Perturbed Fejer processes

E.A. Nurminski

Institute for Automation and Control Processes, Vladivostok
nurmi@dvo.ru

December 2008
Universita della Calabria, Cosenza

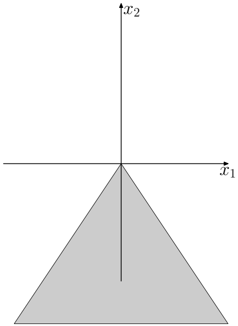
Outline

1. Motivations
   - Convex feasibility problem
   - Convex optimization problem
2. Fejer processes
   - Fejer processes with arbitrary perturbations
   - Fejer processes with attractants
   - Convergence theory
3. Decomposition and parallel computations
   - Sequential projection

Very simple CFP

Find a point in the set

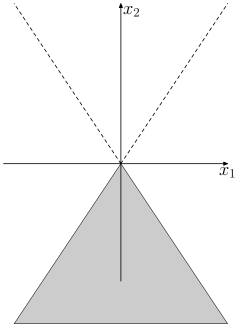
$$2x_1 + x_2 \leq 0$$
$$-2x_1 + x_2 \leq 0$$



Very simple CFP, $X = X_1 \cap X_2$.

Find a point in the set

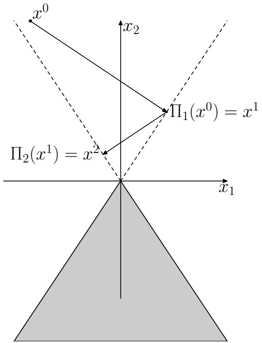
$$2x_1 + x_2 \leq 0$$
$$-2x_1 + x_2 \leq 0$$



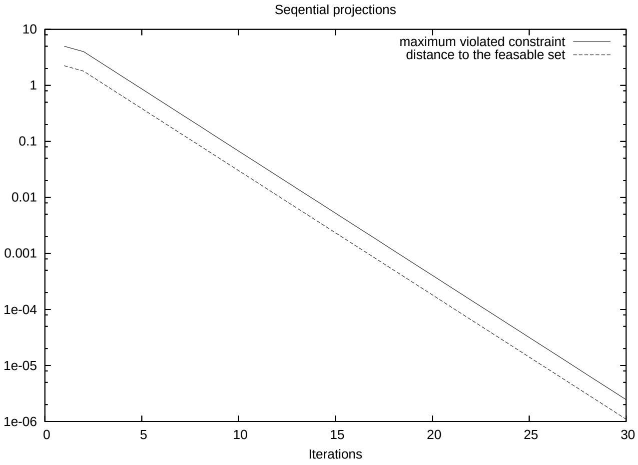
Very simple CFP, $X = X_1 \cap X_2$, sequential projection.

Find a point in the set

$$2x_1 + x_2 \leq 0$$
$$-2x_1 + x_2 \leq 0$$



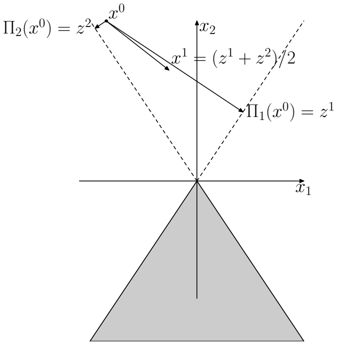
Very simple CFP — sequential projection



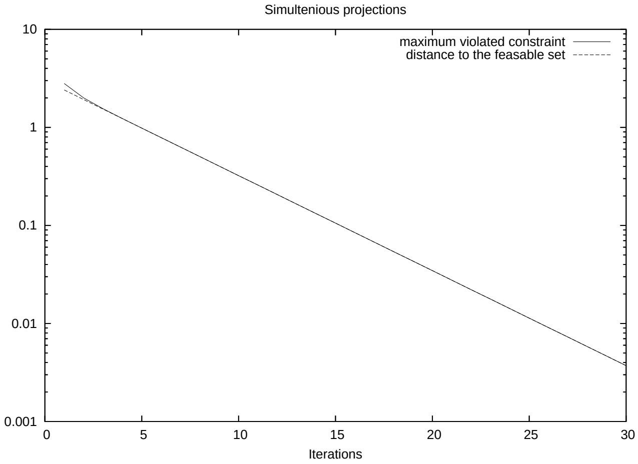
Very simple CFP — simultenious projection

Find a point in the set

$$2x_1 + x_2 \leq 0$$
$$-2x_1 + x_2 \leq 0$$



Very simple CFP — simultenious projection



Projection in optimization and related subjects

Projective equations:

$$x = \Pi_X(x - \lambda G(x)), \quad \lambda > 0$$

G — (sub)gradient field, variational ineqaulity operator, ...

Projection in optimization and related subjects

Projective equations:

$$x = \Pi_X(x - \lambda G(x)), \quad \lambda > 0$$

Simple iteration:

$$x^{k+1} = \Pi_X(x^k - \lambda G(x^k)), \quad \lambda \in (0, \tau),$$

Projection in optimization and related subjects

Projective equations:

$$x = \Pi_X(x - \lambda G(x)), \quad \lambda > 0$$

Simple iteration:

$$x^{k+1} = \Pi_X(x^k - \lambda G(x^k)), \quad \lambda \in (0, \tau),$$

Disadvantages:

1. $G(x)$ needs to be strongly monotone;

Projection in optimization and related subjects

Projective equations:

$$x = \Pi_X(x - \lambda G(x)), \quad \lambda > 0$$

Simple iteration:

$$x^{k+1} = \Pi_X(x^k - \lambda G(x^k)), \quad \lambda \in (0, \tau),$$

Disadvantages:

1. $G(x)$ needs to be strongly monotone;
2. difficult to implement for nontrivial X;

Projection in optimization and related subjects

Projective equations:

$$x = \Pi_X(x - \lambda G(x)), \quad \lambda > 0$$

Simple iteration:

$$x^{k+1} = \Pi_X(x^k - \lambda G(x^k)), \quad \lambda \in (0, \tau),$$

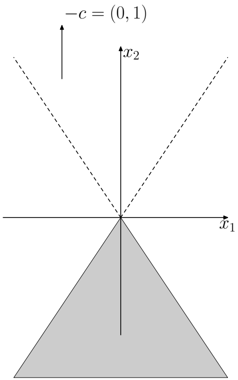
Disadvantages:

1. $G(x)$ needs to be strongly monotone;
2. difficult to implement for nontrivial X;
3. low rate of convergence.

Very simple COP

Solve the optimization problem:

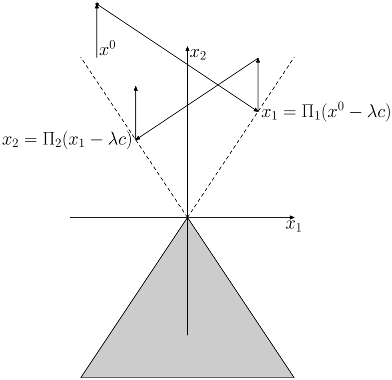
$$\min(-x_2) = \min\ cx$$
$$2x_1 + x_2 \leq 0$$
$$-2x_1 + x_2 \leq 0$$



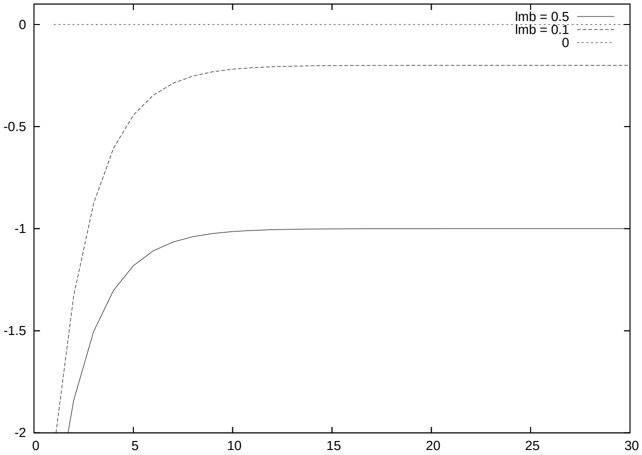
Very simple COP — sequential projection

Solve the optimization problem:

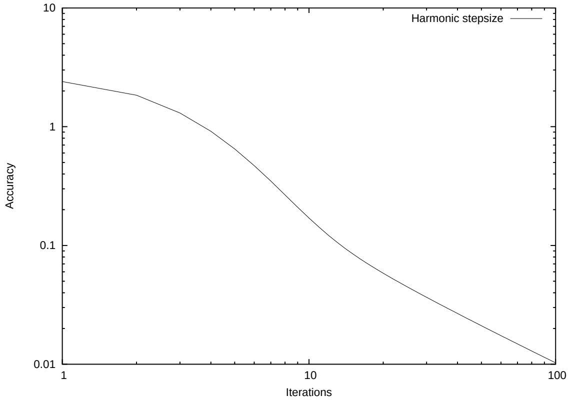
$$\min(-x_2) = \min cx$$
$$2x_1 + x_2 \leq 0$$
$$-2x_1 + x_2 \leq 0$$



Very simple COP — sequential projection



Very simple COP — harmonic stepsize

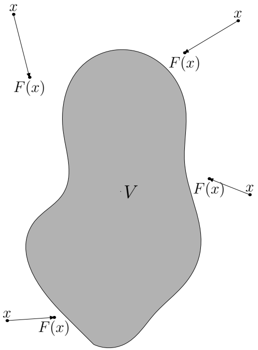


$$x^{k+1} = \Pi_k(x^k - \lambda_k c), k = 0, 1, \ldots, \lambda_k = c/k.$$

Fejer operators

Definition. *An operator F will be called Fejer if for any x $\|F(x) - v\| \leq \|x - v\|$ for all $v \in V, x \in \bar{x} + U$.*

Definition. *Fejer operator F will be called locally strong if for any $\bar{x} \notin V$ there exists a neighborhood of zero U and small enough $\alpha \in [0, 1)$ such that $\|F(x) - v\| \leq \alpha \|x - v\|$ for all $v \in V, x \in \bar{x} + U$.*



Fejer processes

$$x^{k+1} = F(x^k), k = 0, 1, \dots \quad (1)$$

where F is a Fejer operator of any kind.

Theorem. *Let V — closed and bounded, F — locally strong Fejer, and sequence $\{x^k\}$, obtained by (1) with some arbitrary x^0, bounded. Then all limit points of $\{x^k\}$ belong to V.*

Perturbed Fejer processes

Fejer process with small perturbations:

$$x^{k+1} = F(x^k + z^k), k = 0, 1, \ldots \tag{2}$$

Theorem. *Let V — closed and bounded, F — locally strong Fejer, the sequence $\{x^k\}$, obtained by (1) with arbitrary x^0, is bounded, $z^k \to 0$ when $k \to \infty$. Then all limit points $\{x^k\}$ belong to V.*

Collections of Fejer operators

Theorem. *Let $\Phi = \{F_1, F_2, \dots, F_m\}$ is a finite collection of operators F_i such that for any $x \notin V$ there exists F_i locally strong Fejer at x, $z^k \to 0$ when $k \to \infty$ and $\mathcal{F}_k = F_{i_k}$, where F_{i_k} — locally strong Fejer at x^k. If the sequence $\{x^k\}$, constructed by*

$$x^{k+1} = \mathcal{F}_k(x^k + z^k), \quad s = 0, 1, \dots \tag{3}$$

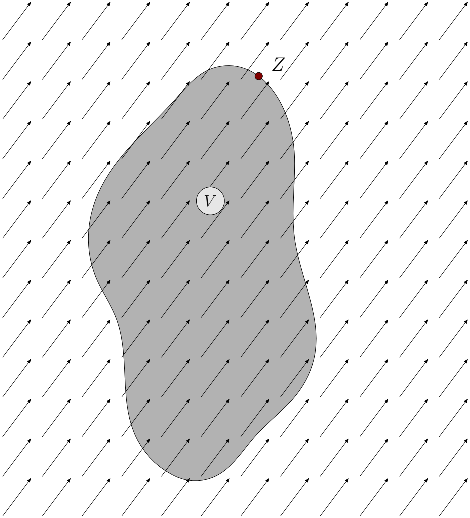
is bounded then all its limit poins belong to V.

Question: What about *infinite* families ?

Attractants

Definition. *Point-to-set mapping* $G : V \to 2^E$ *is callled a locally strong attractant (of some* $Z \subset V$ *) if for any* $x' \in V \setminus Z$ *there is a neighborhood of zero* U *such that* $g(z - x) \geq \delta > 0$ *for all* $z \in Z, x \in x' + U, g \in \Phi(x)$ *and some* $\delta > 0$.

An Attractant vector field



Fejer processes with attractants

Stationary:

$$x^{k+1} = F(x^k + \lambda_k g^k), \; g^k \in G(x^k). \tag{4}$$

Nonstationary:

$$x^{k+1} = F_k(x^k + \lambda_k g^k), \; g^k \in G(x^k). \tag{5}$$

It follows from above that (4) as well as (5) converge to V if $\lambda_k \to 0$ when $k \to \infty$.

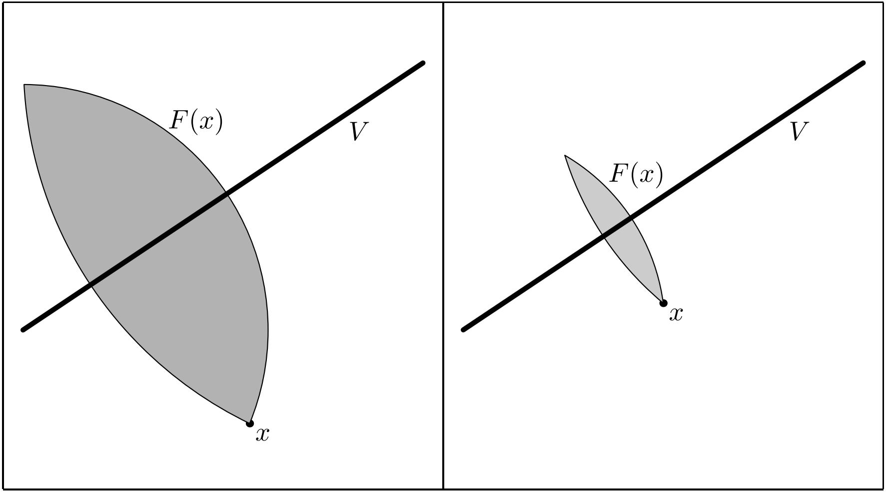
Convergence to Z

Theorem. *Let F is a locally strong Fejer operator, G — locally strong attractant $Z \subset V$, upper semicontinuous on some open $\tilde{V} \supset V$ and sequence $\{x^k\}$, obtained by*

$$x^{k+1} = F(x^k + \lambda_k g^k), \; g^k \in G(x^k), \tag{6}$$

where initial state x^0 arbitrary, $\lambda_k \to +0, \sum \lambda_k = \infty$. If $\{x^k\}$ bounded then any limit point $\{x^k\}$ belongs to Z.

Fejer projective operators



Точка x "far" from V .   Точка x "nearby" V .

Sequential projection

Let

$$V = \cap_{\tau \in T} V_\tau,$$

$V_\tau, \tau \in T$ — convex closed subsets of E.

Theorem. *Let V — closed bounded set, which can be represented as an intersection of a finite family of convex sets $V = \cap_{\tau \in T} V_\tau$ and denote as $\Pi_\tau(x) = x_\tau$ the orthogonal projection of x onto V_τ. If $x \notin V_{\tau'}$ for some $\tau' \in T$, then the operator $F = \Pi_{\tau'}$ is locally strong Fejer at x.*

Sequential projection gradient method

The problem:

$$\min_{x \in V} f(x), V = \cap_{i=1}^{N} V_i.$$

Sequential projection gradient method:

$$x^{k+1} = F_k(x^k - \lambda_k g^k), \; g^k \in \partial f(x^k)$$

где $F_k(x) = \Pi_{i_k}(x)$, а i_k такого, что $x^k \notin V_{i_k}$.
General theory asks for

$$\lambda_k \to +0, \quad \sum_{i=1}^{\infty} \lambda_k = \infty.$$

Can we do better ?

Envelope stepsize control (ESC)

Algorithm model:

$$x^{k+1} = x^k - \lambda_k d^k,\ \ d^k \in D(x^k),$$

$D(x)$ — usc set-valued mapping. Let $D(p, q) = \mathrm{co}\,\{d^t, p < t \leq q\}$. Given

$$0 < \theta_m \to +0, m = 0, 1, \dots \text{ and } q \in (0, 1)$$

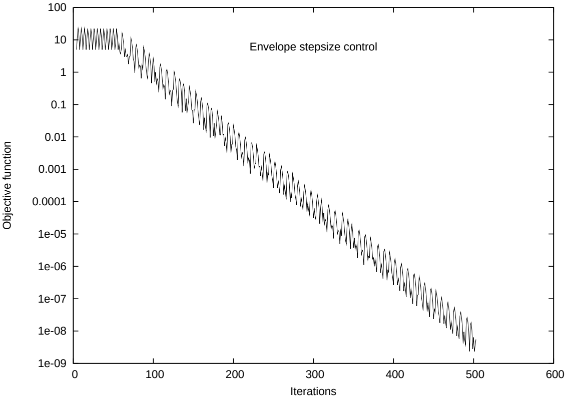
define $\{k_m\}$ and stepsizes $\{\lambda_k\}$ as follows:

- Set $k_0 = 0$ and pick up initial $\lambda_0 > 0$.
- For given m and k_m determine k_{m+1} as the index which satisfies conditions

$$0 \notin D(k_m, s) + \theta_m B, k_m \leq s < k_{m+1},$$
$$0 \in D(k_m, k_{m+1}) + \theta_m B$$

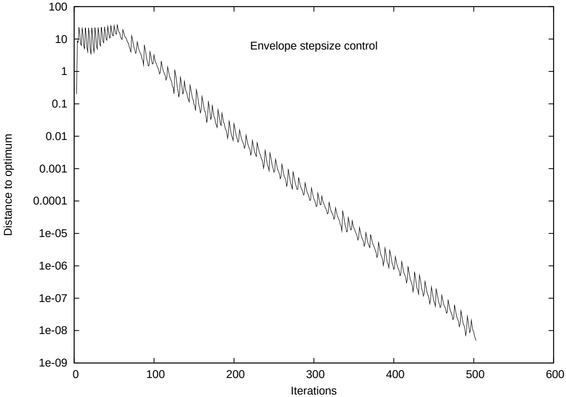
with $\lambda_k = \lambda_{k_m}$. Set $\lambda_{k_{m+1}} = q\lambda_{k_m}$.

Another very simple COP — ESC stepsize



Objective function convergence.

Another very simple COP — ESC stepsize



Distance to optimum convergence.

Summary

- Diminishing additive disturbances in arguments of Fejer operators does not prevent convergence of locally strong Fejer processes.
- Using attractants one can direct Fejer processes to a specific part of attracting set.
- Sequential and simultenious projections are Fejer and can be used to decompose/parallelize projective optimization algorithms.
- It looks like that it is possible to have linear-like convergence, but to prove it we need better convergence theory.