Newton-type methods and generalized solutions to improper linear programs

L. D. Popov*

*Institute of Mathematics and Mechanics UB RAS, popld@imm.uran.ru

A linear (as well as nonlinear) mathematical program is said to be improper (ill-posed, infeasible) if it has no feasible solutions and therefore has no optimal points [1]. Following to [2–3], we introduce some kind of a generalized solution to such a program and propose original Newton-type method that gives us this generalized solution if a program is improper and an ordinary solution if this program is feasible.

1. Let us consider the linear program

$$\min\{(c,x):\ Ax=b,\ x\geq 0\} \tag{1}$$

and the dual one

$$\max\{(b,y):\ A^T y\leq c\}, \tag{2}$$

where matrix $A=(a_{ij})_{m\times n}$ and vectors c and b are given, x and y are vectors of primal and dual unknowns, $(\cdot,\cdot)$ denotes scalar product.

We assume the program (1) is improper of the 1st kind, i. e. it has no feasible solutions but the constraints of the dual program (2) are consistent. As a consequence, the follow perturbed program

$$\min\{(c,x):\ Ax=b-u,\ x\geq 0\} \quad (=:\zeta(u))$$

is solvable iff it is feasible ($u=(u_1,\ldots,u_m)$ is parameter of perturbation).

Denote

$$M(u)=\{x:\ \ Ax=b-u,\quad x\geq 0\}$$

and

$$\bar u=\arg\min\{\,\|u\|:\ M(u)\neq\emptyset\}, \tag{3}$$

where $\|\cdot\|$ is Euclidean norm.

Basing on the vector $\bar u$ let us define a generalized solution to (1) as an ordinary solution to the perturbed program

$$\min\{(c,x):\ Ax=b-\bar u,\ x\geq 0\} \quad (=:\zeta(\bar u)); \tag{4}$$

that is solvable and closest to the original one in the sense (3).

If program (1) is feasible, then $\bar u=0$, so for a proper program our generalized solution coincides with an ordinary one.

2. Let consider the mixed penalty function

$$\bar F_\epsilon(x)=(c,x)-\epsilon\sum_{i=1}^n \ln x_i+\frac{1}{2\epsilon}\|Ax-b\|^2,$$

where $\epsilon>0$ (see [4]).

Let $\bar x_\epsilon$ be an unique minimizer for $\bar F_\epsilon(x)$ over all x positive and $\bar u_\epsilon=b-A\bar x_\epsilon$, $\bar v_\epsilon=\epsilon c-A^T\bar u_\epsilon$. One can verify that the triplet $(\bar x_\epsilon,\bar u_\epsilon,\bar v_\epsilon)$ is an unique solution to the follow system of nonlinear equations

$$\Phi(x,u,v;\epsilon)=\begin{pmatrix} Ax+u-b\\ A^T u+v-\epsilon c\\ Vx-\epsilon^2 e\end{pmatrix}=0. \tag{5}$$

Here $v=(v_1,\ldots,v_n)$, $V=\mathrm{diag}(v_1,\ldots,v_n)$ and $e=(1,\ldots,1)$. Vectors x and v both are positive, v is a slack vector.

The system (5) is just a rewriting of well-known KKT-optimality conditions; it is very similar to the systems which are studying in the theory of central path in mathematical programming [5]. We aim to establish the convergence properties of the sequence $(\bar x_\epsilon,\bar u_\epsilon)$ as $\epsilon\to+0$.

3. Let the dual program (2) satisfy Slater's condition:

$$A^T y_0 < c \quad \text{for some} \quad y_0. \tag{6}$$

This condition implies the program (4) is solvable and its optimal set is bounded.

The convergence properties in question are given in the following statements.

Theorem 1 *If an optimal set of program* (4) *is bounded, then the system* (5) *has an unique solution for every* $\epsilon > 0$.

Theorem 2 *Let the condition* (6) *hold. Then the components* $\bar{u}_\epsilon$ *and* $\bar{x}_\epsilon$ *of the corresponding solutions of the system* (5) *are bounded for all* $0 < \epsilon < \bar{\epsilon}$.

Theorem 3 *If the triplet* $(\bar{x}_\epsilon, \bar{u}_\epsilon, \bar{v}_\epsilon)$ *satisfies to equations* (5), *then*

$$\zeta(\bar{u}_\epsilon) \le (c, \bar{x}_\epsilon) \le \zeta(\bar{u}_\epsilon) + n\epsilon.$$

Theorem 4 *Let the condition* (6) *hold and* $\bar{x}_\epsilon$ *be an unique minimizer of the function* $F_\epsilon(x)$ *over* $x > 0$ *with* $\epsilon > 0$. *Then*

$$b - A\bar{x}_\epsilon = \bar{u}_\epsilon \to \bar{u}, \qquad (c, \bar{x}_\epsilon) \to \zeta(\bar{u})$$

as $\epsilon \to +0$.

Thus, the system (5) appears to be a good tool to find our generalized solution.

4. Given $\epsilon > 0$, one can get a solution to (5) as a limit point of the iterative sequence $\{x^s, u^s, v^s\}$ generated by the formulae

$$x^{s+1} = x^s + \alpha \Delta x, \qquad u^{s+1} = u^s + \alpha \Delta u,$$
$$v^{s+1} = v^s + \alpha \Delta v.$$

The Newton-type descent direction is equal to

$$\begin{pmatrix} \Delta x \\ \Delta u \\ \Delta v \end{pmatrix} = -\nabla \Phi(x_s, u_s, v_s; \epsilon)^{-1} \Phi(x_s, u_s, v_s; \epsilon).$$

The nonsingular Jacobian matrix is as follow

$$\nabla \Phi(x, u, v; \epsilon) = \begin{pmatrix} A & E_{m\times m} & 0 \\ 0 & A^T & E_{n\times n} \\ V_{n\times n} & 0 & X_{n\times n} \end{pmatrix},$$

where $X = \operatorname{diag}(x_1, \ldots, x_n)$, by analogy with V.

For choosing step parameter α the standard procedures may be used similar to Armijo one. In a small neighborhood of a solution $\alpha = 1$. Starting points x^0 and v^0 must be positive. Point u^0 may be arbitrary, e. g. $u^0 = b - Ax^0$.

Note that (according to the implicit function theorem) the system (5) determines three smooth trajectories $x = x(\epsilon)$, $u = u(\epsilon)$, $v = v(\epsilon)$ as $\epsilon \in (0, \bar{\epsilon})$, and for them

$$\begin{pmatrix} \nabla x(\epsilon) \\ \nabla u(\epsilon) \\ \nabla v(\epsilon) \end{pmatrix} = \nabla \Phi(x, u, v; \epsilon)^{-1} \begin{pmatrix} 0 \\ c \\ 2\epsilon e \end{pmatrix} =$$

$$= \begin{pmatrix} (DM - E)Dc + 2\epsilon(E - DM)V^{-1}e \\ HADc - 2\epsilon HAV^{-1}e \\ (E - MD)c + 2\epsilon MV^{-1}e \end{pmatrix};$$

where $D = XV,^{-1}$ $M = A^T HA$, $H = (E + ADA^T)^{-1}$.

These three trajectories may be considered as the generalized central path for the improper program (1).

References

[1] Eremin I.I. *Duality for improper linear and convex programs* // Doklady of AN USSR. 1981. Vol. 256, N 2. P. 272–276. (in Russian)

[2] Eremin I.I., Mazurov Vl.D., Astaf'jev N.N. *Improper linear and convex mathematical programs.* Moscow.: Nauka, 1983. (in Russian)

[3] Eremin I.I. *Contradictory models of optimal planning.* Moscow: Nauka, 1988. (in Russian)

[4] Fiacco A.V., McCormick G.P. *Nonlinear Programming: Sequential Unconstrained Minimization Techniques.* John Wiley & Sons Ltd, 1969.

[5] Roos C., Terlaky T., Vial J.-Ph. *Theory and algorithms for linear optimization.* Chichester: John Wiley & Sons Ltd, 1997.