Envelope stepsize control in Fejer algorithms

E. A. Nurminski[*], N. B. Shamray[†]

[*]Institute for Automation and Control Processes FEB RAS, nurmi@dvo.ru
[†]Institute for Automation and Control Processes FEB RAS, shamray@dvo.ru

Fejer processes can be defined as iterative algorithms of the kind

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

where F is a Fejer operator. They are often used as models for many iterative algorithms in optimization and related areas and can be combined with many kinds of decomposition schemes and generate projection-type methods suitable for parallel computations [1].

To guarantee the convergence of such algorithms we define a strong version of Fejer process.

Definition 1 *An operator F is called locally strong Fejer (with respect to a certain set V) if for any $\bar{x} \notin V$ there exists a neighborhood of zero U and $\alpha < 1$ such that $\|F(x) - v\| \leq \alpha \|x - v\|$ for any $v \in V$ and $x \in \bar{x} + U$.*

A slightly more general model of Fejer-type algorithms may look like following

$$x^{k+1} = F_k(x^k + z^k), k = 0, 1, \dots, \tag{1}$$

where z^k is a diminishing ($z^k \to 0$) disturbance and F_k is selected from some finite collection

$$\mathcal{F} = \{\mathcal{F}_i, i = 1, 2, \dots, N\} \tag{2}$$

of locally strong Fejer operators.

We are especially interested in the case when such collection is associated with representation of V as the intersection of $V_i,\ i = 1, 2, \dots, N$:

$$V = \bigcap_{i=1}^{N} V_i \tag{3}$$

and each ϕ_i is a locally strong Fejer operator for V_i. If disturbances are such that $z^k = \lambda_k g^k$, where λ_k are scalar multipliers, and $g^k \in G(x^k)$ are selections for set-valued attractant mapping (see definition below) then is is possible to prove that (1) converges to certain subset of V which may be a solution of optimization problem, variational inequality [2] etc.

Definition 2 *A set-valued mapping $G(x)$ is called strong locally restricted attractant (of some set $Z \subset V$) if for each $x' \in V \setminus Z$ there exists a neighborhood of zero U such that,*

$$g(z - x) \geq \delta > 0$$

for all $z \in Z,\ x \in x' + U,\ g \in G(x)$ and some $\delta > 0$.

The convergence of (1) is established by the following theorem, see [3].

Theorem 1 *Let $\mathcal{F}$ is a finite family of locally strong with respect to corresponding V_i continuous Fejer operators, and for any $x \notin V = \cup_{i=1}^{N}$ there exists $\iota \in \{1, 2, \dots, N\}$ such that $\mathcal{F}_\iota$ is locally strong at x and $D(\cdot)$ is strong locally restricted attractant of $Z \subset V$. Then the combined process*

$$x^{k+1} = F_k(x^k + \lambda_k d^k), \tag{4}$$

where $d^k \in D(x^k)$, $F_k = \mathcal{F}_{\iota_k}$ and ι_k is such that $x^k + \lambda_k d^k \notin V_{i_k}$, if bounded, converges to the set Z if $\lambda_k \to +0$ and $\sum \lambda_k = \infty$.

Theorem 1 opens many new possibilities for new algorithms of constrained convex optimization and not only, but the diverging series condition for stepsize λ_k used in this theorem is known to result in slow convergence. Therefore it is of theoretical as well as practical interests to search for other stepsize control rules with established convergence and better computational performance. In this report the adaptive stepsize rule for (4) is suggested and its theoretical convergence is established. The main idea of this stepsize control consist in keeping stepsize constant until the process shows the signs of cycling and then divide stepsize by some constant factor. To be precise this rule is formulated as follows.

Let $D(p, q) = \operatorname{co}\{x^{k+1} - x^k,\ k = p, p+1, \dots, q-1\}$. For a given sequence $\theta_m \to +0,\ m = 0, 1, \dots$ determine corresponding sequences of indeces $\{k_m\}$ and numbers $\{\lambda_k\}$ by the following recursive relationships:

1. Set $k_0 = 0$ and pick up initial $\lambda_0 > 0$. Let $q \in (0, 1)$.

2. For given m and k_m determine k_{m+1} as the index which satisfies conditions

$$0 \notin D(k_m, k) + \theta_m B, k_m \leq k < k_{m+1}, \\ 0 \in D(k_m, k_{m+1}) + \theta_m B \tag{5}$$

   with $\lambda_k = \lambda_{k_m}$ for $k_m \leq k < k_{m+1}$. Set

$$\lambda_{k_{m+1}} = q\lambda_{k_m}. \tag{6}$$

The following theorem establishes convergence of the process (1) with the stepsize rule (5) - (6). Denote $X_\star = \{x^\star : 0 \in G(x^\star)\}$. The following theorem holds.

Theorem 2 *Let $G(x)$ is convex-valued, locally bounded upper-semi-continuous set-valued locally strong attractant of $X_\star$, and $x^{k+1} - x^k \in \lambda_k G(x^k)$. Then if the sequence $\{x^k\}$ is bounded then all its limit points belong to $X_\star$.*

Numerical experiments demonstrated quite satisfactory computational performance of this method with nonmonotone but close to linear rate of convergence.

References

[1] V. V. Vasin and I. I. Eremin, *Operators and Fejer Iterative Processes: Theory and Applications*. Ural. Otd. Ross. Acad. Nauk, Yekaterinburg, 2005. [in Russian].

[2] N. B. Shamray, *A sequential projection method for variational inequalities* // Zhurn. Vychisl. Mathem. Matem. Physiki — submitted to.

[3] E. A. Nurminski, *The use of additional small disturbances in Fejer models of iterative algorithms* // Zhurn. Vychisl. Mathem. Matem. Physiki, 2008, v. 48, issue 12, pp. 2121-2128 [In Russian].