Lecture 22

# Algorithms for Monotone VIs

# Regularization and Proximal-Point Methods

November 24, 2008

# Outline

- Tikhonov Regularization
- Proximal-Point Methods

# Tikhonov Regularization

- The basic idea is to approximate the given $VI$ by a sequence of $VI$'s with "better properties" than the original $VI$.

- Here, the mapping $F$ is approximated with a family of mappings $F_\epsilon$ such that $F_\epsilon \longrightarrow F$ as $\epsilon \longrightarrow 0$, where the convergence of the mappings is pointwise

- Regularization corresponds to using $F_\epsilon = F + \epsilon I$

- We study the regularization methods for the class of monotone $VI$'s

- Throughout this lecture (unless stated otherwise), we assume that $K$ **is nonempty closed convex set**

# Tikhonov Trajectory

- Let $F : K \to \mathbb{R}^n$ be continuous and monotone. For $\epsilon > 0$, let $x_\epsilon$ be a solution to $VI(K, F_\epsilon)$. (Exists and unique - why?)
- The set of points $\{x_\epsilon \mid \epsilon > 0\}$ is called the **Tikhonov trajectory of the** $VI(K, F)$.

The Tikhonov trajectory is bounded when $\epsilon$ is bounded away from zero and $F$ is continuous and monotone.

**Proposition 1 (Prop 12.2.1)** *Let $F$ be a continuous monotone mapping on $K$, and let $\epsilon \geq \bar{\epsilon}$. Then, $\{x_\epsilon\}$ is bounded.*

**Proof**: Since $x_\epsilon \in SOL(K, F_\epsilon)$, we have $(\epsilon x_\epsilon + F(x_\epsilon))^T(x - x_\epsilon) \geq 0$ for all $x \in K$. Let $\bar{x} \in K$ be fixed. Then, the preceding relation yields

$$F(x_\epsilon)^T(\bar{x} - x_\epsilon) \geq \epsilon \|x_\epsilon - \bar{x}\|^2 + \bar{x}^T(x_\epsilon - \bar{x}).$$

By monotonicity of $F$, we have $F(\bar{x})^T(\bar{x}-x_\epsilon) \geq F(x_\epsilon)^T(\bar{x}-x_\epsilon)$, implying that

$$F(\bar{x})^T(\bar{x}-x_\epsilon) \geq \epsilon\|x_\epsilon-\bar{x}\|^2 + \bar{x}^T(x_\epsilon-\bar{x}).$$

Therefore,

$$\epsilon\|x_\epsilon-\bar{x}\|^2 \leq \|\bar{x}\| \cdot \|x_\epsilon-\bar{x}\| + \|F(\bar{x})\| \cdot \|\bar{x}-x_\epsilon\|,$$

implying

$$\|x_\epsilon-\bar{x}\| \leq \frac{1}{\epsilon}(\|\bar{x}\| + \|F(\bar{x})\|) \leq \frac{1}{\bar{\epsilon}}(\|\bar{x}\| + \|F(\bar{x})\|),$$

showing that $\{x_\epsilon\}$ is bounded.

# Tikhonov Trajectory Properties

**When $\epsilon \downarrow 0$, the sequence $\{x_\epsilon\}$ may not converge.**
Consider, for example, the $LCP(q, M)$ where

$$M = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \qquad q = \begin{bmatrix} -1 \\ 0 \end{bmatrix}.$$

The solution set is

$$SOL(q, M) = \{(x_1, 1) \mid x_1 \geq 0\} \bigcup \{(0, x_2) \mid x_2 \geq 1\}.$$

The regularized $LCP(q, M + \epsilon I)$ has a unique solution $x_\epsilon = (1/\epsilon, 0)$ for every $\epsilon > 0$, which is unbounded as $\epsilon \to 0$. Interestingly, the distance from $x_\epsilon$ to $SOL(q, M)$ is equal to 1 for all $\epsilon > 0$. Note that in this case $F(x) = Mx + q$ is not monotone; also note the solution set $SOL(q, M)$ is nonconvex.

But for continuous monotone $F$, the situation is better.

**Theorem 2** *Let $F$ be continuous monotone on $K$. The followings statements are equivalent for Tikhonov trajectory:*

*(a) The limit $lim_{\epsilon \downarrow 0} x_\epsilon$ exists.*

*(b) The sequence $\{x_\epsilon\}$ is bounded, i.e.,* $\limsup_{\epsilon \downarrow 0} x_\epsilon < \infty$.

*(c) $SOL(K, F)$ is nonempty.*

**Proof**: (a) $\implies$ (b) is evident.
(b) $\implies$ (c). Let $\bar{x}$ be a limit point of some subsequence $\{x_{\epsilon_k}\}$ of the Tikhonov trajectory as $\epsilon_k \to 0$. Since $\left(F(x_{\epsilon_k}) + \epsilon_k x_{\epsilon_k}\right)^T (x - x_{\epsilon_k}) \geq 0$ for all $k$ and every $x \in K$, y continuity of $F$, we have for any $x \in K$,

$$\lim_{k \to \infty} \left(F(x_{\epsilon_k}) + \epsilon_k x_{\epsilon_k}\right)^T (x - x_{\epsilon_k}) = F(\bar{x})^T (x - \bar{x}).$$

Hence, $F(\bar{x})^T (x - \bar{x}) \geq 0$ showing that $\bar{x} \in SOL(K, F)$.

(c) $\Longrightarrow$ (a). Since $x_\epsilon \in SOL(K, F_\epsilon)$ for any $\epsilon > 0$, it follows that

$$(F(x_\epsilon) + \epsilon x_\epsilon)^T (x - x_\epsilon) \geq 0 \qquad \text{for any } x \in K. \tag{1}$$

Furthermore, for any $x^* \in SOL(K, F)$ and any $\epsilon > 0$, we have $F(x^*)^T(x_\epsilon - x^*) \geq 0$, which by monotonicity of $F$ implies

$$F(x_\epsilon)^T(x_\epsilon - x^*) \geq 0.$$

Therefore, (by the first relation with $x = x^*$), we have

$$\epsilon x_\epsilon^T(x^* - x_\epsilon) \geq 0 \qquad \text{for any } \epsilon > 0 \text{ and } x^* \in SOL(K, F),$$

implying that

$$x_\epsilon^T x^* \geq \|x_\epsilon\|^2 \qquad \text{for any } \epsilon > 0 \text{ and } x^* \in SOL(K, F).$$

**Consider the case when $x_\epsilon \neq 0$ for any $\epsilon > 0$.** Then, by using the Cauchy-Schwartz inequlaity, from the preceding relation we obtain

$$\|x_\epsilon\| \leq \|x^*\| \qquad \text{for any } x^* \in SOL(K, F).$$

Let $x^\star$ be the smallest norm solution in $SOL(K, F)$. In view of the preceding relation, any limit point $\bar{x}$ of the Tikhonov trajectory has to satisfy $\|\bar{x}\| \leq \|x^\star\|$. This and the fact that $x^\star$ is the smallest norm solution imply that $\bar{x} = x^\star$.

**Consider the case when $x_\epsilon = 0$ for some values of $\epsilon > 0$.** Note that this can happen only in the case when $0 \in K$. Since $x_\epsilon \in SOL(K, F_\epsilon)$, it follows that [see (1)]

$$F(0)^T(x - 0) \geq 0 \qquad \text{for any } x \in K,$$

implying that $0 \in SOL(K, F)$. Note that the smallest norm solution in this case is $x^\star = 0$. Thus, for those $\epsilon > 0$ corresponding to $x_\epsilon = 0$, we have that $x_\epsilon$ coincide with the smallest norm solution. By eliminating those, we can focus on values for $\epsilon > 0$ with $x_\epsilon \neq 0$, for which by the analysis of the preceding case, we have $x_\epsilon \to x^\star$ as $\epsilon \downarrow 0$.

# Extension of the Tikhonov Regularization

The idea of using $\epsilon I$ can be generalized to replacing $I$ with a strongly monotone mapping $G$. The trajectory $\{x_\epsilon\}$is referred to as $G$-**trajectory**, and similar results hold for this trajectory.

**Theorem 3 (Theorem 12.2.5)** *Let $F$ be continuous monotone on $K$, and let $G$ be continuous and strongly monotone on $K$. The followings statements are equivalent for $G$-trajectory:*

*(a) The limit $\lim_{\epsilon \downarrow 0} x_\epsilon$ exists. The limit is the unique solution of $VI(K^*, G)$, where $K^* = SOL(K, F)$.*

*(b) The sequence $\{x_\epsilon\}$ is bounded, i.e.,* $\limsup_{\epsilon \downarrow 0} x_\epsilon < \infty$.

*(c) $SOL(K, F)$ is nonempty.*

# Tikhonov Regularization Algorithm

Data: $x_0 \in \mathbb{R}^n$, $\rho_k \downarrow 0$, and $\epsilon_k \downarrow 0$. Set $k = 0$

*Step 1* If $x_k$ solves $VI(K, F)$ stop. $\limsup_{k\to\infty} t_k > 0$

*Step 1* Compute $x_{k+1}$ such that $\|F^{\rm nat}_{\epsilon_k,K}(x_{k+1})\| \le \rho_k$.
Set $k := k + 1$ and go to Step 1.

**Theorem 4 (modified Theorem 12.2.5)** *Let $F$ be continuous monotone on $K$, and let $SOL(K, F)$ be nonempty and bounded. Then, the sequence generated by the Tikhonov Regularization Method is bounded and each of its limit points is in $SOL(K, F)$.*

# Proximal-Point Methods

- Proximal-point Method alleviates the difficulty arising with Tikhonov regularization due to perturbed problems with $\epsilon_k$ small being "too close" to the original problem.

- Proximal point method is based on a similar idea as Tikhonov regularization, namely solving a sequence of perturbed problems approximating the original problem.

- However, Proximal-point method uses a different perturbation that maintains uniformly strong monotonicity of the approximate problems.

- The perturbation is of the form $I - x + cF$, where $c > 0$ and $x$ is fixed.

- When $F$ is monotone on $K$, every mapping $I - x + cF$ is strongly monotone on $K$ with constant 1, i.e., for all $u, v \in K$,

$$(u - v)^T (I - x + cF)(u) - (I - x + cF)(v)) \geq \|u - v\|.$$

## Proximal-Point Algorithm

Data $\bar{c} > 0$, and $r_1, r_2$ such that $0 < r_1 \le r_2 < 2$.
Initialize the iteration: choose $x_0 \in \mathbb{R}^n$, $c_0 > \bar{c}$, $\rho_0 \in (r_1, r_2)$. Set $k = 0$.

*Step 1* If $x_k$ solves $VI(K,F)$ stop. $\limsup_{k\to\infty} t_k > 0$

*Step 1* Solve $VI(K, I - x_k + c_k F)$ and let $w_k$ be the solution $[w_k = \phi_{c_k}(x_k)]$.

*Step 2* Set $x_{k+1} = (1-\rho_k)x_k + \rho_k w_k$.
Select $c_{k+1}$, and $\rho_{k+1} \in (r_1, r_2)$.
Set $k := k+1$ and go to Step 1.

**Theorem 5 (trimmed down Theorem 12.3.7)** *Let $F$ be continuous monotone on $K$, and let the sequence $\{x_k\}$ be generated by the Proximal-Point Method. If $SOL(K,F)$ is nonempty, then the sequence $\{x_k\}$ converges to a point in $SOL(K,F)$.*

**Proof:** Consider $\|x_{k+1} - x^*\|$ for some $x^{\in} SOL(K, F)$.

$$\|x_{k+1} - x^*\|^2 = \|x_k - x^*\|^2 + \rho_k^2 \|x_k - w_k\|^2 - 2\rho_k (x_k - x^*)^T (x_k - w_k).$$

Since $w_k = \phi_{c_k}(x_k)$, we can write

$$x_k - w_k = x_k - \phi_{c_k}(x_k) = (I - \phi_{c_k})(x_k) = Q_k(x_k).$$

Since every solution $x^* \in SOL(K, F)$ also solves $VI(K, I - x^* + cF)$ for any $c > 0$, it follows that $Q_k(x^*) = 0$. Hence, we can write

$$\|x_{k+1} - x^*\|^2 = \|x_k - x^*\|^2 + \rho_k^2 \|Q_k(x_k)\|^2 - 2\rho_k (x_k - x^*)^T (Q_k(x_k) - Q_k(x^*)).$$

Recall that $Q_k$ is co-coercive when $F$ is monotone, so we further have

$$\begin{aligned}\|x_{k+1} - x^*\|^2 &\le \|x_k - x^*\|^2 + \rho_k^2\|Q_k(x_k)\|^2 - 2\rho_k\|Q_k(x_k) - Q_k(x^*)\|^2 \\ &\le \|x_k - x^*\|^2 - \rho_1(2 - \rho_2)\|Q_k(x_k)\|^2,\end{aligned}$$

where in the last relation we use $Q_k(x^*) = 0$ and the bounds on $\rho_k$.
We have that the sequence $\{x_k\}$ is bounded and the scalar sequence $\{\|x_k - x^*\|\}$ converges for every $x^* \in SOL(K, F)$. The proof will be complete if we can show that one of the limit points of $\{x_k\}$ lies in $SOL(K, F)$.

In view of the last relation, we have $\lim_{k\to\infty} Q_k(x_k) = 0$. By the definition of $Q_k(x_k)$, we have that $\|x_k - \phi_{c_k}(x_k)\| \to 0$ as $k \to \infty$.

Let $\bar{x}$ be a limit point of $\{x_k\}$, then $\phi_{c_k}(x_k) - \bar{x}\| \to 0$ along the subsequence $\{x_k\}_{\mathcal{K}}$ converging to $\bar{x}$. Without loss of generality, since $\{c_k\}$ is bounded, we may assume that $\{c_k\}_{\mathcal{K}}$ converges to some $\tilde{c}$.

Since $\phi_{c_k}(x_k) \in SOL(K, I - x_k + c_k F)$, we have for all $k \in \mathcal{K}$,

$$\phi_{c_k}(x_k) = \Pi_K[\phi_{c_k}(x_k) - \phi_{c_k}(x_k) + x_k - c_k F(\phi_{c_k}(x_k))].$$

By letting $k \to \infty$ with $k \in \mathcal{K}$, using the continuity of $F$, we obtain

$$\bar{x} = \Pi_K[\bar{x} - \tilde{c} F(\bar{x})],$$

showing that $\bar{x} \in SOL(K, F)$.

# Approximate Proximal-Point Algorithm

Data $\bar{c} > 0$, and $r_1, r_2$ such that $0 < r_1 \le r_2 < 2$.
Initialize the iteration: choose $x_0 \in \mathbb{R}^n$, $c_0 > \bar{c}$, $\rho_0 \in (r_1, r_2)$, and $\epsilon_0$. Set $k = 0$.

*Step 1* If $x_k$ solves $VI(K, F)$ stop. $\limsup_{k\to\infty} t_k > 0$

*Step 1* Solve $VI(K, I - x_k + c_k F)$ approximately and let $w_k$ be the solution such that

$$\|w_k - \phi_{c_k}(x_k)\| \le \epsilon_k.$$

*Step 2* Set $x_{k+1} = (1 - \rho_k)x_k + \rho_k w_k$.
Select $c_{k+1}$, $\epsilon_{k+1}$, and $\rho_{k+1} \in (r_1, r_2)$.
Set $k := k + 1$ and go to Step 1.

**Theorem 6 (trimmed down Theorem 12.3.7)** *Let $F$ be continuous monotone on $K$, and let the sequence $\{x_k\}$ be generated by the Proximal-Point Method. If $SOL(K, F)$ is nonempty and $\sum_k \epsilon_k < \infty$, then the sequence $\{x_k\}$ converges to a point in $SOL(K, F)$.*

**Proof Sketch**: The idea is to define

$$y(k) = \phi_{c_k}(x_k),$$

and look at the $\|x_{k+1} - y(k)\|$ and $\|y(k) - x^*\|$ in order to estimate $\|x_{k+1} - x^*\|$. The error $\|x_{k+1} - y(k)\|$ is of the order $\rho_k \epsilon_k$, which is summable, and makes the method work. In general, the analysis is very similar to the case without errors $\epsilon_k$.