Lecture 21

Algorithms for Monotone VIs

Projection Methods

November 17, 2008

Outline

- Projection Methods
  - Basic Fixed-Point Iteration
  - Extra-gradient Method
  - Hyperplane Projection Method

Introduction

- Projection methods are conceptually simple methods for solving monotone $VI(K, F)$ for a convex closed set K
- Their advantages are
  - Easily implementable and computationally inexpensive
    - When K has structure that makes the projection on K easy
- Makes them suitable for large scale problems
- Also, they are often used as a sub-procedures in faster and more complex methods (enabling the moves into "promising" regions)
- Their main disadvantage is slow progress since they do not use higher order information

Basic Fixed-Point Iteration

Throughout the rest, we assume that $K \subset \mathbb{R}^n$ is **closed and convex**, and $F : K \to \mathbb{R}^n$ is **continuous mapping**.

- We consider the "skewed" natural map

$$F^{\rm nat}_{K,D}(x) = x - \Pi_{K,D}[x - D^{-1}F(x)],$$

where D is a symmetric and positive definite matrix.

- Such a matrix, induces and inner product and a norm in $\mathbb{R}^n$:

$$< x, y >_D \; = \; x^T Dy, \qquad \|x\| = \sqrt{x^T Dx}.$$

- The skewed projection $\Pi_{K,D}$ on K is just a projection with respect to norm $\|\cdot\|_D$,

$$\hat{x} = \Pi_{K,D}[x] \qquad \Longleftrightarrow \qquad \hat{x} \text{ solves } \quad \min_{y\in K} \|x - y\|^2_D.$$

Fact

The projection mapping $x \mapsto \Pi_{K,D}[x]$ is nonexpansive in the norm $\|\cdot\|_D$

- **Fact** A vector x^* solves $VI(KF)$ if and only if x^* is zero of the skewed natural map, i.e.,

$$x^* = \Pi_{K,D}[x^* - D^{-1}F(x^*)].$$

- Define $\Phi(x) = \Pi_{K,D}[x - D^{-1}F(x)]$ and note that $\Phi : K \to K$.

- When $D = I$ and Φ is contraction with respect to Euclidean norm), we have seen that fixed-point method $x_{k+1} = \Phi(x_k)$ produces a sequence with accumulation points being fixed points of Φ (for any initial $x_0 \in K$)

- The idea of fixed-point here is similar: if we could ensure that $\Phi(x) = \Pi_{K,D}[x - D^{-1}F(x)]$ is a contraction in norm $\|\cdot\|_D$, then we could use "fixed-point" method to find a fixed point of $\Phi(x)$ and hence, a solution to $VI(K, F)$

Basic Projection Method

Choose a symmetric positive definite matrix D. Select an initial vector $x_0 \in K$.

Step 0: Set $k = 0$

Step 1: Compute $x_{k+1} = \Pi_{K,D}[x_k - D^{-1}F(x_k)]$. If $x_{k+1} = x_k$, stop; we have a solution.

Step 2: Otherwise, set $k := k + 1$, and go to Step 1.

In order to ensure the convergence of the sequence $\{x_{k+1}\}$ (or its subsequence) to a fixed point of Φ, we need some conditions of the mapping F and the matrix D.

Convergence for Strongly Monotone Mappings

Theorem 1 *Let $F : K \to \mathbb{R}^n$. Suppose F is strongly monotone and Lipschitz continuous on K,*

$$(F(x) - F(y))^T (x - y) \geq \mu \|x - y\|^2, \qquad \|F(x) - F(y)\| \leq L\|x - y\|^2.$$

Also, let

$$\lambda_{\max} < \frac{2\mu}{L^2} \lambda_{\min}^2,$$

where $\lambda_{\max}$ and $\lambda_{\min}$ are the largest and smallest eigenvalues of D. Then, the mapping $\Pi_{K,D}[x - D^{-1}F(x)]$ is contraction in $\|\cdot\|_D$ with contraction factor

$$\eta = 1 - \frac{L^2}{\lambda_{\max} \lambda_{\min}^2} \left(\frac{2\mu\lambda_{\min}^2}{L^2} - \lambda_{\max} \right).$$

Therefore, the sequence $\{x_k\}$ converges to the unique solution of $VI(K, F)$.

Proof: For any two vectors x, y in K, we have

$$\|\Pi_{K,D}[x - D^{-1}F(x)] - \Pi_{K,D}[y - D^{-1}F(y)]\|_D^2 \le \left\|x - D^{-1}F(x) - \left(y - D^{-1}F(y)\right)\right\|_D^2.$$

Why? By expanding the last term, we have

$$\begin{aligned}\|\Pi_{K,D}[x - D^{-1}F(x)] - \Pi_{K,D}[y - D^{-1}F(y)]\|_D^2 &\le \left\|x - y\right\|_D^2 \\ &- 2(F(x) - F(y))^T(x - y) + \|F(x) - F(y)\|_{D^{-1}}^2.\end{aligned}$$

Using the strong convexity, we have

$$(F(x) - F(y))^T(x - y) \ge \mu\|x - y\|^2 \ge \frac{\mu}{\lambda_{\max}}\|x - y\|_D^2. \tag{1}$$

From

$$\|F(x) - F(y)\|_{D^{-1}}^2 \le \frac{1}{\lambda_{\min}}\|F(x) - F(y)\|^2$$

and Lipschitz continuity of F, we obtain

$$\|F(x) - F(y)\|_{D^{-1}}^2 \le \frac{L^2}{\lambda_{\min}^2}\|x - y\|_D^2 \tag{2}$$

By combining the estimates in (1) and (2), we obtain

$$\|\Pi_{K,D}[x - D^{-1}F(x)] - \Pi_{K,D}[y - D^{-1}F(y)]\|_D^2 \leq \left(1 - \frac{2\mu}{\lambda_{\max}} + \frac{L^2}{\lambda_{\min}^2}\right)\|x - y\|_D^2.$$

Hence, under the given relation for the eigenvalues of D, the mapping $x \mapsto \Pi_{K,D}[x - D^{-1}F(x)]$ is contraction in $\|\cdot\|_D$ and the results follow.

- When $D = \frac{1}{\alpha}I$, the eigenvalue condition reduces to $L^2 < 2\mu/\alpha$. Thus, if we let $\alpha < \frac{2\mu}{L^2}$, the condition is satisfied.

- The corresponding algorithm becomes

$$x_{k+1} = \Pi_K \left[x_k - \alpha F(x_k)\right].$$

- Problem is: we do not always have access to L and μ.
  Hence, we do not know how small α should be to ensure convergence

Co-coercive Mapping

The projection method can be used to solve $VI(K, F)$ with co-coercive mapping. The reason why this work is that the (Euclidean) projection is co-coercive, and when F is co-coercive, then the τ-natural mapping $x \mapsto \Pi_K[x - \tau F(x)]$ is also co-coercive for some range of values of τ.
In particular, we have the following results

Lemma 1 *Let $F : K \to \mathbb{R}^n$ be co-coercive with constant c. If $0 < \tau < 4c$, then $F^{\rm nat}_{K,\tau}(x) = \Pi_K[x - \tau F(x)]$ is co-coercive with constant $1 - \frac{\tau}{4c}$.*

Proof: See Lemma 12.1.7 of FP-II.

Convergence for Co-coercive Mapping

Theorem 2 *Assume that $VI(K,F)$ has a solution. Let $F : K \to \mathbb{R}^n$ be co-coercive with constant c. Consider the projection method*

$$x_{k+1} = \Pi[x_k - \tau F(x_k)],$$

where $\tau < 2c$. Then, the sequence $\{x_k\}$ converges to a solution of $VI(K,F)$.

Proof We have for a fixed point x^* of $F_K^{\rm nat}$ (also a fixed point of $F_{K,\tau}^{\rm nat}$ for any $\tau > 0$),

$$\begin{aligned}
\|x_{k+1} - x^*\|^2 &\le \|\Pi[x_k - \tau F(x_k) - \Pi[x^* - \tau F(x^*)]\|^2 \\
&\le \|x_k - x^* - \tau(F(x_k) - F(x^*))\|^2 \\
&= \|x_k - x^*\|^2 - 2\tau(F(x_k) - F(x^*))^T(x - x^*) \\
&+ \tau^2\|F(x_k) - F(x^*)\|^2
\end{aligned}$$

Since F is co-coercive, we obtain

$$\|x_{k+1} - x^*\|^2 \leq \|x_k - x^*\|^2 - \tau(2c - \tau) \|F(x_k) - F(x^*)\|^2 \qquad (3)$$

Since $2c - \tau > 0$, it follows that

$$dist(x_{k+1}, SOL(K, F)) \leq dist(x_k, SOL(K, F)) \qquad \text{for all } K.$$

(Recall, that $SOL(K, F)$ is convex and closed here - why?)

Furthermore, by summing the relations in (3) over $k = 0, \ldots, N$ for arbitrary N, and letting $N \to \infty$, we obtain

$$\sum_k \|F(x_k) - F(x^*)\|^2 < \infty,$$

implying $F(x_k) \to F(x^*)$.

(This also implies that $\{F(x^*) \mid x^* \in SOL(K,F)\}$ is a singleton - why would this be expected). In view of the preceding two relations, it follows that

$$dist(x_k, SOL(K,F)) \to 0.$$

Note, also that (3) implies that the scalar sequence $\{\|x_k - x^*\|\}$ is nonincreasing for any fixed point x^*. Therefore, the scalar sequence $\{\|x_k - x^*\|\}$ is convergent for any fixed point x^*. This, and $dist(x_k, SOL(K,F)) \to 0$ imply that $\{x_k\}$ is convergent and its limit point is in $SOL(K,F)$.

The proof in the FP-II is given in Lemma 12.1.15 for the case when

- τ is varying and $\tau_k F(x_k)$ is replaced with $F^k(x_k)$, where all $VI(K,F^k)$ have the same solution set.
- $K = \mathbb{R}^n$
- Assuming that each mapping $F^k : \mathbb{R}^n \to \mathbb{R}^n$ is co-coercive with c_k, and the following condition is satisfied: $\inf_k c_k > 1/2$.

Extra-Gradient Method

The name of method comes from its equivalent version in optimization $(F = \nabla f(x))$.
It has two projection steps

$$z_{k+1} = \Pi_K[x - \tau F(x_k)], \qquad x_{k+1} = \Pi_K[x_k - \tau F(z_{k+1})]$$

The main iterate is x_{k+1}. The extra-iterate z_{k+1} is used to construct the direction for moving away from x_k.

- The advantage of taking an extra step is that the algorithm performs better than the projection method
- The convergence analysis still requires F to be Lipschitz
- It can be used to solve $VI(K, F)$ with a pseudo-monotone map F
- We will study it as applied to a monotone $VI(K, F)$

Basic Iterate Relation

Theorem 3 *Let F be monotone and Lipschitz continuous on K with constant L. Let x^* be a solution of $VI(K,F)$. Then, for all k we have*

$$\|x_{k+1} - x^*\|^2 \leq \|x_k - x^*\|^2 - (1 - \tau^2 L^2)\|z_{k+1} - x_{k+1}\|^2.$$

Proof For a projection on convex closed set K, we have for any $x \in \mathbb{R}^n$,

$$\|\Pi_K[x] - z\|^2 \leq \|x - z\|^2 - \|\Pi_K[x] - x\|^2 \qquad \text{for all } z \in K.$$

(can be seen by a more careful analysis in Projection Theorem).
Using this relation with $x = x_{k+1}$ and $z = x^*$, we obtain

$$\|x_{k+1} - x^*\|^2 \leq \|x_k - \tau F(z_{k+1}) - x^*\|^2 - \|x_{k+1} - (x_k - \tau F(z_{k+1}))\|^2.$$

By expanding the terms on the right hand side, we have

$$\|x_{k+1} - x^*\|^2 \leq \|x_k - x^*\|^2 - \|x_{k+1} - x_k\|^2 + 2\tau F(z_{k+1})^T (x^* - x_{k+1}). \quad (4)$$

By the monotonicity of F and $x^* \in SOL(K,F)$, it follows

$$(F(z_{k+1}) - F(x^*))^T(z_{k+1}) - x^*) \geq 0 \quad \Longrightarrow \quad F(z_{k+1})^T(z_{k+1} - x^*) \geq 0.$$

Hence, $F(z_{k+1})^T(z_{k+1} - x_{k+1}) + F(z_{k+1})^T(x_{k+1} - x^*) \geq 0$ implying that

$$F(z_{k+1})^T(z_{k+1} - x_{k+1}) \geq F(z_{k+1})^T(x^* - x_{k+1})$$

Using this relation in (4), we see

$$\|x_{k+1} - x^*\|^2 \leq \|x_k - x^*\|^2 - \|x_{k+1} - x_k\|^2 + 2\tau F(z_{k+1})^T(z_{k+1} - x_{k+1}).$$

By writing $x_{k+1} - x_k = (x_{k+1} - z_{k+1}) + (z_{k+1} - x_k)$ and expanding the squared-norm of this term, and then combining the terms that are in the inner product with $z_{k+1} - x_{k+1}$, we obtain

$$\begin{aligned}\|x_{k+1} - x^*\|^2 \ &\leq \|x_k - x^*\|^2 - \|x_{k+1} - z_{k+1}\|^2 - \|z_{k+1} - x_k\|^2 \\ &+2(x_{k+1} - z_{k+1})^T(x_k - \tau F(z_{k+1}) - z_{k+1}).\end{aligned}$$

We can further write [by adding and subtracting $\tau F(x_k)$]

$$\begin{aligned}(x_{k+1} - z_{k+1})^T \quad & (x_k - \tau F(z_{k+1}) - z_{k+1}) = (x_{k+1} - z_{k+1})^T(x_k - \tau F(z_k) - z_{k+1}) \\ & +\tau(x_{k+1} - z_{k+1})^T(F(x_k) - F(z_{k+1}))\end{aligned}$$

Since $x_{k+1} \in K$ and $z_{k+1} = \Pi_K[x_k - \tau F(x_k)]$, the first term on the right hand side is nonnegative (by projection property). Thus, by using this and Lipschitz continuity of F, we have

$$\begin{aligned}(x_{k+1} - z_{k+1})^T \quad & (x_k - \tau F(z_{k+1}) - z_{k+1}) \\ & \leq \tau(x_{k+1} - z_{k+1})^T(F(x_k) - F(z_{k+1})) \\ & \leq \tau L\|x_{k+1} - z_{k+1}\| \cdot \|x_k - z_{k+1}\| \\ & \leq \frac{1}{2}\left(\|x_{k+1} - z_{k+1}\|^2 + \tau^2 L^2\|x_k - z_{k+1}\|^2\right)\end{aligned}$$

By substituting the preceding estimate in (5), we obtain

$$\begin{aligned}\|x_{k+1} - x^*\|^2 &\le \|x_k - x^*\|^2 - \|x_{k+1} - z_{k+1}\|^2 - \|z_{k+1} - x_k\|^2 \\ &\quad + \|x_{k+1} - z_{k+1}\|^2 + \tau^2 L^2 \|x_k - z_{k+1}\| \\ &= \|x_k - x^*\|^2 - (1 - \tau^2 L^2)\|z_{k+1} - x_k\|^2.\end{aligned}$$

Convergence of Extra-Gradient Method

Theorem 4 *Let F be monotone and Lipschitz continuous over K with constant L. Let $SOL(K, F)$ be nonempty. Then, with $\tau < \frac{1}{L}$, the sequence $\{x_k\}$ generated by the extra-gradient method converges to a solution of $VI(K, F)$.*

Proof The line of proof relies on the basic iterate relation, and follows a line of analysis similar to that of Theorem for co-coercive map and projection method. See FP-II Theorem 12.1.11.

For the estimate in basic relation to result in convergence, we need $\tau < \frac{1}{L}$.

- In practice, often L is not available
- We can use diminishing step τ_k at iteration k, with $\sum_k \tau_k = +\infty$
- But the convergence will slow down
- We next consider a modification of the method not relying on the Lipschitz continuity

Hyperplane Projection Method

Like the extra-gradient method, this method generates an extra-iterate $y_k = \Pi_K[x_k - \tau F(x_k)]$. However, the use of this point in constructing the new iterate is different.
In particular, Armijo search rule is used to determine a point z_k defining a hyperplane

$$H_k = \{x \in \mathbb{R}^n \mid F(z_k)^T(x_k - z_k) = 0\},$$

which separates x^k strongly from the solution set $SOL(K,F)$ [for continuous monotone map, this set is closed and convex - possibly empty].

In particular, z_k is such that, for some positive scalars t, σ, τ,

$$F(z_k)^T(x_k - z_k) \ge \frac{t_k \sigma}{\tau} \|y_k - x_k\|^2,$$

which strongly separates x_k from $SOL(K,F)$ whenever $y_k \neq x_k$ in view of

$$0 \ge F(x^*)^T(x^* - z_k) \ge F(z_k)^T(x^* - z_k) \qquad \text{for all } x^* \in SOL(K,F).$$

The algorithm is initialized with the following parameters:
$k = 0$, $x_0 \in K$, $\tau > 0$, and $\sigma \in (0, 1)$. The algorithm proceeds as follows.

Step 1: Compute $y_k = \Pi_{K,D}[x_k - D^{-1}F(x_k)]$.
If $y_k = x_k$, stop; we have a solution. Else, set $t = 1$ and go to Step 2.

Step 2: (Separation Test) If

$$F(ty_k + (1-t)x_k)^T(x_k - y_k) \geq \sigma\tau \|x_k - y_k\|^2,$$

set $t_k = t$ and go to Step 3; otherwise set $t = \frac{t}{2}$ and repeat the test.

Step 3: Set $z_k = t_k y_k + (1 - t_k)x_k$, and compute

$$w_k = \Pi_{H_k}[x_k) = x_k - \frac{F(z_k)^T(x_k - z_k)}{\|F(z_k)\|^2} F(z_k).$$

Step 4: Set $x_{k+1} = \Pi_K[w_k]$ and set $k := k + 1$, and go to Step 1.

Preliminary Results

Convergence proof uses three preliminary results.

Projection properties (Lemma 12.1.13)

(a) For any $x, y \in \mathbb{R}^n$, we have

$$\|\Pi_K[x] - \Pi_K[y]\|^2 \leq \|x - y\|^2 - \|\Pi_K[x] - x + y - \Pi_K[y]\|^2 .$$

(b) For any $x \in K$ and $y \in \mathbb{R}^n$, we have

$$(x - y)^T (x - \Pi_k[y]) \geq \|x - \Pi_K[y]\|^2.$$

Finite Termination of Loop in Step 2

To have the algorithm well defined, we need to verify that the Step 2 is exited after finitely many trials (test evaluations).
The following lemma provides the guarantee.

Lemma 2 (Based on Lemma 12.1.14) *Let F be continuous over K, and let x_k be such that $x_k \notin SOL(K, F)$. We then have*

- *Step 2 is exited after finitely many trials.*
- *For z_k defined at Step 3, we have*

$$F(z_k)^T(x_k - z_k) > \frac{t_k \sigma}{\tau} \|x_k - y_k\|^2 > 0,$$

Proof Suppose that Step 2 is never exited. Then, it follows that for all i,

$$F(2^{-i} y_k + (1 - 2^{-i}) x_k)^T (x_k - y_k) < \frac{\sigma}{\tau} \|x_k - y_k\|^2.$$

Letting $i \to \infty$, by continuity of F we obtain

$$F(x_k)^T (x_k - y_k) < \frac{\sigma}{\tau} \|x_k - y_k\|^2.$$

Therefore, since $\tau > 0$,

$$\begin{aligned}\sigma\|x_k - y_k\|^2 &\geq \tau F(x_k)^T(x_k - y_k)\\ &= (x_k - (x_k - \tau F(x_k)))^T(x_k - y_k)\\ &\geq \|x_k - y_k\|^2,\end{aligned}$$

where the last inequality follows from $y_k = \Pi_K[x_k - \tau F(x_k)]$, $x_k \in K$, and the projection property (second bullet).

Since $z_k = t_k y_k + (1 - t_k)x_k$, we have $x_k - z_k = t_k(x_k - y_k)$. Hence,

$$F(z_k)^T(x_k - z_k) = t_k\, F(z_k)^T(x_k - y_k) \qquad .$$

By the construction of z_k at Step 3 and from Step 2, we have

$$F(z_k)^T(x_k - y_k) > \sigma\tau\,\|x_k - y_k\|^2.$$

Combining the preceding two relations, we obtain

$$F(z_k)^T(x_k - z_k) > \frac{t_k\sigma}{\tau}\,\|x_k - y_k\|^2 > 0,$$

Auxiliary Lemma

Lemma 3 *Let F be continuous and monotone on K. let $\{x_k\}$ be generated by the Hyperplane Projection Method. We then have*

(a) The sequence $\{\|x_k - x^*\|\}$ is nonincreasing (hence, convergent) for every $x^* \in SOL(K, F)$.

(b) The sequence $\{x_k\}$ is bounded.

(c) If $\{x_k\}$ has an accumulation point in $SOL(K, F)$, then the whole sequence converges to this solution.

(d) $\lim_{k\to\infty} F(z_k)^T(x_k - z_k) = 0$.

Proof Note that (a) implies (b). Also, note that (a) and (b) together imply (c). Thus, only (a) and (d) need a proof.
Define the half-space

$$H_k^- = \{x \in \mathbb{R}^n \mid F(z_k)^T(x - z_k) \le 0\}.$$

As noted earlier, by monotonicity of F, it can be seen that

$$SOL(K,F) \subseteq H_k^-.$$

Note that $x_k \notin H_k^-$ in view of preceding lemma [part (b)]. Therefore,

$$x^* = \Pi_K[x^*], \qquad x^* = \Pi_{H_k^-}[x^*] \qquad \text{for all } x^* \in SOL(K,F).$$

Furthermore, note that by definition of w_k, we have $w_k = \Pi_{H_k^-}[x_k]$.

We now consider $x_{k+1} - x^*$ for arbitrary $x^* \in SOL(K,F)$. By projection property (first bullet), and the preceding relations, we have

$$\begin{aligned}
\|x_{k+1} - x^*\|^2 &= \|\Pi_K[w_k] - \Pi_K[x^*]\|^2 \\
&\le \|w_k - x^*\|^2 - \|\Pi_K[w_k] - w_k\|^2 \\
&= \|\Pi_{H_k^-}[x_k] - \Pi_{H_k^-}[x^*]\|^2 - \|\Pi_K[w_k] - w_k\|^2 \\
&\le \|x_k - x^*\|^2 - \|w_k - x^*\|^2 - \|\Pi_K[w_k] - w_k\|^2 \qquad (7)
\end{aligned}$$

Thus, we have $\|x_{k+1} - x^*\| \leq \|x_k - x^*\|$, showing that the scalar sequence $\{\|x_k - x^*\|\}$ is nonincreasing, and hence convergent. This proves part (a). From (7), it also follows

$$\|w_k - x^k\|^2 \leq \|x_k - x^*\|^2 - \|x_{k+1} - x^*\|^2.$$

Since the scalar sequence $\{\|x_k - x^*\|\}$ is convergent, we have

$$\lim_{k\to\infty} \|w_k - x_k\| = 0.$$

By the definition of w_k, from the preceding relation we see that

$$\lim_{k\to\infty} \frac{F(z_k)^T(x_k - z_k)}{\|F(z_k)\|} = 0. \tag{8}$$

Since $\{x_k\}$ is bounded, $\{F(x_k)\}$ is also bounded by continuity of F. By $y_k = \Pi_K[x_k - \tau F(x_k)]$, we obtain that $\{y_k\}$ is also bounded. Since z_k is a convex combination of y_k and x_k, we see that $\{z_k\}$ is bounded, which by continuity of F implies that $\{F(z_k)\}$ is bounded. Therefore, in view of (8),

$$\lim_{k\to\infty} F(z_k)^T(x_k - z_k) = 0,$$

showing part (d).

Theorem 5 *Let F be a continuous monotone mapping on K, and assume that $SOL(K,F)$ is nonempty. Then, the sequence $\{x_k\}$ generated by Hyperplane Projection Method converges to a solution $x^* \in SOL(K,F)$.*

Proof Idea

The key is to show that a subsequence of $\{x_k\}$ converges to some point in $SOL(K,F)$, and then use the preceding lemma to conclude that the entire sequence converges to that point.

Showing that a subsequence of $\{x_k\}$ converges to a point in $SOL(K,F)$ is done by examining the following two possibilities for the stepsize sequence:

(1) $\limsup_{k\to\infty} t_k > 0$

(2) $\limsup_{k\to\infty} t_k = 0$, which implies $\lim_{k\to\infty} t_k = 0$.