Differential Equations, Vol.31, No.11, 1995, pp. 1754–1763

ON DIFFERENTIAL PREDICTION-TYPE GRADIENT METHODS FOR COMPUTING FIXED POINTS OF EXTREMAL MAPPINGS [1]

A.S. Antipin

UDC 517.977

(Revised version 1 December 2003)

1. STATEMENT OF THE PROBLEM

Consider the problem of computing a fixed point of the extremal mapping

$$v^* \in \operatorname{Argmin}\{\Phi(v^*, w) \mid w \in \Omega\}, \quad v^* \in \Omega^*, \tag{1.1}$$

where the function $\Phi(v, w)$ is defined on the product $\mathbb{R}^n \times \mathbb{R}^n$ and $\Omega \subset \mathbb{R}^n$. We assume that $\Phi(v, w)$ is convex with respect to the variable $w \in \Omega$ for each chosen $v \in \Omega$, the extremal (marginal) mapping $G(v) \equiv \operatorname{argmin}\{\Phi(v, w) \mid w \in \Omega\}$ is defined for all $v \in \Omega$, and the set $\Omega^* \subset \Omega$ of solutions to the original problem is nonempty. By (1.1), each point of Ω^* satisfies the inequality

$$\Phi(v^*, v^*) \le \Phi(v^*, w) \quad \forall w \in \Omega, \quad v^* \in \Omega^*. \tag{1.2}$$

If the function $\Phi(v, w)$ is differentiable with respect to the second variable, then (1.2) is equivalent to the variational inequality [1]

$$\langle \nabla\Phi_w(v^*, v^*), w - v^* \rangle \ge 0 \quad \forall w \in \Omega, \quad v^* \in \Omega^*. \tag{1.3}$$

where $\nabla\Phi_w(v, w)$ is the vector gradient of the function with respect to the second variable.

Inequality (1.2) is related to the existence of a fixed point and has not relation to the stability of the point, that is, the possibility to get into a small neighborhood of this point by some method. To specify the stability of the point, we suggest another inequality [2, 3], namely,

$$\Phi(w, v^*) \le \Phi(w, w) \quad \forall w \in \Omega, \quad v^* \in \Omega^*, \tag{1.4}$$

which, together with (1.2), generalizes the notion of a saddle point [4, 5]. Indeed, if $\Phi(w, w) \equiv \equiv \Phi(v^*, v^*) = \mathit{const}$ for all $w \in \Omega$, i.e., if the function $\Phi(v, w)$ is identically equal to the constant $\Phi(v^*, v^*)$ on the diagonal of the square $\Omega \times \Omega$, then inequality (1.4) becomes

$$\Phi(w, v^*) \le \Phi(v^*, v^*) \quad \forall w \in \Omega, \quad v^* \in \Omega^*. \tag{1.5}$$

The system of inequalities (1.2) and (1.5) defines a saddle point, in addition, its first component (as well as the second one) is a fixed point.

Inequality (1.4) can be considered as a consequence of the more restrictive inequality

$$\Phi(w, w) - \Phi(w, v^*) - \Phi(v^*, w) + \Phi(v^*, v^*) \ge 0 \quad \forall w \in \Omega, \quad v^* \in \Omega^*, \tag{1.6}$$

for all $w \in \Omega$ and all $w^* \in \Omega^*$. Indeed, by Eqs. (1.2), relation (1.4) directly follows from (1.6).

The aim of the present paper is to prove the convergence of a predictive gradient method to a solution to problem (1.1) and to obtain estimates for the rate of convergence of this method under conditions (1.4) and (1.6) and their various modifications.

[1] Computational Center of Russian Academy of Sciences, Moscow. Translated from *Differentsial'nye Uravneniya*, 31, No.11, pp. 1786–1795, November, 1995. Original article submitted June 30, 1995.

2. PROBLEMS WITH SKEW-SYMMETRIC GOAL FUNCTION

Below we show that inequalities (1.4) and (1.6) allow one to establish convergence properties of the considered processes. However, these inequalities are not constructive, since their formal definition contains the unknown vector v^*. Therefore, the validity of inequality (1.4) or (1.6) cannot be tested for a given particular problem. Hence, we face an important problem of describing classes of functions $\Phi(v, w)$ for which the above inequalities automatically hold. Such classes of functions exist. Let us describe one of these classes. Consider the class of functions such that

$$\Phi(w, v) + \Phi(v, w) \le 0 \quad \forall w \in \Omega, \quad \forall v \in \Omega. \tag{2.1}$$

Recall that the points (v, w) and (w, v) are symmetric with respect to the diagonal $w = v$ of the square. We especially note the case in which (2.1) becomes an equality:

$$\Phi(w, v) + \Phi(v, w) = 0 \quad \forall w \in \Omega, \quad \forall v \in \Omega. \tag{2.2}$$

For $w = v$, it follows from Eq. (2.2) that $\Phi(w, w) = 0$ on the diagonal of the square. If $\Phi(v, w)$ is defined on a finite set of values v_i, w_j, $i = 1, \dots, n$, $j = 1, \dots, m$, then it can be considered as the matrix $\Phi_{i,j}$; in this case, relation (2.2) is reduced to the well-known definition of the skew-symmetric matrix; $\Phi_{i,j} + \Phi_{j,i} = 0$ for all i, j. It is natural to call a function $\Phi(v, w)$ with property (2.2) an *skew-symmetric* function. Similarly, using the condition

$$\Phi(w, v) - \Phi(v, w) = 0 \quad \forall w \in \Omega, \quad \forall v \in \Omega \tag{2.3}$$

one can introduce the notion of a symmetric function $\Phi(v, w)$. If to the point (v, w) we assign the value of the function $\Phi(\cdot, \cdot)$ at the point (w, v), then we obtain the function that can be called the *transposed* function $\Phi^\top(v, w)$. In terms of this function, conditions (2.2) and (2.3) become $\Phi(v, w) = -\Phi^\top(v, w)$ and $\Phi(v, w) = \Phi^\top(v, w)$, respectively. We can readily verify that any real function $\Phi(v, w)$ can be represented in the form of the sum $\Phi(v, w) = \Psi_1(v, w) + \Psi_2(v, w)$, where the function $\Psi_1(v, w)$ is symmetric and the function $\Psi_2(v, w)$ is skew-symmetric. This expansion is unique, and we have $\Psi_1(v, w) = 0.5(\Phi(v, w) + \Phi^\top(v, w))$ and $\Psi_2(v, w) = 0.5(\Phi(v, w) - \Phi^\top(v, w))$.

In the present paper we restrict ourselves to the case of an skew-symmetric function $\Phi(v, w)$. We shall consider the condition of being skew-symmetric in a somewhat more general form than (2.1), namely,

$$\Phi(w, w) - \Phi(w, v) - \Phi(v, w) + \Phi(v, v) \ge 0 \quad \forall w \in \Omega, \quad \forall v \in \Omega. \tag{2.4}$$

If $\Phi(w, w) = 0$, then (2.4) reduces to (2.1).

The foregoing consideration shows that the notion of skew-symmetry covers equilibrium problems and equilibrium solutions that inherit properties of optimization problems and saddle problems.

Inequality (2.4) holds for all $v \in \Omega$ and $w \in \Omega$, in particular, for $v = v^* \in \Omega^*$. By taking into account (1.2), relation (2.4) immediately implies (1.4). Thus, if the goal function of problem (1.1) have the skew-symmetry property (2.4), then the equilibrium solution to problem (1.1) satisfies condition (1.4).

Let us show that a normalized function of the saddle problem is skew-symmetric. The saddle problem is to solve the system of inequalities

$$L(x^*, p) \le L(x^*, p^*) \le L(x, p^*), \qquad x \in Q \subset R^n, \quad p \in P \subset R^m, \tag{2.5}$$

where the function $L(x, p)$ is convex with respect to x and concave with respect to p. Applying the normalized function of the form $\Phi(v, w) = L(z, p) - L(x, y)$ [6], $w = (z, y)$, $v = (x, p)$, we can reduce system (2.5) to the computation of a fixed point of extremal inclusion (1.1). Since the function $\Phi(v, w)$ is separable with respect to the variables z and y and the set $\Omega = Q \times P$ has block structure, problem (1.1) is equivalent to problem (2.4), and the sets of solutions to these problems coincide.

Let us verify that the normalized function $\Phi(v, w)$ of saddle problem (2.5) satisfies the following conditions [4, 5]:

$$\Phi(w, w) = 0 \quad \forall w \in \Omega, \tag{2.6}$$

$$\Phi(w, v) + \Phi(v, w) = 0 \quad \forall w \in \Omega, \quad \forall v \in \Omega. \tag{2.7}$$

The first property means that the function $\Phi(v, w)$ vanishes on the diagonal of the square, i.e., for $v = w$. This clearly holds for the function $\Phi(v, w) = L(z, p) - L(x, y)$, where $w = (z, y)$ and $v = (x, p)$, because $\Phi(w, w) = L(z, y) - L(z, y) = 0$ for $v = w$.

We can readily verify the second property as well. Indeed, let $\Phi(v, w) = L(z, p) - L(x, y)$. Since the domain of the variables $w \in \Omega$ and $v \in \Omega$ is the same, we can set $v = w$ and $w = v$ in $\Phi(v, w)$; then $\Phi(w, v) = L(x, y) - L(z, p)$. Therefore, $\Phi(w, v) + \Phi(v, w) = L(x, y) - L(z, p) + L(z, p) - L(x, y) = 0$.

Summing (2.6) and (2.7) we obtain

$$\Phi(w, w) - \Phi(w, v) - \Phi(v, w) + \Phi(v, v) = 0 \quad \forall w \in \Omega, \quad \forall v \in \Omega. \tag{2.8}$$

According to (1.2), for $v = v^*$, relation (2.8) immediately implies (1.4). Since $\Phi(w, w) = 0$ and $\Phi(v^*, v^*) = 0$, we have

$$\Phi(w, v^*) \le \Phi(v^*, v^*) \le \Phi(v^*, w), \qquad w \in \Omega, \quad v^* \in \Omega^*. \tag{2.9}$$

Thus, (v^*, v^*) is a saddle point of the normalized function $\Phi(v, w)$, and $\Phi(v, v) = 0$, where v is a fixed point. The left inequality in system (2.9) coincides with condition (1.4).

Equations (2.6) and (2.7) can be generalized up to the inequalities

$$\Phi(w, w) \ge 0 \quad \forall w \in \Omega. \tag{2.10}$$

$$\Phi(w, v) + \Phi(v, w) \le 0 \quad \forall w \in \Omega, \quad \forall v \in \Omega. \tag{2.11}$$

This generalization extends the class of equilibrium problems under consideration. Note that the system of inequalities (2.10), (2.11) was used in [7] to prove the Ky Fan inequality, which is equivalent to the Kakutani fixed point theorem for a continuous point-to-set mapping on a bounded closed convex set [8].

If the function $\Phi(v, w)$ is differentiate with respect to the second variable, then, by the convexity inequalities

$$\langle \nabla f(x), y - x \rangle \le f(y) - f(x) \le \langle \nabla f(y), y - x \rangle, \tag{2.12}$$

for all y and x from some set, relation (1.6) yields the inequality

$$\langle \nabla \Phi_w(w, w) - \nabla \Phi_w(v^*, v^*), w - v^* \rangle \ge 0 \quad \forall w \in \Omega, \quad v^* \in \Omega^*,$$

which means that the partial gradient $\nabla \Phi_w(v, w)$ is a monotone operator with respect to the equilibrium v^*.

There exist classes of equilibrium problems for which inequalities (1.4) and (1.6) can be strengthened. Such problems are similar to optimization problems with strongly convex goal functions, and their solutions satisfy the following inequalities [2, 3]:

1) if the solution to the problem is unique, then

$$\Phi(w, w) - \Phi(w, v^*) \ge \gamma |w - v^*|^{1+\nu} \quad \forall w \in \Omega, \tag{2.13}$$

the constant γ is positive and the parameter ν belongs to $[0, \infty]$; for $\nu = 0$ we have a sharp isolated equilibrium;

2) if the solutions to the problem form a closed connected set, then

$$\Phi(w, w) - \Phi(w, \pi_{\Omega^*}(w)) \geq \gamma |w - \pi_{\Omega^*}(w)|^{1+\nu} \quad \forall w \in \Omega, \tag{2.14}$$

$\pi_{\Omega^*}(w)$ is the projection operator of the vector w to the set Ω^*. If the set Ω^* is a singleton, then (2.14) becomes (2.13). We also note that (2.14) implies (1.4).

An analog of Eq. (1.6) has the form

$$\begin{aligned} (\Phi(w, w) - \Phi(w, \pi_{\Omega^*}(w))) \quad &- \quad (\Phi(\pi_{\Omega^*}(w), w) - \Phi(\pi_{\Omega^*}(w), \pi_{\Omega^*}(w)) \geq \\ &\geq \quad \gamma |w - \pi_{\Omega^*}(w)|^{1+\nu} \quad \forall w \in \Omega. \end{aligned} \tag{2.15}$$

If the function $\Phi(w, v)$ is differentiable and convex with respect to the second variable, then inequality (2.15) can be represented in the form [2, 3]

$$\langle \nabla \Phi_w(w, w) - \nabla \Phi_w(\pi_{\Omega^*}(w), \pi_{\Omega^*}(w)), w - \pi_{\Omega^*}(w) \rangle \geq \gamma |w - \pi_{\Omega^*}(w)|^{1+\nu} \quad \forall w \in \Omega.$$

3. LINEAR AND BILINEAR CLASSES OF PROBLEMS

In this section we study two popular classes of problems as examples.

3.1. Quadratic Equilibrium

Consider the fixed point problem for the quadratic extremal inclusion

$$v^* \in \text{Argmin}\{0.5\langle Nw, w \rangle + \langle Mv^* + m, w \rangle \mid w \in \Omega\}, \tag{3.1}$$

where N and M are nonnegative matrices, i.e., $\langle Nv, v \rangle \geq 0$ and $\langle Mv, v \rangle \geq 0$ for all $v \in \mathbb{R}^n$. In addition we assume that the matrix N is symmetric. In particular, if $\Omega = \mathbb{R}^n$, then problem (3.1) reduces to the solution of the linear system of equations $(N + M)w = -m$.

Let us verify that the goal function of problem (3.1) satisfies the skew-symmetry condition (2.4) and condition (1.6) holds at the points of a solution. Consider

$$\begin{aligned} &\Phi(w, w) - \Phi(w, v) - \Phi(v, w) + \Phi(v, v) = \\ &= 0.5\langle Nw, w \rangle + \langle Mw, w \rangle + \langle m, w \rangle - 0.5\langle Nv, v \rangle - \langle Mw, v \rangle - \langle m, v \rangle - \\ &- 0.5\langle Nw, w \rangle - \langle Mv, w \rangle - \langle m, w \rangle + 0.5\langle Nv, v \rangle + \langle Mv, v \rangle + \langle m, v \rangle = \\ &= \langle M(w - v), w - v \rangle \geq 0 \qquad \forall w \in \Omega, \quad \forall v \in \Omega. \end{aligned} \tag{3.2}$$

Therefore, if the matrix M is nonnegative, then $\Phi(w, v)$ from (3.1) is skew-symmetric, and hence a solution to this problem satisfies either condition (2.13) or condition (2.14), depending on the behavior of the eigenvalues of the matrix M.

Indeed, let $v = v^*$ in (3.2):

$$\Phi(w, w) - \Phi(w, v^*) - \Phi(v^*, w) + \Phi(v^*, v^*) = \langle M(w - v^*), w - v^* \rangle \geq 0. \tag{3.3}$$

Let the matrix M be degenerate. Consider the decomposition of the space $\mathbb{R}^n$ into the direct sum $\mathbb{R}^n = H_1 + H_2$, where H_1 is the kernel of the matrix M and H_2 is the orthogonal complement of H_1. In this case, any vector $w - v^* \in \mathbb{R}^n$ can be represented as $w - v^* = h_1 + h_2$, where

$h_1 = \pi_{H_1}(w - v^*)$ and $h_2 = \pi_{H_2}(w - v^*)$; moreover, $Mh_1 = 0$ and $Mh_2 \in H_2$. Taking into account these facts, we continue the calculations:

$$\begin{aligned}
&\langle M(w - v^*), w - v^*\rangle = \langle M^{1/2}(w - v^*), M^{1/2}(w - v^*)\rangle = \\
&= \langle M^{1/2}(h_1 + h_2), M^{1/2}(h_1 + h_2)\rangle = \langle M^{1/2}h_2, M^{1/2}h_2\rangle = \\
&= \langle Mh_2, h_2\rangle \geq \mu|h_2|^2 = \mu|w - v^* - h_1|^2 = \mu|w - v^* - \pi_{H_1}(w - v^*)|^2 = \\
&= \mu|w - v^* - \pi_{H_1}(w) - \pi_{H_1}(v^*)|^2 = \mu|w - \pi_{H_1}(w)|^2.
\end{aligned} \tag{3.4}$$

Here μ is the minimal nonzero eigenvalue of the degenerate matrix M. In addition, in this line of reasoning we used the existence of a square root of the symmetric matrix M and the fact that the projection operator in $\pi_{H_1}(w - v^*)$ is linear and $\pi_{H_1}(v^*) = v^*$. Combining (3.3) and (3.4), we obtain Eq. (2.15) with $n = 1$, and taking into account Eq. (1.2), we obtain Eq. (2.14).

3.2. Sharp Equilibrium

Assume that in problem (3.1), the matrices N and M have the form $N = 0$ and $M = \begin{pmatrix} 0 & -L^\top \\ L & 0 \end{pmatrix}$, where L is some submatrix and Ω is endowed with the structure of a polyhedral set:

$$\Omega = \left\{ \begin{pmatrix} z \\ y \end{pmatrix} \mid \begin{pmatrix} 0 & -B \\ A & 0 \end{pmatrix} \begin{pmatrix} z \\ y \end{pmatrix} \leq \begin{pmatrix} -c \\ b \end{pmatrix} \begin{pmatrix} z \\ y \end{pmatrix} \geq \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}.$$

Introduce the vector notation $v = (x, p)$, $w = (z, y)$, and $m = (-c, b)$ and rewrite problem (3.1) in the vector-matrix form

$$(x^*, p^*) \in \mathrm{Argmin} \left\{ \begin{array}{c} \begin{pmatrix} 0 & -L^\top \\ L & 0 \end{pmatrix} \begin{pmatrix} z \\ y \end{pmatrix} + (-c, b) \begin{pmatrix} z \\ y \end{pmatrix} \mid \\ \begin{pmatrix} 0 & -B \\ A & 0 \end{pmatrix} \begin{pmatrix} z \\ y \end{pmatrix} \leq \begin{pmatrix} -c \\ b \end{pmatrix} \begin{pmatrix} z \\ y \end{pmatrix} \geq \begin{pmatrix} 0 \\ 0 \end{pmatrix} \end{array} \right\} \tag{3.5}$$

or, in notation of (3.1),

$$v^* \in \mathrm{Argmin}\{\langle Mv^* + m, w\rangle \mid w \in \Omega\}. \tag{3.6}$$

Using the linearity of the problem and the block structure of its restrictions, we represent (3.5) in the form of a two-person game with nonzero sum:

$$\begin{array}{lcl} x^* & \in & \mathrm{Argmin}\{-\langle c, z\rangle + \langle Lz, p^*\rangle \mid Az \leq b,\ z \geq 0\}, \\ p^* & \in & \mathrm{Argmin}\{\langle b, y\rangle + \langle Lx^*, y\rangle \mid By \geq c,\ y \geq 0\}. \end{array} \tag{3.7}$$

It is assumed that the dimensions of the matrices L, A, and B agree with dimensions of the variables x and y and the vectors c and b. If $L = 0$ and $B = A^\top$, then (3.7) coincides with the a primal linear program and a dual linear program.

Each of the problems in (3.7) is a linear program. It was shown in [9] that the set of optima of any of them is the set of sharp minima, i.e., it satisfies the conditions

$$\begin{aligned} &\gamma|z - x^*| \leq (-\langle c, z\rangle + \langle Lz, p^*\rangle) - (-\langle c, x^*\rangle + \langle Lx^*, p^*\rangle), \\ &\gamma|y - p^*| \leq (\langle b, y\rangle - \langle Lx^*, y\rangle) - (\langle b, p^*\rangle - \langle Lx^*, p^*\rangle) \end{aligned} \tag{3.8}$$

for all $z \in \{z \mid Az \leq b,\ z \geq 0\}$ and for all $y \in \{y \mid By \geq c,\ y \geq 0\}$, where $\gamma > 0$. We add inequalities (3.8) and represent the resultant inequality in the notation of problem (3.6):

$$\langle Mv^*, w\rangle + \langle m, w\rangle - \langle Mv^*, v^*\rangle - \langle m, v^*\rangle \geq \gamma|w - v^*|. \tag{3.9}$$

This inequality holds for all $w \in \Omega$. Since the matrix M in problem (3.6) is nonnegative, i.e., $\langle Mv, v\rangle \geq 0$ for all $v \in \mathbb{R}^n$, the goal function of this problem satisfies the skew-symmetry condition (1.6). Combining (1.6) and (3.9) we obtain the main condition (1.4): $\langle Mw, w\rangle + \langle m, w\rangle - \langle Mw, v^*\rangle - \langle m, v^*\rangle \geq |w - v^*|\ \forall w \in \Omega$. Thus, a fixed point of the linear two-person game (3.7) is a sharp equilibrium.

4. CONTROLLED METHODS

The notion of convergence of a method on the phase portrait of a dynamical system is geometrically associated with a "stable node" singularity to which all trajectories that start from any point of some neighborhood of the equilibrium are convergent. In this connection we face the natural idea to transform the phase portrait of a dynamical system in such a way that an unstable equilibrium transforms into an asymptotically stable node with the same coordinates. In the present paper we realize this idea with the use of the feed-back control [10].

Let us describe this approach. Suppose that the function $\Phi(v, w)$ is differentiate with respect to the second variable, and hence a necessary (and sufficient) condition for the minimum of problem (1.1) can be expressed as the equation

$$v^* = \pi_\Omega(v^* - \alpha\nabla\Phi_w(v^*, v^*)), \tag{4.1}$$

where $\pi_\Omega(\ldots)$ is the projection operator onto the set Ω, $\alpha > 0$ is a parameter of step length type, and $\nabla\Phi_w(v, w)$ is the vector gradient of the function $\Phi(v, w)$ with respect to the second variable w.

The point v^* is either a fixed point or an equilibrium point. If we make a step from the point v^* along the partial antigradient of the function $\Phi(v, w)$, then, after projecting, we arrive at the point v^* again. Problems (1.1) and (4.1) are equivalent.

The residual, that is, the difference between the left- and right-hand sides of (4.1), which vanishes at the point v^* and does not vanish at an arbitrary point v, specifies a transformation of $\mathbb{R}^n$ into itself. The image of this transformation can be considered as a vector field whose fixed point is v^*. Let us state the problem to find a trajectory such that its tangent coincides with the given direction of the field at this point. Formally, this problem is described by the system of differential equations

$$\frac{dv}{dt} + v = \pi_\Omega(v - \alpha\nabla\Phi_w(v, v)), \qquad v(t_0) = v^0. \tag{4.2}$$

Since the right-hand side of this system satisfies all conditions of the existence and uniqueness theorem, system (4.2), for all $v(t_0) = v^0$, generates a trajectory $v(t)$ for all $t \geq t_0$.

The gradient $\nabla\Phi_w(v, v)$ with respect to the second variable is not a potential operator, since, in general, there is no function whose gradient coincides with $\nabla\Phi_w(v, v)$. The examples in [4, 5] show that the process must not converge to an equilibrium, and to provide the convergence, we introduce an additive control [4, 5] into the right-hand side of the dynamical system (4.2):

$$\frac{dv}{dt} + v = \pi_\Omega(v - \alpha\nabla\Phi_w(v + u, v + u)), \qquad v(t_0) = v^0, \tag{4.3}$$

and state the following problem. In some class of feed-backs, choose a control $u = u(v, \dot{v})$ (where $\dot{v} = dv/dt$), as a function of the state of the dynamical system, to provide the convergence of the corresponding trajectory to the equilibrium. At the equilibrium point, the object does not move and its velocities vanish; therefore, $u = u(v^*, v^*) = 0$.

Consider two kinds of control with respect to the derivative: $u = \dot{v}$, and with respect to the residual:

$$u = \pi_\Omega(v - \alpha\nabla\Phi_w(v, v)) - v. \tag{4.4}$$

The closure of system (4.3) by the control with respect to the derivative $u = \dot{v}$ leads to the implicit (unsolved with respect to the derivative) differential system

$$\frac{dv}{dt} + v = \pi_\Omega(v - \alpha\nabla\Phi_w(v + \dot{v}, v + \dot{v})), \qquad v(t_0) = v^0, \tag{4.5}$$

The iterative analog of this system is the implicit iterative process

$$v^{n+1} = \pi_\Omega(v^n - \alpha\nabla\Phi_w(v^{n+1}, v^{n+1})), \tag{4.6}$$

Here v^n is the preceding approximation, and (4.6) must be solved with respect to the variables v^{n+l}. In turn, other iterative subprocesses are necessary to solve this equation.

Taking into account the advantages of process (4.6), it is necessary to note its disadvantage, namely, it is implicit (or unsolved) with respect to the derivative. After the closure of system (4.3) by control (4.4) we obtain

$$\frac{dv}{dt} + v = \pi_\Omega(v - \alpha\nabla\Phi_w(\bar u, \bar u)), \qquad \bar u = \pi_\Omega(v - \alpha\nabla\Phi_w(v, v)). \tag{4.7}$$

System (4.7) is explicit, which becomes especially apparent in the iterative analog

$$\bar u^n = \pi_\Omega(v^n - \alpha\nabla\Phi_w(v^n, v^n)), \qquad v^{n+1} = \pi_\Omega(v^n - \alpha\nabla\Phi_w(\bar u^n, \bar u^n)),$$

This is an explicit iterative scheme with a preliminary (predictive) step on which we first compute the prediction $\bar u^n$ and then the subsequent approximation v^{n+1}.

5. CONVERGENCE TO A SHARP EQUILIBRIUM

Consider the behavior of the predictive method of gradient projection given by (4.7) in the case of a sharp equilibrium [2, 11].

Let us represent the process (4.7), according to the definition of a projection operator, in the form of the following variational inequalities:

$$\langle \dot v + v - v + \alpha\nabla\Phi_w(\bar u, \bar u), w - v - \dot v\rangle \geq 0 \tag{5.1}$$

for all $v \in \Omega$ and

$$\langle \bar u - v + \alpha\nabla\Phi_w(v, v), w - \bar u\rangle \geq 0 \tag{5.2}$$

for all $v \in \Omega$. Moreover, to prove the convergence of the method, we need the Lipschitz condition in the form

$$|\nabla\Phi_w(w, w) - \nabla\Phi_w(v, v)| \leq |\Phi|\,|w - v| \qquad \forall w, v \in \Omega, \tag{5.3}$$

where $|\Phi|$ is a Lipschitz constant for the operator $\nabla\Phi_w(v, w)$. This condition always holds if the gradient of the function $\Phi(v, w)$ is (jointly) Lipschitzian.

Taking into account (5.3), we obtain the estimate for the deviation between $v + \dot v$ and the vector $\bar u$ given by (4.7):

$$\begin{aligned} |v + \dot v - \bar u| &\leq |\pi_\Omega(v - \alpha\nabla\Phi_w(\bar u, \bar u)) - \pi_\Omega(v - \alpha\nabla\Phi_w(v, v))| \leq \\ &\leq |\nabla\Phi_w(\bar u, \bar u) - \nabla\Phi_w(v, v)| \leq \alpha|\Phi|\,|\bar u - v|, \qquad \forall w, v \in \Omega, \end{aligned} \tag{5.4}$$

where $|\Phi|$ is a Lipschitz constant for the vector function $\Phi(v, v)$.

For the gradient $\nabla\Phi_w(v, w)$, replace the (absent) condition of being monotone by the sharpness condition (2.13) with $\nu = 0$ and assume that the original problem has an isolated equilibrium:

$$\Phi(w, w) - \Phi(w, v^*) \geq \gamma|w - v^*|, \quad \forall w \in \Omega. \tag{5.5}$$

In this case, process (4.7) converges at a finite time.

Theorem 1. *If the set of solutions to problem (1.1) is nonempty and satisfies the sharpness condition (5.5), the goal function $\Phi(v, w)$ is convex with respect to the variable w for each chosen*

v, Ω is a closed convex set, and the partial gradient of the function $\Phi(v,w)$ satisfies the Lipschitz condition (5.3), then the trajectory $v(t)$ of process (4.7) with parameter α, $0<\alpha<1/(\sqrt{2}|\Phi|)$, where $|\Phi|$ is the constant occurring in (5.3), converges at a finite time, i.e., there exists an index n_f such that $\bar{u}^{n_f}=v^*$.

Proof. We set $w=v^*$ in Eq. (5.1) and represent the obtained expression in the form

$$\langle\dot{v},v^*-v\rangle-|\dot{v}|^2+\alpha\langle\nabla\Phi_w(\bar{u},\bar{u}),v^*-\bar{u}\rangle+\alpha\langle\nabla\Phi_w(\bar{u},\bar{u}),\bar{u}-v-\dot{v}\rangle\geq 0.$$

Hence, by the convexity condition (2.12) we have

$$\begin{aligned}\langle\dot{v},v^*-v\rangle-|\dot{v}|^2 &+ \alpha(\Phi(\bar{u},v^*)-\Phi(\bar{u},\bar{u}))-\alpha\langle\nabla\Phi_w(v,v)-\nabla\Phi_w(\bar{u},\bar{u}),\bar{u}-v-\dot{v}\rangle+\\ &+ \alpha\langle\nabla\Phi_w(v,v),\bar{u}-v-\dot{v}\rangle\geq 0.\end{aligned}$$

We write $w=v+\dot{v}$ in Eq. (5.2) and add the obtained inequality and the previous one. Taking into account (5.3) and (5.4), we see that

$$\langle\dot{v},v^*-v\rangle-|\dot{v}|^2+\alpha(\Phi(\bar{u},v^*)-\Phi(\bar{u},\bar{u}))+\alpha^2|\Phi|^2|\bar{u}-v|^2+\langle\bar{u}-v,v+\dot{v}-\bar{u}\rangle\geq 0. \tag{5.6}$$

We transform the last inner product on the right-hand side of Eq. (5.6) by using the identity

$$|v_1-v_2|^2=|v_1-v_3|^2+2\langle v_1-v_3,v_3-v_2\rangle+|v_3-v_2|^2, \tag{5.7}$$

then we have

$$\langle\bar{u}-v,v+\dot{v}-\bar{u}\rangle=0.5|\dot{v}|^2-0.5|\bar{u}-v|^2-0.5|\bar{u}-v-\dot{v}|^2.$$

Now we can write (5.6) in the form

$$\langle\dot{v},v-v^*\rangle+|\dot{v}|^2+\alpha(\Phi(\bar{u},\bar{u})-\Phi(\bar{u},v^*))-\alpha^2|\Phi|^2|\bar{u}-v|^2-0.5|\dot{v}|^2+0.5|\bar{u}-v|^2+0.5|\bar{u}-v-\dot{v}|^2\leq 0$$

or

$$\langle\dot{v},v-v^*\rangle+0.5|\dot{v}|^2+\alpha(\Phi(\bar{u},\bar{u})-\Phi(\bar{u},v^*))+(0.5-\alpha^2|\Phi|^2)|\bar{u}-v|^2\leq 0. \tag{5.8}$$

By Eq. (5.5), the third term in the last inequality is nonnegative, and, therefore,

$$0.5\frac{d}{dt}|v-v^*|^2+0.5|\dot{v}|^2+(0.5-\alpha^2|\Phi|^2)|\bar{u}-v|^2\leq 0. \tag{5.9}$$

Since $0<\alpha<1/(\sqrt{2}|\Phi|)$, we have $d=0.5-\alpha^2|\Phi|^2>0$. Integrate inequality (5.9) from t_0 to t:

$$|v-v^*|^2+\int_{t_0}^{t}|\dot{v}|^2d\tau+(1-2\alpha^2|\Phi|^2)\int_{t_0}^{t}|\bar{u}-v|^2d\tau\leq|v_0-v^*|^2.$$

It follows that $|v(t)-v^*|$ is monotone decreasing and the trajectory $v(t)$ is bounded and the integrals

$$\int_{t_0}^{t}|\dot{v}|^2d\tau<\infty,\qquad \int_{t_0}^{t}|\bar{u}-v|^2d\tau<\infty$$

are convergent as $t\to\infty$; hence,

$$\liminf|\dot{v}|^2=0,\qquad \liminf|\bar{u}-v|^2=0,\qquad t\to\infty. \tag{5.10}$$

Let us obtain another assertion of the form (5.10). To this end we consider the inequality $0.5\frac{d}{dt}|v-v^*|^2+0.5|\dot{v}|^2 \le 0$, which follows from (5.9); its differentiation yields

$$|v-v^*|\frac{d}{dt}|v-v^*|+0.5|\dot{v}|^2 \le 0.$$

If we assume that $|v(t)-v^*| \ne 0$ for all $t \ge t_0$, then we have $\frac{d}{dt}|v-v^*|^2+0.5\frac{|\dot{v}|^2}{|v-v^*|} \le 0$.

Integrating this inequality from t_0 to t we obtain the estimate

$$|v-v^*|^2+\frac{1}{2}\int_{t_0}^{t}\frac{|\dot{v}|^2}{|v-v^*|}d\tau \le |v_0-v^*|^2,$$

which implies the convergence of the integral; therefore, $\liminf(|\dot{v}|^2/|v-v^*|)=0$, $t\to\infty$. Since we have the estimate $|\dot{v}|^2/(|v-\bar{u}|+|\bar{u}-v^*|) \le |\dot{v}|^2/|v-v^*|$, it follows from the convergence $|v-\bar{u}|\to 0$ as $t\to\infty$ that

$$\liminf(|\dot{v}|^2/|\bar{u}-v^*|)=0, \qquad t\to\infty. \tag{5.11}$$

Let us return to inequality (5.8). Taking into account the sharpness condition (5.5), we represent (5.8) in the form

$$\langle\dot{v},\bar{u}-v^*\rangle+\langle\dot{v},v-\bar{u}\rangle+0.5|\dot{v}|^2+\alpha\gamma|\bar{u}-v^*|+d|\bar{u}-v|^2 \le 0.$$

In the second and third summands we single out a square:

$$\langle\dot{v},\bar{u}-v^*\rangle+|(0.5/\sqrt{d})\dot{v}+\sqrt{d}(v-\bar{u})|^2-(0.25/d)|\bar{u}|^2+0.5|\dot{v}|^2+\alpha\gamma|\bar{u}-v^*| \le 0.$$

Hence $\alpha\gamma|\bar{u}-v^*| \le 0.25(1/d-2)|\dot{v}|^2+|\dot{v}|\,|\bar{u}-v^*|$. Assuming that $|\bar{u}-v^*|\ne 0$ for all $t\ge t_0$ we obtain

$$\alpha\gamma \le 0.25(1/d-2)|\dot{v}|^2/|\bar{u}-v^*|+|\dot{v}|^2. \tag{5.12}$$

By combining inequalities (5.12), (5.10), and (5.11) we obtain a contradiction. Therefore, the assumption that $|\bar{u}-v^*|\ne 0$ for all $t\ge t_0$ leads to a contradiction; hence, there exists $t=t_f$ such that $\bar{u}(t_f)=v^*$. This completes the proof of the theorem. $\square$

Similarly, one can prove that the implicit process (4.5) also converges to an equilibrium at a finite time under the assumptions of Theorem 1.

6. CONVERGENCE TO A QUADRATIC EQUILIBRIUM

In this section we obtain an exponential estimate for the rate of convergence of process (4.7) under the assumption that the original problem (1.1) has an isolated quadratic equilibrium. Thus, let condition (2.13) with $n=1$ be valid:

$$\Phi(w,w)-\Phi(w,v^*) \ge \gamma|w-v^*|^2, \quad \forall w\in\Omega. \tag{6.1}$$

Theorem 2. *If inequality (6.1) is used instead of (5.5) in the assumption of Theorem 1, then the trajectory $v(t)$ converges to an equilibrium solution with exponential rate, i.e.,*

$$|v(t)-v^*|^2 \le |v^0-v^*|^2\exp(2s(\alpha)(t_0-t)), \tag{6.2}$$

where $s(\alpha) = \alpha\gamma(d/d_1) \geq 0$, $d = 0.5 - \alpha^2|\Phi|^2$, *and* $d_1 = 0.5 + \alpha\gamma - \alpha^2|\Phi|^2$.

Proof. Using (6.1), we estimate the third summand of inequality (5.8) of Theorem 1. Then we obtain

$$\langle \dot{v}, v - v^* \rangle + 0.5|\dot{v}|^2 + \alpha\gamma|\bar{u} - v^*| + d|\bar{u} - v|^2 \leq 0. \tag{6.3}$$

We transform the third summand on the left-hand side of (6.3) by means of identity (5.10) with $v_1 = \bar{u}$, $v_2 = v^*$, and $v_3 = v$. Then (6.3) becomes

$$\langle \dot{v}, v - v^* \rangle + 0.5|\dot{v}|^2 + d_1|\bar{u} - v|^2 + 2\alpha\gamma\langle \bar{u} - v, v + v^* \rangle + \alpha\gamma|v - v^*|^2 \leq 0. \tag{6.4}$$

In the sum of the third term and the fourth term of (6.4) we single out a square:

$$\langle \dot{v}, v - v^* \rangle + 0.5|\dot{v}|^2 + |\sqrt{d_1}(\bar{u} - v^*) + (\alpha\gamma/\sqrt{d_1})(v - v^*)|^2 - ((\alpha\gamma)^2/d_1)|v - v^*|^2 + \alpha\gamma|v - v^*| \leq 0,$$

hence, we have

$$\langle \dot{v}, v - v^* \rangle + 0.5|\dot{v}|^2 + s(\alpha)|v - v^*| \leq 0, \tag{6.5}$$

where $s(\alpha) = \alpha\gamma(d/d_1) > 0$. By assumption, v^* is a unique minimum. Therefore, (6.5) can be represented in the form $d|v - v^*|^2/dt + 2s(\alpha)|v - v^*|^2 \leq 0$. Let us rewrite the obtained inequality in the form

$$\exp(-2s(\alpha)t)\frac{d(\exp(2s(\alpha)t)|v - v^*|^2)}{dt} \leq 0 \quad \text{or} \quad \frac{d(\exp(2s(\alpha)t)|v - v^*|^2)}{dt} \leq 0,$$

which implies estimate (6.2). We can readily see that $s(\alpha) = \alpha\gamma(d/d_1) > 0$ is nonnegative for $0 < \alpha < 1/(\sqrt{2}|\Phi|)$. The proof of the theorem is completed. $\square$

Thus, if the value of the parameter $\alpha > 0$ is not very large, then the trajectory $v(t)$ converges to the equilibrium solution with exponential rate. The implicit process (4.5) also converges with exponential rate if the assumptions of Theorem 2 are satisfied.

7. CONVERGENCE TO A DEGENERATE EQUILIBRIUM

Let us state a theorem on the convergence of process (4.7) in the degenerate case, that is, for the case in which condition (1.4) holds.

Theorem 3. *If the set of solutions to problem (1.1) is nonempty and satisfies condition (1.4), the goal function $\Phi(v, w)$ is convex with respect to the variable w for each chosen v, Ω is a closed convex set, and the vector function $\nabla\Phi_w(v, w)$ satisfies Lipschitz condition (5.3), then the trajectory $v(t)$ of process (4.7) with parameter α, $0 < \alpha < 1/(\sqrt{2}|\Phi|)$, where $|\Phi|$ is the constant in (5.3), is monotone convergent, with respect to the norm, to an equilibrium state, i.e., $v(t) \to v^* \in$ $\in \Omega^*$ as $t \to \infty$ for all $v^0 \in \mathbb{R}^n$.*

Proof. Using transformations similar to those in the proof of Theorem 1 we establish assertion (5.10).

Let a sequence $t_i \to \infty$ satisfy the conditions $v(t_i) \to v'$, $\bar{u}(t_i) \to v'$, and $\dot{v}(t_i) \to 0$. Consider (5.1) and (5.2) for all $t \to \infty$. Passing to the limit, we obtain the limit inequality of the form $\langle \nabla\Phi_w(v', v'), w - v' \rangle \geq 0$ for all $w \in \Omega$, which coincides with inequality (1.3). The last relation means that v' is an equilibrium solution to the problem.

Thus, each limit point of the trajectory $v(t)$ is a solution to the original problem. This fact, together with the remark that $|v(t) - v^*|$ is monotone decreasing as $t \to \infty$, implies that $v(t)$ has a unique limit point, i.e., the trajectory $v(t)$ is monotone convergent, with respect to the norm, to a solution of the original problem, $v(t) \to v^* \in \Omega^*$ as $t \to \infty$ for any $v^0 = v(t_0) \in \mathbb{R}^n$. The theorem is proved. $\square$

The convergence of the implicit process (4.5) under the assumption of Theorem 3 can be proved similarly.

Thus, we proved that in the case of a sharp equilibrium, the considered processes converge at a finite time, the methods have an exponential rate of convergence for a quadratic equilibrium, and the processes can converge as slowly as desired for a degenerate equilibrium.

The research was partially supported by the Russian Foundation for Basic Research (grant No. 94-01-00005).

LITERATURE CITED

1. F.P. Vasil'ev, *Numerical Methods for Solution of Extremal Problems* [in Russian], Moscow (1988).

2. A.S. Antipin, *Operations Research Proceedings*, Eds.: U. Derigs, A. Bachem, and A. Drexl, Berlin, Heidelberg (1994), pp. 11–15.

3. A.S. Antipin, *Zh. Vychislit. Mat. Mat. Fiz.*, **35**, No.5, pp. 688–704 (1995).

4. A.S. Antipin, *Differents. Uravn.*, **28**, No.11, pp. 1846–1861 (1992).

5. A.S. Antipin, *Avtomat. Telemekh.*, No.3, pp. 12–23 (1994).

6. H. Nikaido and K. Isoda, *Pacific Journal of Math.*, **5**, Suppl. 1, pp. 807–815 (1955).

7. J.-P. Aubin and I. Ekeland, *Applied Nonlinear Analysis*, John Wiley & Sons, N.Y. (1984).

8. J.-P. Aubin and H. Frankowska, *Set Valued Analysis*, Birkhauser, Boston, Basel, Berlin (1990).

9. B.T. Polyak, *Introduction to Optimization* [in Russian], Moscow (1983).

10. S.V. Emel'yanov and S.K. Korovin, *Dynamics of Inhomogeneous Systems* [in Russian], Proceedings of Seminar in VNIISI, Moscow (1982), pp. 5–27.

11. A.S. Antipin, *Dokl. RAN*, **342**, No.3, pp. 1–4 (1995).