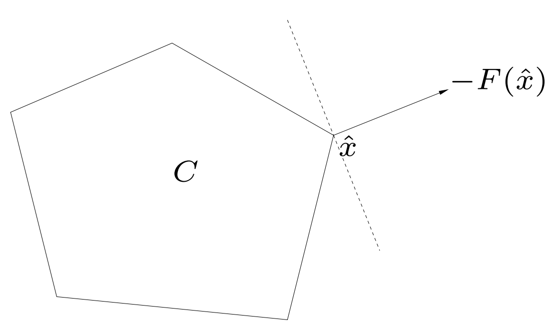
16. Variational inequalities

- variational inequality
- monotonicity
- examples
- linear complementarity problem
- analytic center cutting-plane method
- extragradient method

Variational inequality

given closed convex set C, mapping $F : \mathbf{R}^n \to \mathbf{R}^n$, find $\hat{x} \in C$ such that

$$F(\hat{x})^T(x - \hat{x}) \geq 0 \quad \forall x \in C$$



equivalently, solve

$$\hat{x} = P_C(\hat{x} - F(\hat{x}))$$

where P_C is projection on C

Monotonicity

we will focus on variational inequalities with *monotone* F:

$$\left(F(u) - F(v)\right)^T (u - v) \geq 0 \quad \forall u, v$$

- F is strictly monotone if

$$\left(F(u) - F(v)\right)^T (u - v) > 0 \quad \forall u, v, u \neq v$$

- F is strongly monotone if there exists a $\sigma > 0$ such that

$$\left(F(u) - F(v)\right)^T (u - v) \geq \frac{\sigma}{2} \|u - v\|_2^2 \quad \forall u, v$$

Convex optimization

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & x \in C \end{array}$$

with f convex, C a convex set

optimality condition (236B lecture 4-9)

$\hat{x} \in C$ is optimal if

$$\nabla f(\hat{x})^T (x - \hat{x}) \geq 0 \quad \forall x \in C$$

this is a variational inequality with $F(x) = \nabla f(x)$

monotonicity

$F(x) = \nabla f(x)$ is monotone if and only if f is convex

- if f is convex, then for all $u, v \in \mathbf{dom}\, f$,

$$\begin{array}{rcl} (\nabla f(u) - \nabla f(v))^T (u - v) & = & -\nabla f(u)^T (v - u) - \nabla f(v)^T (u - v) \\ & \geq & (f(u) - f(v)) + (f(v) - f(u)) \\ & = & 0 \end{array}$$

- if $\nabla f(x)$ is monotone, then for all $x, y \in \mathbf{dom}\, f$,

$$\begin{array}{rcl} f(y) & = & f(x) + \int_0^1 \nabla f(x + t(y - x))^T (y - x)\, dt \\ & \geq & f(x) + \nabla f(x)^T (y - x) \end{array}$$

Convex-concave saddle-point problems

suppose $f(u, v)$ is convex-concave and U and V are convex

saddle-point: $(\hat{u}, \hat{v}) \in U \times V$ is a saddle-point

$$f(\hat{u}, v) \leq f(\hat{u}, \hat{v}) \leq f(u, \hat{v}) \quad \forall (u, v) \in U \times V \qquad (1)$$

variational inequality formulation $(\hat{u}, \hat{v})$ is a saddle-point iff

$$\left[\begin{array}{c} \nabla_u f(\hat{u}, \hat{v}) \\ -\nabla_v f(\hat{u}, \hat{v}) \end{array} \right]^T \left[\begin{array}{c} u - \hat{u} \\ v - \hat{v} \end{array} \right] \geq 0 \quad \forall (u, v) \in U \times V \qquad (2)$$

this is a variational inequality with $F(u, v) = (\nabla_u f(u, v), -\nabla_v f(u, v))$

proof

- if $(\hat{u}, \hat{v})$ satisfies the variational inequality, then for all $(u, v) \in U \times V$

$$f(\hat{u}, \hat{v}) \leq f(\hat{u}, \hat{v}) + \nabla f_u(\hat{u}, \hat{v})^T (u - \hat{u}) \leq f(u, \hat{v})$$

$$f(\hat{u}, \hat{v}) \geq f(\hat{u}, \hat{v}) + \nabla f_v(\hat{u}, \hat{v})^T (v - \hat{v}) \geq f(\hat{u}, v)$$

  therefore (1) holds

- if $(\hat{u}, \hat{v})$ is a saddle-point, then $\hat{u}$ minimizes $f(u, \hat{v})$ over $u \in U$, *i.e.*,

$$\nabla f_u(\hat{u}, \hat{v})^T (u - \hat{u}) \geq 0 \quad \forall u \in U$$

  $\hat{v}$ maximizes $f(\hat{u}, v)$ over $v \in V$, *i.e.*,

$$\nabla f_v(\hat{u}, \hat{v})^T (v - \hat{v}) \leq 0 \quad \forall v \in V$$

  therefore (2) holds

monotonicity

$$F(u, v) = \left[\begin{array}{c} \nabla_u f(u, v) \\ -\nabla_v f(u, v) \end{array} \right]$$

is monotone if and only if f is convex-concave

- if f is convex-concave, then for all $w = (u, v)$, $\hat{w} = (\hat{u}, \hat{v})$

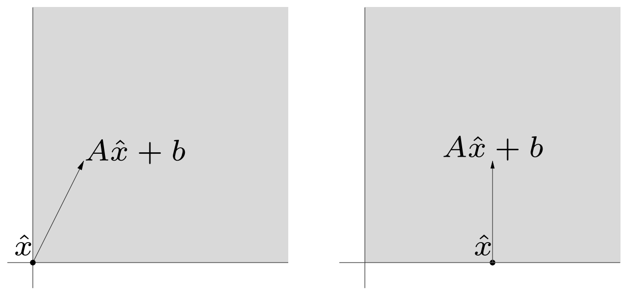
$$\begin{aligned} &(F(w) - F(\hat{w}))^T (w - \hat{w}) \\ &\quad = \quad (\nabla_u f(w) - \nabla_u f(\hat{w}))^T (u - \hat{u}) - (\nabla_v f(w) - \nabla_v f(\hat{w}))^T (v - \hat{v}) \\ &\quad \geq \quad -f(\hat{u}, v) + f(u, v) - f(u, \hat{v}) + f(\hat{u}, \hat{v}) + f(u, \hat{v}) - f(u, v) \\ &\qquad\quad + f(\hat{u}, v) - f(\hat{u}, \hat{v}) \\ &\quad = \quad 0 \end{aligned}$$

- proof of converse is similar as proof on page 16–5

Linear complementarity problem

given $A \in \mathbf{R}^{n \times n}$, $b \in \mathbf{R}^n$, find $\hat{x}$

$$\hat{x} \succeq 0, \qquad A\hat{x} + b \succeq 0, \qquad (A\hat{x} + b)_i \hat{x}_i = 0, \quad i = 1, \ldots, n$$



- this is a variational inequality with $F(x) = Ax + b$, $C \in \mathbf{R}^n_+$
- F is monotone if $A + A^T \succeq 0$
- for certain classes of A, can be solved using simplex-like algorithms

Quadratic program as LCP

$$\begin{array}{ll} \text{minimize} & (1/2)x^T P x + q^T x \\ \text{subject to} & Gx \preceq h, \quad x \succeq 0 \end{array}$$

(with $P \succeq 0$)

optimality conditions

$$Px + G^T z - u + q = 0, \qquad Gx + s = h, \qquad \begin{bmatrix} x \\ s \end{bmatrix} \succeq 0, \quad \begin{bmatrix} u \\ z \end{bmatrix} \succeq 0$$

$$x^T u + s^T z = 0$$

this can be written as a (monotone) linear complementarity problem with

$$F(x, s) = \begin{bmatrix} P & G^T \\ -G & 0 \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix} + \begin{bmatrix} q \\ h \end{bmatrix}$$

Bimatrix game as LCP

- two players, using randomized strategies

$$x \in X = \{x \succeq 0 \mid \mathbf{1}^T x = 1\}, \qquad y \in Y = \{y \succeq 0 \mid \mathbf{1}^T y = 1\}$$

- cost to player 1 is $x^T A y$; cost to player 2 is $x^T B y$
- we assume A and B have nonnegative entries

Nash equilibrium: $\hat{x} \in X$, $\hat{y} \in Y$ with

$$\hat{x}^T A \hat{y} \le x^T A \hat{y} \quad \forall x \in X, \qquad \hat{x}^T B \hat{y} \le \hat{x}^T B y \quad \forall y \in Y$$

in other words, $\hat{x}^T A \hat{y} = \min_k (A\hat{y})_k = \min_k (B^T \hat{x})_k$

bimatrix game as linear complementarity problem

$$F(x,y) = \begin{bmatrix} 0 & A \\ B^T & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} - \begin{bmatrix} \mathbf{1} \\ \mathbf{1} \end{bmatrix}$$

- if $\hat{x}$, $\hat{y}$ is a Nash equilibrium, then

$$x = \frac{\hat{x}}{\hat{x}^T B \hat{y}}, \qquad y = \frac{\hat{y}}{\hat{x}^T A \hat{y}}$$

  solve the linear complementarity problem

- if x, y solve the complementarity problem, then

$$\hat{x} = \frac{x}{\mathbf{1}^T x}, \qquad \hat{y} = \frac{y}{\mathbf{1}^T y}$$

  is a Nash equilibrium

(this LCP is not monotone, but can still be solved efficiently)

Duality gap function

$$\eta(x) = \sup_{y \in C} F(y)^T (x - y), \qquad \mathbf{dom}\, \eta = C$$

is called the *duality gap function* of the variational inequality

properties

- $\eta(x)$ is a convex function
- $\eta(x) \geq 0$ for all $x \in C$
- (for F monotone and continuous) $\eta(x) = 0$ if and only if x solves the variational inequality (proof on next page)

consequence

solution set of monotone variational inequality is convex (possibly empty)

proof of 3rd property

- if x solves the variational inequality, then, by monotonicity,

  $$F(y)^T (x - y) \leq F(x)^T (x - y) \leq 0 \quad \forall y \in C$$

  therefore $\eta(x) = \sup_{y \in C} F(y)^T (x - y) = 0$

- suppose $\eta(x) = 0$; consider $w = x + t(z - x)$ with $z \in C$, $0 < t \leq 1$:

  $$0 = \eta(x) \geq F(w)^T (x - w) = tF\left(x + t(z - x)\right)^T (x - z)$$

  taking the limit for $t \to 0$ gives

  $$F(x)^T (z - x) \geq 0$$

Analytic centering cutting-plane method

to generate cutting-plane at x

- if $x \notin C$, use feasibility cut (cutting-plane that separates x from C)
- if $x \in C$ and not a solution, use the cutting-plane

$$F(x)^T(z-x) \le 0$$

proof: if $F(x)^T(z-x) > 0$ then, by monotonicity,

$$F(z)^T(z-x) \ge F(x)^T(z-x) > 0$$

therefore z is not a solution of the variational inequality

Extragradient method

choose $x^{(0)} \in C$
repeat for $k = 1, 2, \ldots$

$$\begin{aligned} y^{(k)} &= P_C\left(x^{(k-1)} - t_k F(x^{(k-1)})\right) \\ x^{(k)} &= P_C\left(x^{(k-1)} - t_k F(y^{(k)})\right) \end{aligned}$$

- step size t_k is fixed ($e.g.$, $t_k = 1/L$) or determined by line search
- a 2-step variation of the projection algorithm

$$x^{(k)} = P_C\left(x^{(k-1)} - t_k F(x^{(k-1)})\right)$$

Assumptions

in the analysis of the extragradient method we assume:

- C is bounded
- F is monotone

$$(F(x) - F(y))^T (x - y) \geq 0 \quad \forall x, y$$

- F is Lipschitz continuous

$$\|F(x) - F(y)\|_2 \leq L\|x - y\|_2 \quad \forall x, y$$

Analysis

denote $y = y^{(i)}$, $x = x^{(i-1)}$, $x^+ = x^{(i)}$, $t = t_i$

1. x^+ is Euclidean projection of $x - tF(y)$; hence for all $z \in C$

$$\left(x^+ - x + tF(y)\right)^T (z - x^+) \geq 0$$

   equivalently

$$2tF(y)^T(z - x^+) + \|z - x\|_2^2 \geq \|x^+ - x\|_2^2 + \|z - x^+\|_2^2 \quad \forall z \in C$$

2. similarly, since y is the Euclidean projection of $x - tF(x)$ on C

$$2tF(x)^T(z - y) + \|z - x\|_2^2 \geq \|y - x\|_2^2 + \|z - y\|_2^2$$

3. apply 2. to $z = x^+$

$$\begin{aligned} &2tF(y)^T(x^+ - y) + \|x^+ - x\|_2^2 \\ &\quad = \quad 2tF(x)^T(x^+ - y) + \|x^+ - x\|_2^2 + 2t(F(y) - F(x))^T(x^+ - y) \\ &\quad \geq \quad \|y - x\|_2^2 + \|x^+ - y\|_2^2 + 2t\left(F(y) - F(x)\right)^T\left(x^+ - y\right) \end{aligned}$$

4. if $0 < t \leq 1/L$, we can use the Lipschitz continuity of F to show

$$\begin{aligned} &2tF(y)^T(x^+ - y) + \|x^+ - x\|_2^2 \\ &\quad \geq \quad \|y - x\|_2^2 + \|x^+ - y\|_2^2 - 2tL\|y - x\|_2\,\|x^+ - y\|_2 \\ &\quad \geq \quad 0 \end{aligned}$$

5. combine this with the inequality in 1. and use monotonicity of F

$$\begin{aligned} 2tF(z)^T(z - y) + \|z - x\|_2^2 \quad &\geq \quad 2tF(y)^T(z - y) + \|z - x\|_2^2 \\ &\geq \quad \|z - x^+\|_2^2 \quad \forall z \in C \end{aligned}$$

conclusion

- if $0 < t_i \leq 1/L$, then for all $z \in C$

$$t_i F(z)^T(z - y^{(i)}) + \frac{1}{2}\|z - x^{(i-1)}\|_2^2 - \frac{1}{2}\|z - x^{(i)}\|_2^2 \geq 0$$

- add the inequalities for $i = 1, \ldots, k$, and divide by $s_k = \sum_{i=1}^k t_i$

$$F(z)^T(\bar{y}^{(k)} - z) \leq \frac{1}{2s_k}\|z - x^{(0)}\|_2^2 \qquad \forall z \in C$$

where

$$\bar{y}^{(k)} = \frac{1}{s_k}\sum_{i=1}^k t_i y^{(i)}$$

Iteration complexity for bounded C

assume C is bounded, with

$$\sup_{z \in C} \|z - x^{(0)}\|_2^2 \leq 2R$$

- if $0 < t_i \leq 1/L$ for $i = 1, \ldots, k$,

$$\eta(\bar{y}^{(k)}) = \sup_{z \in C} F(z)^T (\bar{y}^{(k)} - z) \leq \frac{R}{s_k}$$

- if $t_i = 1/L$, $i = 1, \ldots, L$

$$\eta(\bar{y}^{(k)}) \leq \frac{LR}{k}$$

$O(L/\epsilon)$ iterations to reduce duality gap function below ϵ

Line search

- choose $t_0 > 0$
- at iteration k, set $t_k := t_{k-1}$ and backtrack ($t_k := \beta t_k$) until

$$2t_k F(y^{(k)})^T (x^{(k)} - y^{(k)}) + \|x^{(k)} - x^{(k-1)}\|_2^2 \geq 0$$

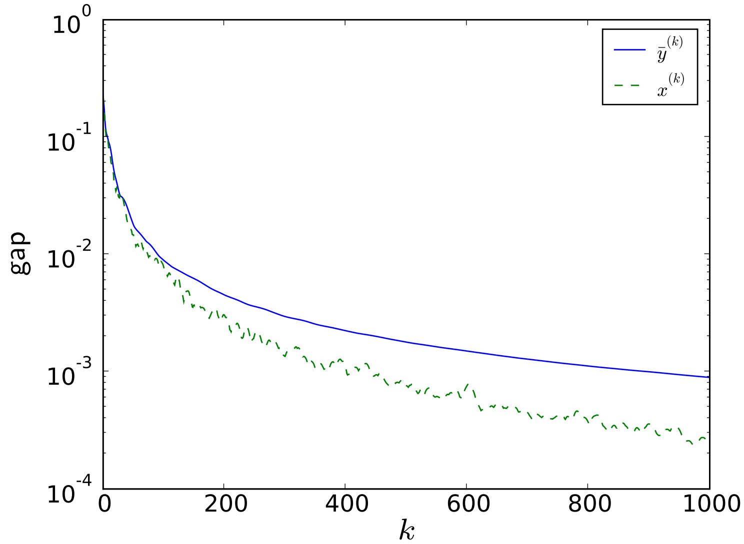
($i.e.$, the inequality in 4. on page 16–19 holds)

consequences for analysis

- step sizes t_k are bounded below by $t_k \geq \min\{t_0, \beta/L\}$ (see p.16–19)
- conclusion on 16–20 still holds:

$$\eta(\bar{y}^{(k)}) \leq \frac{R}{s_k} \leq \frac{R}{k \min\{t_0, \beta/L\}}$$

Matrix game



$A \in \mathbf{R}^{500\times 500}$; fixed step size $t = 1/L = 1/\|A\|_2$

gap is $\max_i(A^T u)_i - \min_i(Av)_i$ for $(u, v) = \bar{y}^{(k)}$ and $(u, v) = x^{(k)}$

Nuclear norm approximation

$$\begin{array}{ll} \text{minimize} & \|A(u) + B\|_* \\ \text{subject to} & \|u\|_\infty \leq 1 \end{array}$$

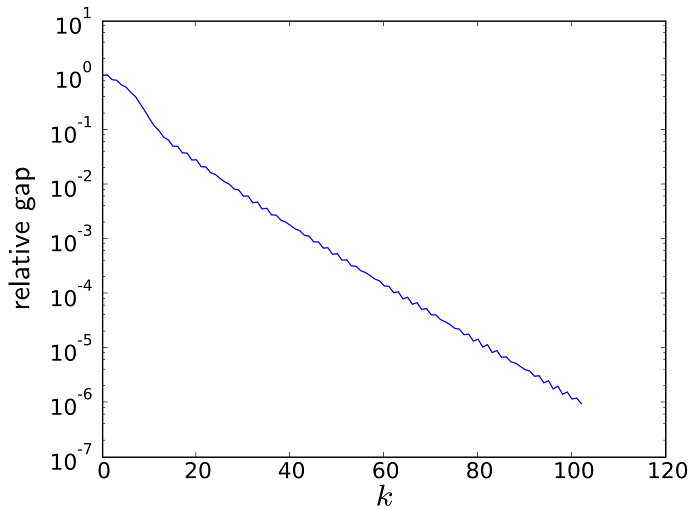
- $A(u) = u_1 A_1 + \cdots + u_n A_n \in \mathbf{R}^{p\times q}$, $B \in \mathbf{R}^{p\times q}$
- $\|\cdot\|_*$ is sum of singular values

minimax formulation

$$\begin{array}{ll} \text{minimize} & \sup_{\|V\|_2\leq 1} \mathbf{tr}(V^T(A(u) - B)) \\ \text{subject to} & \|u\|_\infty \leq 1 \end{array}$$

- $\|\cdot\|_2$ is maximum singular value
- $f(u, V) = \mathbf{tr}(V^T(A(u) + B))$ is convex-concave (bilinear)

example ($p = 1000$, $q = 100$, $n = 500$)



- gap is $\|A(u) - B\|_* - \mathbf{tr}(B^T V) + \sum_i |\mathbf{tr}(A_i^T V)|$ at $(u, V) = x^{(k)}$
- relative gap is gap divided by $\|A(u) - B\|_*$
- one iteration requires two SVDs for projections on $\{Z \mid \|Z\|_2 \le 1\}$

References

- P. Tseng, *On accelerated proximal gradient methods for convex-concave optimization* (2008)

  §5 was used for the analysis of the extragradient method

- A. Nemirovski, *Prox-method with rate of convergence $O(1/t)$ for variational inequalities with Lipschitz continuous monotone operators and smooth convex-concave saddle point problems*, SIAM J. Optim. (2004)