ALGORITHMS FOR SOLVING NON-LINEAR PROGRAMMING PROBLEMS

Yu.G. Evtushenko

Computing Center USSR Academy of Sciences.

Moscow

1. We consider the following problems

$$\min_{x \in X_1} F(x), \quad X_1 = \{x \in E_n : g(x) = 0, h(x) \leq 0, 0 \leq x\} \quad (1)$$

$$\min_{x \in X_2} F(x), \quad X_2 = \{x \in E_n : g(x) = 0, h(x) \leq 0\} \quad (2)$$

where F, g, h are continuously differentiable functions defined on E_n , Euclidean n space, $x = [x^1, x^2, \dots, x^n]$ is a point in E_n ,

functions F, g, h define the mapping $F: E_n \rightarrow E_1$,
 $g: E_n \rightarrow E_c, h: E_n \rightarrow E_c, R = [g, h] \in E_{2c}$. We

define two open sets

$$X_1^0 = \{x: g(x) = 0, h(x) < 0, x > 0\}, X_2^0 = \{x: g(x) = 0, h(x) < 0\}.$$

R_x is the $n \times m$ matrix whose ij -th element is
 $\partial R^i / \partial x^j$; F_x is the $n \times 1$ column matrix whose ij -th
 element is $\partial F / \partial x^i$; $D(x), D(\sqrt{x}), D(R), D(\sqrt{-R})$ are diagonal
 matrices of an order of $n \times n, n \times n, m \times m, m \times m$
 respectively whose j -th diagonal elements are $x^j, \sqrt{x^j}, R^j, \sqrt{-R^j}$
 respectively.

For numerical solution of (1) we propose to find the points
 which are the limits (when $t \rightarrow \infty$) of a solution to the following
 Cauchy problem [1]

$$\dot{x} = -D(x) [F_x + R_x v], (\cdot) = d/dt, x(0) = x_0 \in X_1^0 \quad (3)$$

where $v \in E_m, m = l + k$ is defined from a linear set of equations

$$A v + R_x' D(x) F_x = 0, A = R_x' D(x) R_x - D(R).$$

We suppose that the sets X_1^0, X_2^0 are non-empty and the set $Z = \{x: F(x) \leq F(x_0), x \in X_1\}$ is bounded.

Let e^i denote a unit vector whose i -th component is equal to unity.

Definition: The constraint qualification holds at point x , if the vectors $R_x^j(x)$ subject to $R^j(x) = 0$, $j \in [1:m]$ and vectors e^i subject to $x^i = 0$ are linearly independent.

Lemma 1: If constraint qualification holds at any point $x \in X_1$, then $A(x)$ is a nonsingular positive definite matrix at any point $x \in X_1$.

The right sides of system (3) are continuous throughout X_1 , and there exists a solution to Cauchy problem (3). We assume, moreover, that it is unique. Calculating the first derivative of $R(x)$, with respect to system (3), we obtain

$$\dot{R} = -D(R) \vartheta, \dot{F} = -\|D(\sqrt{x}) H_x\|^2 - \|D(\sqrt{-R}) v\|^2, \\ \text{where } \|z\| = z'z, H_x = F_x + R_x v.$$

Starting at any point $x_0 \in X_1^0$, a continuous solution $x(x_0, t)$ to (3) exists, and $x(x_0, t)$ remains in X_1^0 , for all $t \geq 0$. Let $\tilde{x}$ denote a solution to (3) and V be a neighborhood of $\tilde{x}$.

Theorem 1: Let the set Z be compact, X_1^0 be non-empty, the set of equilibrium points of system (3) be finite on X_1 , the constraint qualification hold on the set X_1 . Then a solution $x(x_0, t)$ to system (3) converges to the point $\tilde{x}$ for any $x_0 \in V \cap X_1^0$. Moreover, if the function $F(x)$ and the set X_1 are strictly convex, then the solution $x(x_0, t)$ converges to a global solution to (1) for any $x_0 \in X_1^0$.

We can also use a discrete version of (3)

$$x_{s+1} = x_s - \alpha M(x_s) F_x(x_s), M = D(x) - D(x) R_x [R_x D(x) R_x - D(R)]^{-1} D(x) R_x$$

We denote $\beta = \max_{x \in Z} \|\vartheta(x)\|$, $\alpha = \max_{x \in Z} \|H_x(x)\|$, $\lambda = \max_{y \in E_n} \max_{x \in Z} y' M(x) y / \|y\|^2$.

We assume that the gradient F'_x satisfies Lipschitz condition with constant L on X_1 .

Theorem 2: If the assumptions of the previous theorem are satisfied, the function $R(x)$ is linear, $0 < \alpha < \min[1/\alpha, 1/(2/L)]$ then a solution x_s to (4) converges to $\tilde{x}$ for any $x_0 \in V \cap X_1^0$ and $F(x_{s+1}) \leq F(x_s)$, $x_s \in X_1^0$ for $s = 0, 1, 2, \dots$.

It is possible to change α at each step using a method of steepest descent under the condition $x_s \in X_1^0$. In some particular problems method (3) coincides with the methods proposed in [2, 3].

2. Associated with the general non-linear programming problem (2) are various dual problems. This section develops a new form of a dual problem. Define a modified Lagrangian

$$N(z) = F(x) + \sum_{i=1}^l p^i g^i(x) + \sum_{i=1}^c h^i(x) (w^i)^2$$

where $w \in E_c$, $p \in E_l$, $z = (x, p, w) \in E_{n+l+c}$.

Consider the max-min problem

$$J = \max_{p \in E_l} \max_{w \in E_c} \min_{x \in E_n} N(x, p, w) \quad (5)$$

The necessary conditions for a max-min are given by

$$N_x(z) = N_p(z) = N_w(z) = 0 \quad (6)$$

We shall solve max-min problem (5) instead of (2). The application of a modified Lagrangian w permits us to transform problem (2) to an unconstrained max-min problem. In dealing with N , the multiple w is not constrained to be non-negative, in contrast to an ordinary Lagrangian. Using gradient, max-min algorithms [4, 5] yield the following three methods

$$\dot{x} = -N_x, \quad \dot{p} = N_p, \quad \dot{w} = N_w \quad (7)$$

$$\dot{x} = -N_x, \quad \dot{p} = g - g'_x N_{xx}^{-1} N_x, \quad \dot{w} = 2 D(w) [h - h'_x N_{xx}^{-1} N_x] \quad (8)$$

$$\dot{x} = -N_x - N_{xx}^{-1} (g_x g + 4 h_x D(w) D(w) h), \quad \dot{p} = g, \quad \dot{w} = 2 D(w) h \quad (9)$$

Newton's method gives

$$N_{\bar{x}\bar{z}}(\bar{z}) \dot{\bar{z}} = -N_{\bar{z}}(\bar{z}), \quad (10)$$

$$\bar{z}_{s+1} = \bar{z}_s - N_{\bar{z}\bar{z}}^{-1}(\bar{z}_s) N_{\bar{z}}(\bar{z}_s). \quad (11)$$

Let us denote the points satisfying (6) by $\bar{z}_* = (x_*, p_*, w_*)$.

Assume now that $N(\bar{z})$ is twice continuously differentiable in all arguments in some neighborhood of $\bar{z}_*$.

Lemma 2: Let the multiple w_* satisfy the strict complementary slackness conditions, the constraint qualification hold at a feasible point x_* , $N_{\bar{x}\bar{x}}(\bar{z}_*)$ be a positive definite matrix. Then $\bar{z}_*$ is an isolated local saddle point of $N(\bar{z})$ and x_* is a local solution to (2).

We can use the finite difference approximation to (7)-(10). For example, (7) yields

$$x_{s+1} = x_s - \alpha N_x(\bar{z}_s), \quad p_{s+1} = p_s + \alpha N_p(\bar{z}_s), \quad w_{s+1}^j = w_s^j + \alpha N_w^j(\bar{z}_s) \quad (12)$$

where α is the step length.

Theorem 3: If the assumptions of Lemma 2 are satisfied, then the algorithms (7)-(10) converge locally to $\bar{z}_*$.

The finite difference approximations to (7)-(10), similar to (12), converge locally to $\bar{z}_*$ if α is sufficiently small.

The method (11) converges quadratically to $\bar{z}_*$.

3. In this section two modifications of a penalty function algorithm are presented. Consider a max-min problem associated with (2)

$$\max_{0 \leq \tau \leq T} \min_{x \in E_n} \mathcal{P}(x, \tau), \quad \mathcal{P}(x, \tau) = F(x) + \tau \left[\sum_{i=1}^{\ell} (g^i)^2 + \sum_{i=2}^c [p^i + |h^i|]^2 \right] \quad (13)$$

The pair $\tau = T, x \in \bar{Z}$ will be a solution to this problem, where

$$\bar{Z} = \{ \bar{x} \in E_n : \min_{x \in E_n} \mathcal{P}(x, T) = \mathcal{P}(\bar{x}, T) \}.$$

Let X_* denote the set of solutions to problem (2). It is easy to show that if Z, X_* are non-empty, then there exists such a scalar γ that the following inequality holds for any $\tilde{x} \in Z$

$$F(x_*) - \gamma/\tau \leq \mathcal{P}(\tilde{x}, T) \leq F(x_*).$$

If $F(x)$ is bounded below $F(x) \geq \delta$ then $\sum_{i=1}^L [g^i(x)]^2 + \sum_{i=1}^L [h^i(x) + |h^i(x)|]^2 \leq [F(\tilde{x}) - \delta]/T$. By taking T sufficiently large we can thereby find an approximate solution to problem (2) with the required accuracy.

We shall call the constraints essential in problem (2) if the unconstrained infimum of $F(x)$ differs from solution to (2). To solve (13) we shall use a method similar to (3)

$$\dot{x} = -\mathcal{P}_x(x, \tau), \quad \dot{\tau} = \mathcal{P}_\tau(T - \tau), \quad \tau(0) = 0, \quad x(0) = x_0. \quad (14)$$

Theorem 4: Let F, h be convex, continuously differentiable functions, $g(x)$ be a linear function, Z and X_* be non-empty compact sets, the constraints be essential. Then a solution $x(x_0, t)$ to (14) converges to Z for arbitrary $x_0 \in E_n$.

Another algorithm can be used

$$\dot{x} = -\mathcal{P}_x(x, \tau) \quad \text{where} \quad 0 \leq \tau(t) \leq d\tau/dt, \quad x(0) = x_0 \quad (15)$$

Theorem 5: If the assumptions of the previous theorem are satisfied and X_2^0 is a non-empty set, then a solution $x(x_0, t)$ to (15) converges to X_* for any $x_0 \in E_n$.

R e f e r e n c e s

1. Ю. Г. Евтушенко. Два численных метода решения задач нелинейного программирования. Доклады АН СССР, 1974, т. 215, № 1, 38-40.
2. Ю. Г. Евтушенко, В. Г. Жадан. Численные методы решения некоторых задач исследования операций. Ж. вычисл. матем. и матем. физ., 1973, 13, № 3, 583-598.
3. И. И. Дикин. О непрерывных аналогах метода внутренних точек. Управляемые системы, Новосибирск, 1971, № 9.
4. Ю. Г. Евтушенко. Некоторые локальные свойства минимаксных задач. Ж. вычисл. матем. и матем. физ., 1974, 14, № 3, 669-679.

5.Ю.Г.Евтушенко. Итеративные методы решения минимаксных задач. Ж. вычисл. матем. и матем. физ. 1974, 14, № 5.