq b\}$ is no longer a Cartesian product, and

$$\widehat{F}(\pi) \triangleq b - \left[A \circ \widehat{x} \circ A^T \right] \pi \quad \text{is provably co-coercive},$$

with $\widehat{x}(r) \triangleq$ unique solution of the VI $(X, r + F)$.

The upshot is that the basic projection method can not be applied to the original VI (K, F) but can be applied to the equivalent co-coercive NCP $(\widehat{F})$.

While applicable to the equivalent VI $(\widehat{K}, F)$, a projection method can not take advantage of the Cartesian structure of X.

The Extragraident Method

Given $x^\nu \in K$, let $x^{\nu+1/2} \triangleq \Pi_K(x^\nu - \tau\Phi(x^\nu))$ and $x^{\nu+1} \triangleq \Pi_K(x^\nu - \tau\Phi(x^{\nu+1/2}))$

We say that the function Φ is pseudo monotone on K with respect to $\text{SOL}(K,\Phi)$ if the latter set is nonempty and for every $x^* \in \text{SOL}(K,\Phi)$ it holds that

$$(x - x^*)^T\Phi(x) \geq 0, \quad \forall\, x \in K.$$

Theorem. Let Φ be Lipschitz with constant $L > 0$ and pseudo-monotone on K with respect to $\text{SOL}(K,\Phi)$. If $\tau < 1/L$, then the extragradient sequence $\{x^\nu\}$ converges to a solution of the VI. $\square$

A hyperplane projection method allows the removal of the Lipschitz continuity requirement.

Tikhonov Regularization

Let $\{\varepsilon_\nu\} \downarrow 0$. Let $x^\nu \in \text{SOL}(K, \Phi + \varepsilon_\nu I)$.

• If Φ is continuous and monotone on K, then $\text{SOL}(K, \Phi + \varepsilon I)$ is a singleton for every $\varepsilon > 0$.

• The same is true if $K \triangleq \prod_{\alpha=1}^{Q} K^\alpha$ and if Φ is a continuous P_0-function on K, i.e., for every pair $x \triangleq (x^\alpha)_{\alpha=1}^Q \neq y \triangleq (y^\alpha)_{\alpha=1}^Q$, both in K, an index α exists such that $x^\alpha \neq y^\alpha$ and

$$(x^\alpha - y^\alpha)^T (\Phi^\alpha(x) - \Phi^\alpha(y)) \geq 0,$$

• The asymptotic behavior of the Tikhonov trajectory $\{x(\varepsilon) \mid \varepsilon > 0\}$ as $\varepsilon \downarrow 0$ is of interest, where $x(\varepsilon)$ is the unique solution of the regularized VI $(K, \Phi + \varepsilon I)$.

• Continuity of the trajectory is easily shown.

• If Φ is continuous and monotone, then SOL(K, Φ) is a closed convex set, if nonempty.

• Thus, if Φ is continuous and monotone and if SOL$(K, \Phi) \neq \emptyset$, then SOL(K, Φ) has a unique element with the least Euclidean norm, which we call the least-norm solution of the VI (K, Φ).

Convergence of Tikhonov.

• If Φ is continuous and monotone on K, then TFAE:

(i) $\lim_{\varepsilon \downarrow 0} x(\varepsilon)$ exists; (ii) $\limsup_{\varepsilon \downarrow 0} \|x(\varepsilon)\| < \infty$; and (iii) SOL$(K, \Phi) \neq \emptyset$.

If any one of the above holds, then $\lim_{\varepsilon \downarrow 0} x(\varepsilon) =$ least-norm solution of the VI (K, Φ).

• If Φ is a continuous P_0-function on K, then

$$\emptyset \neq \text{SOL}(K, \Phi) \text{ bounded} \quad \Rightarrow \quad \limsup_{\varepsilon \downarrow 0} \|x(\varepsilon)\| < \infty$$

$$\Updownarrow \ \Phi \text{ analytic}$$

$$\lim_{\varepsilon \downarrow 0} x(\varepsilon) \text{ exists.}$$

The Proximal-Point Method

Given $x^\nu \in K$, let $x^{\nu+1} \in \text{SOL}(K, \Phi + \tau(\bullet - x^\nu))$, $\tau > 0$.

If Φ is continuous and monotone on K, then

- $\text{SOL}(K, \Phi) \neq \emptyset$ implies $\{x^\nu\}$ converges to a solution of the VI;
- $\text{SOL}(K, \Phi) = \emptyset$ implies $\lim_{\nu \to \infty} \|x^\nu\| = \infty$.

There is a local convergence theory without requiring monotonicity.

Partial regularization

Suppose F is strongly monotone. As an alternative to the previously discussed projection approach, we may partially regularize the map

$$\Phi(x, \pi) \triangleq \begin{pmatrix} F(x) + A^T \pi \\ b - Ax \end{pmatrix}$$

by considering, for $\tau > 0$,

$$\Phi^\tau(x, \pi) \triangleq \begin{pmatrix} F(x) + A^T \pi \\ b - Ax + \tau\,(\,\pi - \pi^\nu\,) \end{pmatrix}.$$

Matrix Splitting Methods for the LCP

Consider the LCP: $0 \leq z \perp q + Mz \geq 0$, and let $M \triangleq B + C$.

$$\boxed{\text{Given } z^\nu \geq 0, \text{ let } z^{\nu+1} \in \text{SOL}(q + Cz^\nu, B)} \; .$$

Some common choices of the splitting: $M = D + L + U$, where D, L, and U are the diagonal, strictly lower triangular, and strictly upper triangular parts of M, respectively. Suppose that M has positive diagonals.

- $B = \tau I$ yields a simple projection method
- $B = D$ yields the projected Jacobi method:

$$z_i^{\nu+1} = \max \left\{ 0, \, -m_{ii}^{-1} \left(q_i + \sum_{j \neq i} m_{ij} z_j^\nu \right) \right\}, \quad j = 1, \cdots, n.$$

- $B = D + L$ yields the projected Gauss-Seidel method:

$$z_i^{\nu+1} = \max \left\{ 0, \, -m_{ii}^{-1} \left(q_i + \sum_{j<i} m_{ij} z_j^{\nu+1} + \sum_{j>i} m_{ij} z_j^\nu \right) \right\}, \quad j = 1, \cdots, n.$$

- $B = L + \omega^{-1} D$ the projected successive overrelaxation method, $\omega \in (0, 2)$.

A Diagonal Dominance Based Convergence Theory

Given $E \in \mathbb{R}^{n \times n}$, define the comparison matrix $\bar{E} \triangleq [\bar{e}_{ij}]_{i,j=1}^{n}$, where

$$\bar{e}_{ij} \triangleq \begin{cases} e_{ii} & \text{if } i = j \\ -|\, e_{ij} \,| & \text{if } i \neq j \end{cases}.$$

Note that $\bar{E}$ has all off-diagonal entries non-positive.

A matrix E (with $e_{ii} > 0$ for all i) is an H-matrix if $\bar{E}$ is a P-matrix, i.e., all principal minors of $\bar{E}$ are positive; equivalently, if $\exists\ d \geq 0$ such that $\bar{E}d > 0$.

Theorem. Let B be an H-matrix with positive diagonals. If $\rho(\bar{B}^{-1}|C|) < 1$, then

- M is an H-matrix with positive diagonals, and
- the splitting method converges to the unique solution of the LCP (q, M).

□

Nonlinear Jacobi and Gauss-Seidel

Consider a partitioned VI with $K \triangleq \prod_{i=1}^{Q} K^i$. Write x and $F(x)$ accordingly:

$$x \triangleq \left(x^i \right)_{i=1}^{Q} \quad \text{and} \quad F(x) \triangleq \left(F^i(x) \right)_{i=1}^{Q}.$$

Assume for each i and fixed $x^{-i} \triangleq (x^j)_{j \neq i}$, the function $F^i(\bullet, x^{-i})$ is strongly monotone. The entire vector function F needs not be monotone, however.

Given $x^\nu \triangleq (x^{\nu,i})_{i=1}^{Q} \in K$.

Jacobi: Let $x^{\nu+1,i} \in \text{SOL}(K^i, F^i(\bullet, x^{\nu,-i}))$ for each $i = 1, \cdots Q$.

Gauss-Seidel: Let $x^{\nu+1,i} \in \text{SOL}(K^i, F^i(x^{\nu+1,1}, \cdots, x^{\nu+1,i-1}, \bullet, x^{\nu,i+1}, \cdots, x^{\nu,n}))$ for each $i = 1, \cdots, Q$.

While the GS method is sequential, the Jacobi method can be implemented in a distributed manner.

These methods have in recent days received significant interests in many game-theoretic applications, where multi-agent optimization is a central mathematical paradigm for analyzing the non-cooperative behavior of the agents.

Convergence Theory

Suppose that F has bounded derivatives on K. Let

$$\zeta^k_{\text{min}} \triangleq \inf_{x\in K} \text{smallest eigenvalue of } J_{x^k} F^k(x)$$

$$\xi^{kk'}_{\text{max}} \triangleq \sup_{x\in K} \left\| J_{x^{k'}} F^k(x) \right\|, \quad \text{for all } k' \neq k.$$

Define the $Q \times Q$ matrix

$$\mathbf{\Gamma} \triangleq \begin{bmatrix} \frac{1}{1+\zeta^1_{\text{min}}} & \frac{\xi^{12}_{\text{max}}}{1+\zeta^1_{\text{min}}} & \frac{\xi^{13}_{\text{max}}}{1+\zeta^1_{\text{min}}} & \cdots & \frac{\xi^{1Q}_{\text{max}}}{1+\zeta^1_{\text{min}}} \\ \frac{\xi^{21}_{\text{max}}}{1+\zeta^2_{\text{min}}} & \frac{1}{1+\zeta^2_{\text{min}}} & \frac{\xi^{23}_{\text{max}}}{1+\zeta^2_{\text{min}}} & \cdots & \frac{\xi^{2Q}_{\text{max}}}{1+\zeta^2_{\text{min}}} \\ \vdots & \vdots & \ddots & & \vdots \\ \frac{\xi^{Q-11}_{\text{max}}}{1+\zeta^{Q-1}_{\text{min}}} & \frac{\xi^{Q-12}_{\text{max}}}{1+\zeta^{Q-1}_{\text{min}}} & \cdots & \frac{1}{1+\zeta^{Q-1}_{\text{min}}} & \frac{\xi^{Q-1Q}_{\text{max}}}{1+\zeta^{Q-1}_{\text{min}}} \\ \frac{\xi^{Q1}_{\text{max}}}{1+\zeta^{Q}_{\text{min}}} & \frac{\xi^{Q1}_{\text{max}}}{1+\zeta^{Q}_{\text{min}}} & \cdots & \frac{\xi^{QQ-1}_{\text{max}}}{1+\zeta^{Q}_{\text{min}}} & \frac{1}{1+\zeta^{Q}_{\text{min}}} \end{bmatrix}.$$

- If $\rho(\mathbf{\Gamma}) < 1$, then iteration is a contraction.

- If $|||\mathbf{\Gamma}||| \leq 1$ for a matrix norm induced by some monotonic vector norm, then iteration is nonexpansive. in this case, an averaging scheme can compute a solution of the VI if one exists.

Numerical Methods for Complementarity Problems, Variational Inequalities, and Extended Systems

Jong-Shi Pang

Department of Industrial and Enterprise Systems Engineering

University of Illinois at Urbana-Champaign

presented at

CIMPA-UNESCO-VIETNAM School
Variational Inequalities and Related problems
Monday May 17, 2010, 2:00–3:45 PM

Lecture IV: Differential Variational Inequalities

A new paradigm of dynamical systems

Contents of Presentation

- **D**ifferential **V**ariational **I**nequalities: What and why?

— a general motivation

- Several perspectives:

— piecewise systems

— discontinuous systems

— optimal control

— complex systems in evolution

- Formal definitions
- A Time-stepping method illustrated

Why a new paradigm of dynamical system?

On one hand, ordinary differential equations (ODEs) with smooth right-hand sides have a long history in mathematics:

$\dot{x}(t)$	$\triangleq$	$\dfrac{dx(t)}{dt} = f(t, x(t))$	modeling systems in evolution
0	$=$	$\Phi(x(0), x(T))$	prescribed initial and terminal conditions,

where $f : \mathbb{R}^{1+n} \to \mathbb{R}^n$ and $\Phi : \mathbf{R}^{2n} \to \mathbf{R}^n$ are differentiable vector functions.

On the other hand, mathematical programming is a post-world war II subject, typically thought of as an optimization problem in finite dimensions:

$$\boxed{\text{minimize } \theta(x) \,:\, x \in X \subseteq \mathbb{R}^n} \quad ;$$

many of its subfields—e.g., computational game theory—are more recent areas of focus.

Mathematical programming (MP) in a nutshell

Strengths:

- treatment of algebraic inequalities (as opposed to equations)

- consequences thereof—piecewise properties, non-smoothness, multi-valuedness, disjunctions, and logical relations

- variational conditions for equilibrium problems

- computational emphasis

Major deficiency:

- static – discrete-time at best

What are Differential Variational Systems?

Bridging smooth ODEs with MP, DVIs form a novel mathematical paradigm offering a broad, unifying framework for modeling many applied (dis)equilibrium problems containing

- dynamics (pathway to equilibrium)
- inequalities (unilateral constraints), and
- disjunctions (paradigm switch).

These dynamical systems exhibit the following non-traditional characteristics:

- state-triggered mode transition, distinct from arbitrary switchings
- non-smoothness, extending smooth dynamical systems
- an endogenous auxiliary variable, satisfying a variational condition.

One one hand, these are hybrid systems of a particular kind; on the other hand, they are special cases of differential inclusions.

Source Problems

- piecewise smooth ODEs
- ODEs with discontinuous (or even set-valued) right-hand sides
- optimal control problems with control/state constraints
- differential Nash games
- multi-body mechanical systems under frictional contact
- dynamic traffic equilibria
- electric circuits with ideal diodes
- hybrid engineering systems
- biological synthesis modeling, and many more.

A piecewise systems perspective

Consider the linear ODE:

$$\boxed{\dot{x} = Ax, \quad x(0) = x^0}\ , \quad \text{where } A \in \mathbb{R}^{n \times n}.$$

The unique solution: $\boxed{x(t) = \exp(At)x^0}$, for all $t \geq 0$.

Now consider a simple variant:

$$\dot{x} = \begin{cases} Ax & \text{if } c^T x < 0 \\ ? & \text{if } c^T x = 0 \\ Bx & \text{if } c^T x > 0 \end{cases}$$

Two cases:

- $Ax = Bx$ on $c^T x = 0$, or
- Ax not necessarily equal to Bx on $c^T x = 0$.

The continuous case:

$$\dot{x} = \begin{cases} Ax & \text{if } c^T x \leq 0 \\ Bx & \text{if } c^T x \geq 0 \end{cases}$$

where $Ax = Bx$ on $c^T x = 0$, implying $B = A + bc^T$.

Equivalently, $\boxed{\dot{x} = Ax + b \max(0, c^T x)}$.

- The right-hand side is piecewise linear, but not differentiable.
- A solution $x(t, x^0)$ exists and is unique for all $t \geq 0$ and all x^0.
- There is no longer an explicit expression for $x(t, x^0)$.

Some questions:

- How often does the trajectory $x(\bullet, x^0)$ switch between the two halfspaces? In finite time? In infinite time?

- What does "switch" mean formally? Is touching the hyperplane considered a switch?

- The Zeno behavior: infinitely many mode switches in finite time.

- Are there bimodal systems with (in)finite switches in long time?

- Can we characterize bimodal systems with finite switches, including zero, in infinite time?

- Does a bimodal system inherit the same system-theoretic properties as the two modes?

- In particular, is the Lyapunov stability of a bimodal system equivalent to that of its two modes?

- How is such a dynamical system related to mathematical programming?

Relation to Mathematical Programming

• via the piecewise linearity, and more generally, non-differentiability;

• via complementarity, and more generally, disjunction:
let $\tau \equiv \max(0, c^T x)$,

$$\begin{array}{rcll} \dot{x} & = & Ax + b\tau & \text{an ODE} \\ 0 \leq \tau & \perp & -c^T x + \tau \geq 0 & \text{a (trivial) LCP} \end{array}$$

where $\perp$ is the orthogonality condition, which in the scalar case is equivalent to the **disjunctive "or"** condition; i.e., either $\tau = 0$ or $-c^T x + \tau = 0$.

• Above is an ODE coupled with a variational condition;

• a (trivial) state-parameterized 1-dimensional linear complementarity problem defined by the scalar 1;

• thereby merging the classical ODE domain with the contemporary optimization/equilibrium field.

The Filippov multi-valued convexification

$$\dot{x} \in F(x) \triangleq \begin{cases} \{Ax\} & \text{if } c^T x > 0 \\ \{\lambda Ax + (1-\lambda)Bx \mid \lambda \in [0,1]\} & \text{if } c^T x = 0 \\ \{Bx\} & \text{if } c^T x < 0 \end{cases}$$

Conversion to a linear complementarity system:

Write $c^T x = \eta^+ - \eta^-$ with $0 \leq \eta^+ \perp \eta^- \geq 0$; get

$$\begin{array}{rcl} \dot{x} & = & \lambda Ax + (1-\lambda)Bx \\ 0 \leq \begin{pmatrix} \lambda \\ \eta^+ \end{pmatrix} & \perp & \begin{pmatrix} -c^T x \\ 1 \end{pmatrix} + \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{pmatrix} \lambda \\ \eta^+ \end{pmatrix} \geq 0; \end{array}$$

• The variational condition is a state-parameterized 2-dimensional linear complementarity problem defined by the positive semidefinite (albeit asymmetric) matrix $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$.

• The (discontinuous) step-function and the signum function can similarly be modeled by the complementarity condition.

A linear-quadratic optimal control problem

extendable to an open-loop differential linear-quadratic game

Given a time horizon $T > 0$ and an initial state $\xi \in \mathbb{R}^n$, find an absolutely continuous function $x : [0,T] \to \mathbb{R}^n$ and an integrable function $u : [0,T] \to \mathbb{R}^m$:

$$\underset{x,u}{\text{minimize}} \quad V(x,u) \equiv c^T x(T) + \tfrac{1}{2} x(T)^T W x(T) +$$

$$\int_0^T \left[\begin{pmatrix} x(t) \\ u(t) \end{pmatrix}^T \begin{pmatrix} p(t) \\ q(t) \end{pmatrix} + \tfrac{1}{2} \begin{pmatrix} x(t) \\ u(t) \end{pmatrix}^T \begin{bmatrix} P & Q \\ Q^T & S \end{bmatrix} \begin{pmatrix} x(t) \\ u(t) \end{pmatrix} \right] dt$$

subject to $x(0) = \xi$

and for almost all $t \in [0,T]$:

$$\dot{x}(t) = r(t) + Ax(t) + Bu(t) \quad \text{and} \quad f + Cx(t) + Du(t) \geq 0,$$

where W and $\begin{bmatrix} P & Q \\ Q^T & S \end{bmatrix}$ are symmetric positive semidefinite.

Necessary and Sufficient Optimality Conditions

A pair of functions (x, u) is optimal if and only if $\exists$ an absolutely continuous function $\lambda(t)$ such that for all almost all $t \in [0, T]$:

$$\begin{pmatrix} \dot{\lambda}(t) \\ \dot{x}(t) \end{pmatrix} = \begin{pmatrix} -p(t) \\ r(t) \end{pmatrix} + \begin{bmatrix} -A^T & -P \\ 0 & A \end{bmatrix} \begin{pmatrix} \lambda(t) \\ x(t) \end{pmatrix} + \begin{bmatrix} -Q \\ B \end{bmatrix} u(t)$$

$$\boxed{\begin{array}{lll} u(t) \in & \underset{u}{\text{argmin}} & u^T \left[q(t) + B^T \lambda(t) + Q^T x(t) \right] + \frac{1}{2} u^T S u \\ & \text{subject to} & f + Cx(t) + Du \geq 0 \end{array}}$$

$$x(0) = \xi \quad \text{and} \quad \lambda(T) = c + Wx(T).$$

- Above is a differential affine variational system with boundary conditions;
- in this case, the variational condition is a convex quadratic program in the control variable parameterized by the state and adjoint variable;
- nonlinear generalizations.

Complex systems in evolution

illustrated by a dynamic traffic equilibrium problem

- A traffic network represented as a directed graph with node set $\mathcal{N}$, link set $\mathcal{L}$, trip destination set S.

- For a given link $(i, j) \in \mathcal{L}$ from node i to node j, denote by f_{ij}^s the flow on link (i, j) to destination $s \in \mathcal{N}$.

- The travel time along a link $(i, j) \in \mathcal{L}$ is a given function $\tau_{ij}(f)$ of the total flow vector f with components f_ℓ^s for all $\ell \in \mathcal{L}$ and all $s \in S$.

- For each $i \in \mathcal{N}$ and $s \in S$, a function $d_i^s(\pi)$ is given that represents the travel demand from i to s where π is the (unknown) vector of minimum travel times $\pi_j^{s'}$ from node $j \in \mathcal{N}$ to destination $s' \in S$.

"Dynamicizing" the static user equilibrium

A static user equilibrium is a pair (f,π) of vectors satisfying

- Wardrop's route-choice principle (formulated using link flows):

$$0 \le f_{ij}^s \perp \tau_{ij}(f) + \pi_j^s - \pi_i^s \ge 0, \quad \text{for all } (i,j) \in \mathcal{L} \text{ and } s \in S$$

$$\Leftrightarrow \quad 0 = \min\left(f_{ij}^s,\ \tau_{ij}(f) + \pi_j^s - \pi_i^s \right)$$

- demand satisfaction:

$$0 \le \pi_i^s \perp \sum_{j:(i,j)\in\mathcal{L}} f_{ij}^s - \sum_{k:(k,i)\in\mathcal{L}} f_{ki}^s - d_i^s(\pi) \ge 0, \quad \text{for all } i \in \mathcal{N} \text{ and } s \in S,$$

A continuous-time flow-adjusting user (dis)equilibrium:

$$\frac{df_{ij}^s(t)}{dt} = -\min\left(f_{ij}^s(t),\ \tau_{ij}(f(t)) + \pi_j^s(t) - \pi_i^s(t) \right),$$

$$0 \le \pi_i^s \perp \sum_{j:(i,j)\in\mathcal{L}} f_{ij}^s - \sum_{k:(k,i)\in\mathcal{L}} f_{ki}^s - d_i^s(\pi) \ge 0.$$

- Above is a differential complementarity system with a piecewise ODE and a nonlinear complementarity condition.

Frictional contact in engineering mechanics

Multiple bodies come into contact under external forces (gravitation e.g.); goal is to understand body motion (possibly deformations) and interaction of forces, some internal.

Friction at contact points induces rolling and slipping, leading to system mode changes, among other things. Also contacts can break.

Initiated by Löstedt (1981), rigid-body paradigm have been extensively examined; yet they induce discontinuity in velocities and impulsive forces, thus are challenging to analyze rigorously and simulate numerically.

Mathematically, rigid-body models require the theory and methods of measure differential inclusions.

Locally compliant models were proposed in Song's 2002 Ph.D. thesis as an alternative to ease some deficiencies associated with the rigid-body models, without sacrificing the physics of contact.

Friction model with compliance:
5 main components

- Newton's law to describe force-induced body dynamics
- Kinematics to specify body orientation due to rotational motion
- Constitutive law of compliance
- Principle of non-penetration
- Friction law

The model equations

- Force equilibrium: $M(q)\dot{\nu} = f(q,\nu) + W(q)\lambda$, where $M(q)$ is the mass matrix, q and ν are, resp. the (generalized) position and velocity vector, $\lambda \triangleq (\lambda_{\mathrm{n}}, \lambda_{\mathrm{t}}\lambda_{\mathrm{o}})$ is the tuple of normal contact and two tangential friction forces;

- Kinematics: $\dot{q} = G(q)\nu$;

- Definition of separation: $s \equiv \delta + \Psi(q)$;

- Constitutive law of local compliance: $\lambda = K(q)\delta + C(q)\dot{\delta}$, where $K(q)$ and $C(q)$ are, resp., the stiffness and compliance matrix;

- Normal non-penetration and contact: $0 \leq \lambda_{\mathrm{n}} \perp s_{\mathrm{n}} \geq 0$;

- Tangential friction principle: for each contact point i,

$(\lambda_{i\mathrm{t}}, \lambda_{i\mathrm{o}}) \in \underset{(\widehat{\lambda}_{i\mathrm{t}}, \widehat{\lambda}_{i\mathrm{o}}) \in \mathcal{F}_i(\mu_i \lambda_{i\mathrm{n}})}{\arg\min} \left\{ \frac{ds_{i\mathrm{t}}}{dt} \widehat{\lambda}_{i\mathrm{t}} + \frac{ds_{i\mathrm{o}}}{dt} \widehat{\lambda}_{i\mathrm{o}} \right\}$, where $\mathcal{F}_i$ is the friction map such as the Coulomb friction: $\mathcal{F}_i(\rho) \equiv \{(a,b) \in \Re^2 : \sqrt{a^2 + b^2} \leq \rho\}$, $\rho \geq 0$.

Differential Complementarity Systems (DCSs)

Linear. Given $A \in \Re^{n \times n}$, $B \in \Re^{n \times m}$, $C \in \Re^{m \times m}$, and $D \in \Re^{m \times m}$,

$$\boxed{\begin{array}{rcl} \dot{x} & = & Ax + By \\ 0 \leq y & \perp & Cx + Dy \geq 0 \end{array}} \quad \Leftrightarrow \quad \boxed{\begin{array}{l} \text{differential inclusion} \\ \dot{x} \in Ax + B\,\text{SOL}(Cx, D) \end{array}}$$

Nonlinear. Given $(F, G) : \Re^{n+m} \to \Re^{n+m}$,

$$\boxed{\begin{array}{rcl} \dot{x} & = & F(x, y) \\ 0 \leq y & \perp & G(x, y) \geq 0 \end{array}}$$

Some issues

- Well-posedness (i.e., existence + uniqueness of a "solution")
- Regularity (C^1, absolutely continuous, L^2); dependence on initial conditions
- Zenoness (i.e., infinite mode switches in finite time)
- System-theoretic issues (e.g., observability, stability, controllability)
- Numerics (time-stepping methods and their convergence)
- Boundary-value problems $\Phi(x(0), x(T)) = 0$
- Extensions and applications.

Differential Quasi-Variational Inequalities

model for frictional contact problems

$\dot{x} = F(x, y, u)$	an ODE (dynamics of motion)
$0 \leq y \perp G(x, y) \geq 0$	complementarity (non-penetration)
$u \in \text{SOL}(K(x, y), \Psi(x, y, \bullet))$	variational inequality (friction forces)

where $\text{SOL}(K(x, y), \Psi(x, y, \bullet))$ is the solution set of a variational inequality; i.e.,

$$u \in K(x, y), \quad \text{and} \quad (u' - u)^T \Psi(x, y, u) \geq 0, \quad \forall u' \in K(x, y),$$

where $K(x, y)$ is a moving closed convex set

$$K(x, y) \equiv \{ u \in \Re^\ell : \widehat{H}(x, y, u) \geq 0 \},$$

parameterized by the state variable x and the algebraic variable y that satisfies a state-dependent NCP.

A time-stepping method illustrated

The initial-value linear complementarity system (LCS)

$$\begin{aligned} \dot{x} \quad &= \quad Ax + By, \quad x(0) = \xi \\ 0 \leq y \quad &\perp \quad Cx + Dy \geq 0. \end{aligned}$$

Let $\theta \in [0,1]$; $h > 0$, $N_h \triangleq T/h$, and $x^{h,0} \triangleq \xi$. Generate a sequence of iterates

$$\left\{ x^{h,1}, \cdots, x^{h,N_h} \right\} \quad \text{and} \quad \left\{ y^{h,1}, \cdots, y^{h,N_h} \right\}$$

by discretizing the time derivative and solving a sequence of finite-dimensional linear complementarity problems: iteratively, for $k = 0, \cdots, N_h - 1$,

$$\begin{aligned} x^{h,k+1} \quad &= \quad x^{h,k} + h \left\{ A \left[\theta\, x^{h,k} + (1-\theta)\, x^{h,k+1} \right] + By^{h,k+1} \right\} \\ 0 \leq y^{h,k+1} \quad &\perp \quad C \left[\theta\, x^{h,k} + (1-\theta)\, x^{h,k+1} \right] + Dy^{h,k+1} \geq 0; \end{aligned}$$

$\theta = 0 \Rightarrow$ an implicit scheme

$\theta = 1 \Rightarrow$ an explicit scheme.

Construct by linear/constant interpolation discrete-time trajectories $\widehat{x}^h(t)$ and $\widehat{y}^h(t)$ for $t \in (0,T]$; specifically,

A convergence theorem

$$\begin{aligned} \widehat{x}^h(t) &\equiv x^{h,k} + \frac{t - t_{h,k}}{h}(x^{h,k+1} - x^{h,k}) \quad \forall\, t \in [\, t_{h,k}, t_{h,k+1}\,], \\ \widehat{y}^h(t) &\equiv y^{h,k+1} \quad \forall\, t \in (t_{h,k}, t_{h,k+1}], \end{aligned}$$

where $t_{h,k} \triangleq kh$ for all $k = 0, 1, \cdots, N_h$.

Theorem. Let D be positive semidefinite. Suppose $\{x^{h,k}, y^{h,k}\}_{k=1}^{N_h}$ are well defined for all $h > 0$ sufficiently small and positive scalars ρ_y, ψ_x, and $\bar{h}$ exist such that for all $h \in (0, \bar{h}]$ and all $k = 0, 1, \cdots, N_h - 1$,

$$\|\, y^{h,k+1} \,\| \leq \rho_y \,(\, 1 + \|\, x^{h,k} \,\|\,) \quad \text{and} \quad \|\, x^{h,k+1} - x^{h,k} \,\| \leq h\, \psi_x \,(\, 1 + \|\, x^{h,i} \,\|\,).$$

There is a sequence $\{h_\nu\} \downarrow 0$ such that the following two limits exist: $\widehat{x}^{h_\nu} \to \widehat{x}$ uniformly on $[0, T]$ and $\widehat{y}^{h_\nu} \to \widehat{y}$ weakly in $L^2(0, T)$. Furthermore, all such limits $(\widehat{x}, \widehat{y})$ are weak solutions of the initial-value homogeneous LCS with $x(0) = x^0$. □