Dokl. Akad. Nauk SSSR
T.233 (1977), No.3,
Pp. 277–280

Soviet Math. Dokl.
Vol.18 (1977), No.2 ,
Pp. 370–373

UDC 512.98
AMS (MOS) subject classifications (1970). Primary 46D05, 46B40.

APPLICATION OF THE SINGULAR PERTURBATION METHOD FOR SOLVING MINIMAX PROBLEMS

N.I. GRAČEV AND Ju.G. EVTUŠENKO

Revised 8 April 2001

1. Consider the problem of finding the minimax

$$\min_{x\in E_n}\max_{y\in E_m} F(x,y), \tag{1}$$

where $F(x,y)$ is a continuously differentiable function, E_i is i-dimensional Euclidean space, and $x=[x^1,\dots,x^n]$, $y=[y^1,\dots,y^m]$.

Let $z_*=(x_*,y_*)$ be a strict local solution point of (1). This means that there exist neighborhoods X and Y of the points x_* and y_*, respectively, such that the following inequalities hold for any $x\in X$, $y\in Y$, $x\neq x_*$, $y\neq y_*$:

$$F(x_*,y)<F(x_*,y_*)<F(x,\bar{y}(x))=\sup_{y\in Y}F(x,y).$$

According to [1], for z_* to be a strict local solution of (1) it is necessary that

$$F_x(z_*)=F_y(z_*)=0, \tag{2}$$

and sufficient that the following quadratic forms be negative definite:

$$y^\top F_{yy}(z_*)y<0,\quad x^\top[F_{xy}(z_*)F_{yy}^{-1}(z_*)F_{yx}(z_*)-F_{xx}(z_*)]x<0 \tag{3}$$

for $\forall y\in E_m$, $\|y\|\neq 0$, $\forall x\in E_n$, $\|x\|\neq 0$.

In [1] and [2], iterative methods for solving (1) were proposed. However, these methods turned out to be laborious for problems of high dimensionality, since they required multiple inversion of the matrices of second derivatives. The following easily implemented method can be employed for approximate computations. We shall be looking for limit points (as $t\to\infty$) of the solution of the Cauchy problem

$$\dot{x}=-\varepsilon F_x(x,y),\quad \dot{y}=F_y(x,y),\quad (\cdot)=d/dt,\quad x(0)=x_0,\quad y(0)=y_0. \tag{4}$$

A discrete variant is possible, wherein

$$x_{s+1}=x_s-\varepsilon\alpha F_x(x_s,y_s),\qquad y_{s+1}=y_s+\alpha F_y(x_s,y_s),\qquad s=0,1,2,\dots; \tag{5}$$

here, $0<\varepsilon<1$ and $0<\alpha$.

We denote the solutions of (4) by $x(z_0,t,\varepsilon)$, $y(z_0,t,\varepsilon)$, where $z_0=(x_0,y_0)$. Making use of the results of [3] and [4], one can show that the following theorem holds.

Theorem 1. *Let $F(x,y)$ be twice continuously differentiable in a neighborhood of a point z_* at which the conditions* (2) *and* (3) *hold. Then there exist a neighborhood W of z_* and numbers $\bar{\alpha}>0$ and $\bar{\varepsilon}>0$ such that, for all $z_0\in W$, $0<\alpha<\bar{\alpha}$, $0<\varepsilon<\bar{\varepsilon}$, the solutions of* (4) *and* (5) *converge to z_*, i.e.*

$$\lim_{t\to\infty} x(z_0,t,\varepsilon)=x_*,\qquad \lim_{t\to\infty} y(z_0,t,\varepsilon)=y_*,$$
$$\lim_{s\to\infty} x_s=x_*,\qquad \lim_{s\to\infty} y_s=y_*.$$

Theorem 1 leads to a simple scheme for solving (1). The existence of variables that change both slowly and rapidly in (4) and (5) complicates the computations; however, this approach turns out to be very effective in a number of higher-dimensional problems. In fact, we have made a transition opposite to that usually made in the theory of singular equations. Namely, instead of solving a degenerate problem (which may be very complicated), we are solving the problem (4), which is equivalent to the singularly perturbed problem.

Example. Let $F(x,y)=e^{x^2}\sin 2\pi(x-y)$. It is easy to see that

$$\min_{x\in E_1}\max_{y\in E_1} F(x,y)=1,\qquad \max_{y\in E_1}\min_{x\in E_1} F(x,y)=-1.$$

Necessary and sufficient conditions for a minimax and a maximin, respectively, are satisfied at the solution points of both problems. Therefore, the solutions of (4) and (5) converge to a minimax solution for $\varepsilon\ll 1$, and to a maximin solution for $\varepsilon\gg 1$.

Let us present two more methods of solving (1), in which

$$\dot{x}=-F_x(x,y(x)),\qquad F(x,y(x))=\max_{y\in E_m}F(x,y);\tag{6}$$

$$\dot{x}=-\varepsilon F_x(x,y),\qquad \dot{y}=-F_{yy}^{-1}(x,y)F_y(x,y).\tag{7}$$

Their discrete versions have the form

$$x_{k+1}=x_k-\varepsilon F_x(x_k,y_k),\qquad F(x_k,y_k)=\max_{y\in E_m}F(x_k,y);\tag{8}$$

$$x_{k+1}=x_k-\varepsilon F_x(x_k,y_k),\qquad y_{k+1}=y_k-F_{yy}^{-1}(x_k,y_k)F_y(x_k,y_k).\tag{9}$$

Sufficient conditions for convergence are formulated in the same way as in Theorem 1 (except that one does not need to introduce a small step α in discrete variants).

$\boxed{\textbf{2.}}$ The methods must be changed significantly to solve problems with hierarchic structure and games with nonconflicting interests [5]. The methods for solving such problems are presented in [6] and [7]. Following [7], we express the problem of finding a guaranteeing strategy in the following form:

$$J=\min_{x\in E_n}F(x,y)\min_{y\in E_m}K(x,y).\tag{10}$$

Here, for every fixed x one seeks a set of $y(x)$ from the condition

$$K(x,y(x))=\min_{y\in E_m}K(x,y),$$

and then computes

$$J=\min_{x\in E_n}\max_{y\in y(x)}K(x,y(x))=K(x_*,y(x_*))=K(x_*,y_*).$$

The point (x_*, y_*) will be called a *solution of problem* (10). If the function $K(x,y)$ attains a strict local minimum in y at the points $y(x)$ for all x in a neighborhood of x_*, and if the function $F(x,y(x))$ attains a strict local minimum at x_*, then we shall say that $z_* = (x_*, y_*)$ is a strict local solution of (10). We set

$$z = (x,y), \quad B(z) = K_{yy}^{-1}(z)K_{yx}(z), \quad H(z,\lambda) = F(z) + K_y^\top(z)\lambda,$$

$$N(z,\lambda) = H_{xx}(z,\lambda) - B^\top(z)H_{yx}(x,\lambda) - H_{yx}(z,\lambda)B(z) + B^\top(z)H_{yy}(z,\lambda)B(z),$$

where $\lambda \in E_m$.

According to [7], for z_* to be a strict local solution of (10) it is necessary that

$$K_y(z_*) = 0, \qquad F_x(z_*) - B^\top(z_*)F_y(z_*) = 0, \tag{11}$$

and sufficient that $\forall y \in E_m,\ \forall x \in E_n,\ \|x\| \neq 0,\ \|y\| \neq 0,$

$$y^\top K_{yy}(z_*)y > 0, \qquad x^\top N(z_*, \lambda_*)x > 0. \tag{12}$$

Here, λ_* is determined from the condition

$$H_x(z_*, \lambda_*) = 0.$$

We present three methods of solving (10):

$$\dot{x} = -\varepsilon[F_x(x,y) - B^\top(x,y)F_y(x,y)] = -\varepsilon\varphi(x,y), \qquad \dot{y} = -K_y(x,y); \tag{13}$$

$$\dot{x} = -\varphi(x,y(x)), \qquad K(x,y(x)) = \min_{y \in E_m} K(x,y); \tag{14}$$

$$\dot{x} = -\varepsilon\varphi(x,y), \qquad \dot{y} = -K_{yy}^{-1}(x,y)K_y(x,y). \tag{15}$$

Theorem 2. *Let the functions $F(z)$ and $K(z)$ be twice continuously differentiable in a neighborhood of a point z_* at which the conditions* (11) *and* (12) *are satisfied. Then the methods* (13), (14), *and* (15), *and their discrete variants of the form* (5), (8), *and* (9), *converge locally to the point z_*.*

$\boxed{\textbf{3.}}$ Let us apply the approach presented above to the solution of nonlinear programming problems:

$$\min_{x \in X} F(x), \qquad X = \{x \in E_n : g(x) = 0,\ \ h(x) \leq 0\}, \tag{16}$$

where $x \in E_n,\ g \in E_c$, and $h \in E_m$. Following [8] and [9], we form a modified Lagrangian

$$L(x,p,w) = F(x) + \sum_{i=1}^{c} p^i g^i(x) + \sum_{i=1}^{m} (w^i)^2 h^i(x).$$

We shall seek

$$\max_{p \in E_c} \max_{w \in E_m} \min_{x \in E_n} L(x,p,w).$$

The methods presented above yield three schemes for solving problem (16):

$$\dot{p} = \varepsilon L_p(x,p,w), \qquad \dot{w} = \varepsilon L_w(x,p,w), \qquad \dot{x} = -L_x(x,p,w); \tag{17}$$

$$\dot{p} = L_p(x(p,w),p,w), \quad \dot{w} = L_w(x(p,w),p,w), \quad L(x(p,w),p,w) = \min_{x \in E_n} L(x,p,w); \tag{18}$$

$$\dot{p} = L_p(x,p,w), \quad \dot{w} = L_w(x,p,w), \quad \dot{x} = -L_{xx}^{-1}(x,p,w)L_x(x,p,w). \tag{19}$$

We assume that there exists a solution of problem (16), namely a point x_* at which the condition for regularity of constraints (CRC) holds, i.e. all the vectors $g_x^i(x_*)$ and $h_x^j(x_*)$ such that $h^j(x_*) = 0$ are linearly independent. Moreover, at a stationary point $z_* = (x_*, p_*, w_*)$ (a point such that $L_x = L_p = L_w$), the condition for strict complementary slackness (CSCS) holds, i.e. $h^j(x_*) = 0$ implies $w^j \neq 0$.

Theorem 3. *Let the function $L(x,p,w)$ be twice continuously differentiable in a neighborhood of an admissible stationary point z_*, let the matrix $L_{xx}(z_*)$ be positive definite, and let the CRC and CSCS hold. Then there exist sufficiently small numbers $\bar{\varepsilon}$ and $\bar{\alpha}$ such that, for $0 < \varepsilon < \bar{\varepsilon}$ and $0 < \alpha < \bar{\alpha}$, the methods* (17), (18) *and* (19), *and their discrete variants of the forms* (5), (8) *and* (9), *locally converge (exponentially and geometrically, respectively) to the point z_*.*

The convergence of (17) was proved in [7] for $\varepsilon = 1$. The introduction of a small parameter ε improves the convergence, which makes the method closer to the method (18). The latter, although more laborious, converges more rapidly in practice.

BIBLIOGRAPHY

1. Ju.G. Evtušenko, Ž. Vyčisl. Mat. i Mat. Fiz., 14 (1974), 669 = U.S.S.R. Comput. Math. and Math. Phys., 14 (1974), No.3, 129. MR 50 #10965.

2. Ju.G. Evtušenko, Ž. Vyčisl. Mat. i Mat. Fiz., 14 (1974), 1138 = U.S.S.R. Comput. Math. and Math. Phys., 14 (1974), No.5, 52. MR 51 #4672.

3. A.N. Tihonov, Mat. Sb., 31 (73) (1952), 575. (Russian) MR 14, 1085.

4. A.B. Vasil'eva and V.F. Butuzov, Asymptotic expansions of solutions of singularly perturbed equations, "Nauka", Moscow, 1973. (Russian)

5. Ju.G. Germeĭer and N.N. Moiseev, Problems in Appl. Math. and Mech., "Nauka", Moscow, 1971, p. 30. (Russian) Zbl. 293 #93014.

6. I.I. Eremin, Nonstationary Processes in Mathematical Programming, Sverdlovsk, 1974. (Russian)

7. N.I. Gračev and Ju.G. Evtušenko, Operations Research, vyp.4, Vyčisl. Centr Akad. Nauk SSSR, Moscow, 1974, p. 219. (Russian) RŽ Mat. 1975 #461246.

8. Ju.G. Evtušenko, Dokl. Akad. Nauk SSSR, 221 (1975), 1016 = Soviet Math. Dokl., 16 (1975), 443. MR 52 #2201.

9. Ju.G. Evtušenko, Ž. Vyčisl. Mat. i Mat. Fiz., 16 (1976), 307 = U.S.S.R. Comput. Math. and Math. Phys., 16 (1976), No.2, 24.