Computational Mathematics and Mathematical Physics, Vol. 44, No. 9, 2004, pp. 1484–1493.
Translated from Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki, Vol. 44, No. 9, 2004, pp. 1564–1573.

Application of Newton's Method for Solving Large Linear Programming Problems

A.I. Golikov, Yu.G. Evtushenko, and N. Mollaverdi

Computing Center, Russian Academy of Sciences, ul. Vavilova 40, Moscow, 119991 Russia
e-mail: gol@ccas.ru, evt@ccas.ru

Revised version

Abstract — To simultaneously solve the primal and dual linear programming (LP) problems, it is proposed to use a new auxiliary function that is similar to the modified Lagrangian function and then apply the generalized Newton's method to its unconstrained minimization. This approach is applicable to solving LP problems with a large number (up to several millions) of nonnegative variables and a moderate number (several thousands) of equality constraints. The results of test calculations on a Pentium-IV computer are presented. They show that the solution of problems of the size indicated above takes from several dozen to several thousand seconds.

Keywords: large linear programming problems, Newton's method, Lagrangian function

1. INTRODUCTION

It is well known that the exact solution to a primal linear programming (LP) problem can be found by using a smooth penalty function with a finite penalty coefficient applied to the dual LP problem (see, e.g., [1] – [3]). Such a function is convex, piecewise quadratic, and continuously differentiable, but it does not have the Hessian matrix. However, introducing the generalized Hessian matrix, one can construct the generalized Newton's method for this penalty function. In [4] – [6], the finite global convergence of the generalized Newton's method as applied to the minimization of a convex piecewise quadratic function was proved. The minimization of this penalty function applied to the dual LP problem makes it possible to obtain the exact normal solution (i.e., a solution with the minimal Euclidean norm) to the primal problem beginning from a certain finite value of the penalty coefficient.

In this paper, we propose to use an auxiliary function similar to the modified Lagrangian function (see, e.g., [7] – [9]) rather than a piecewise quadratic penalty function, which is traditionally used. This approach can be described as follows. Starting from a certain fixed value of the penalty coefficient obtained by the single unconstrained maximization of the auxiliary function, we calculate the exact projection of the given point on the solution set of the primal LP problem, using simple formulas provided by Theorem 1 below. Under a certain assumption, this theorem yields a formula for the threshold value of the penalty coefficient. Substituting the determined projection into the auxiliary function and then maximizing it, we obtain the exact solution to the dual LP problem (Theorem 2). Theorem 3 asserts that the iterative process described below produces the exact solutions to the primal and dual problems in a finite number of steps; this assertion is true for any penalty coefficient and any initial vector of the primal problem. The auxiliary maximization problem is solved using the generalized Newton's method, which converges for this problem in a finite number of steps.

The proposed method was implemented in Matlab 6.5 on a Pentium-IV computer with 1 Gb RAM. The numerical experiments with random LP problems demonstrated that the method is highly efficient in solving LP problems with a large number (up to several million) of nonnegative variables and a moderate number (several thousand) of equality constraints. The execution time for these problems was in the range from several dozen to several thousand seconds. These results can be explained by the fact that the basic computational effort in our method is consumed by solving the auxiliary unconstrained maximization problem. Its size is determined by the number of equality constraints, which is substantially less than the number of variables in the original LP problem.

2. BASIC THEOREMS

Let the primal LP problem be given in the canonical form :

$$f_* = \min_{x \in X} c^\top x, \qquad X = \{x \in \mathbb{R}^n : Ax = b, \ \ x \geq 0_n\}. \tag{P}$$

The problem dual to (P) has the form

$$f_* = \max_{u \in U} b^\top u, \qquad U = \{u \in \mathbb{R}^m : A^\top u \leq c\}. \tag{D}$$

Here, $A \in \mathbb{R}^{m \times n}$, $c \in \mathbb{R}^n$, and $b \in \mathbb{R}^m$ are given; x is the vector of primal variables; u is the vector of dual variables; and 0_i is the zero vector of dimension i. Assume that the solution set X_* of the primal problem (P) is nonempty; hence, the solution set U_* of the dual problem (D) also is nonempty. The necessary and sufficient optimality conditions (the Kuhn–Tucker conditions) for problems (P) and (D) have the form

$$Ax_* - b = 0_m, \qquad x_* \geq 0_n, \qquad x_*^\top v_* = 0, \tag{1}$$

$$v_* = c - A^\top u_* \geq 0_n. \tag{2}$$

Here, the nonnegative vector of auxiliary variables $v = c - A^\top u \geq 0_n$ is introduced into the constraints of the ratal problem (D).

Let $\hat{x}$ be an arbitrary vector in $\mathbb{R}^n$. We consider the problem of finding the projection $\hat{x}_*$ of $\hat{x}$ on the solution set X_* of the primal problem (P):

$$\frac{1}{2}\|\hat{x}_* - \hat{x}\|^2 = \min_{x \in X_*} \frac{1}{2}\|x - \hat{x}\|^2, \qquad X_* = \{x \in \mathbb{R}^n : Ax = b, \ \ c^\top x = f_*, \ \ x \geq 0_n\}. \tag{3}$$

Henceforth, we use the Euclidean norm of vectors.

Define the Lagrangian function for problem (3):

$$L(x, p, \beta, \hat{x}) = \frac{1}{2}\|x - \hat{x}\|^2 + p^\top (b - Ax) + \beta(c^\top x - f_*).$$

Here, $p \in \mathbb{R}^m$ and $\beta \in \mathbb{R}^1$ are the Lagrange multipliers, and $\hat{x}$ will be interpreted as a fixed vector parameter. The dual problem for (3) has the form

$$\max_{p \in \mathbb{R}^m} \ \max_{\beta \in \mathbb{R}^1} \ \min_{x \in \mathbb{R}^n_+} \ L(x, p, \beta, \hat{x}). \tag{4}$$

The Kuhn–Tucker conditions for problem (3) are as follows:

$$x - \hat{x} - A^\top p + \beta c \geq 0_n, \qquad D(x)(x - \hat{x} - A^\top p + \beta c) = 0_n, \qquad x \geq 0_n, \tag{5}$$

$$Ax = b, \qquad c^\top x = f_*, \tag{6}$$

where $D(z)$ denotes the diagonal matrix whose ith diagonal entry is the ith component of the vector z. It is easy to verify that formulas (5) are equivalent to

$$x = (\hat{x} + A^\top p - \beta c)_+, \tag{7}$$

where a_+ is obtained by replacing the negative components in a by zeros.

Formula (7) yields the solution to the inner minimization problem in (4). Substituting (7) into the lagrangian function $L(x, p, \beta, \hat{x})$, we obtain the dual function

$$\tilde{L}(p, \beta, \hat{x}) = b^\top p - \frac{1}{2}\|(\hat{x} + A^\top p - \beta c)_+\|^2 - \beta f_* + \frac{1}{2}\|\hat{x}\|^2.$$

The function $\tilde{L}(p, \beta, x)$ is concave, piecewise quadratic, and continuously differentiable. The dual problem (4) is reduced to solving the outer maximization problem

$$\max_{p\,\in\,\mathbb{R}^m} \ \max_{\beta\in\mathbb{R}^1} \tilde{L}(p, \beta, \hat{x}). \tag{8}$$

Having solved problem (8), we find the optimal p and β. Substituting them into (7), we obtain the projection $\hat{x}_*$, i.e., the solution to problem (3). The necessary and sufficient optimality conditions for problem (8) have the form

$$\begin{aligned} \tilde{L}_p(p, \beta, \hat{x}) &= b - A(\hat{x} + A^\top p - \beta c)_+ = b - Ax = 0_m, \\ \tilde{L}_\beta(p, \beta, \hat{x}) &= c^\top(\hat{x} + A^\top p - \beta c)_+ - f_* = c^\top x - f_* = 0, \end{aligned}$$

where x is defined by formula (7). These conditions are fulfilled if and only if $x \in X_*$ and $x = \hat{x}_*$.

Unfortunately, the unconstrained optimization problem (8) contains the a priori unknown quantity f_*, which is the optimal value of the objective function of the LP problem. However, one can avoid this drawback by simplifying problem (8). To this end, we propose replacing (8) by the following simplified unconstrained maximization problem:

$$I_1 = \max_{p\in\mathbb{R}^m} S(p, \beta, \hat{x}). \tag{9}$$

Here, $\hat{x}$ is a fixed vector, β is a fixed scalar, and the function $S(p, \beta, \hat{x})$ is defined by

$$S(p, \beta, \hat{x}) = b^\top p - \frac{1}{2}\|(\hat{x} + A^\top p - \beta c)_+\|^2. \tag{10}$$

Without loss of generality, one can assume that the first ℓ components of the vector $\hat{x}_*$ are strictly positive. Accordingly, we write the vectors $\hat{x}_*$, $\hat{x}$, c, and the matrix A as

$$\hat{x}_*^\top = \left[[\hat{x}_*^\ell]^\top, [\hat{x}_*^d]^\top\right], \quad \hat{x}^\top = \left[[\hat{x}^\ell]^\top, [\hat{x}^d]^\top\right], \quad c^\top = \left[[c^\ell]^\top, [c^d]^\top\right], \quad A = [A_\ell \mid A_d], \tag{11}$$

where $\hat{x}_*^\ell > 0_\ell$, $\hat{x}_*^d = 0_d$, and $d = n - \ell$.

Using this notation, we can rewrite the necessary and sufficient optimality conditions (5), (6) for problem (3) in a more detailed form

$$\hat{x}_*^\ell = \hat{x}^\ell + A_\ell^\top p - \beta c^\ell > 0_\ell, \tag{12}$$

$$\begin{gathered} \hat{x}_*^d = 0_d, \qquad \hat{x}^d + A_d^\top p - \beta c^d \le 0_d, \\ A_\ell \hat{x}_*^\ell = b, \qquad c^{\ell\top} \hat{x}_*^\ell = f_*. \end{gathered} \tag{13}$$

The system of linear equations in the unknown p appearing in (12) is consistent. Assuming that A_ℓ has the full rank m and $\ell \geq m$, we can find the unique solution p to this system by the formula

$$p = (A_\ell A_\ell^\top)^{-1} A_\ell (\hat{x}_*^\ell - \hat{x}^\ell + \beta c^\ell). \tag{14}$$

Substituting this expression into (13), we obtain the inequality

$$q \leq \beta z, \tag{15}$$

where we use the notation $q = \hat{x}^d + A_d^\top (A_\ell A_\ell^\top)^{-1} A_\ell (\hat{x}_*^\ell - \hat{x}^\ell)$ and $z = c^d - A_d (A_\ell A_\ell^\top)^{-1} A_\ell c^\ell$.

If p is defined by (14) and β satisfies inequality (15), then the pair $[p, \beta]$ is a solution to the dual problem (8). Let us find the minimal value of β satisfying (15).

In accordance with partition (11), we write the optimal vector of the auxiliary variables v_* in the Kuhn–Tucker conditions (1), (2) for problems (P) and (D) in the form $v_*^\top = [v_*^{\ell\top}, v_*^{d\top}]$. Then, by the complementary slackness condition $x_*^\top v_* = 0$, $x_* \geq 0_n$, $v_* \geq 0_n$, and expression (2) can be written as

$$\begin{aligned} v_*^\ell &= c^\ell - A_\ell^\top u_* = 0_\ell, && (16) \\ v_*^d &= c^d - A_d^\top u_* = 0_d. && (17) \end{aligned}$$

From (16), we obtain $u_* = (A_\ell A_\ell^\top)^{-1} A_\ell c^\ell$. Substituting this expression into (17), we find that $v_*^d = z \geq 0_d$. Now, we define the index set $\sigma = \{\ell + 1 \leq i \leq n : (v_*^d)^i > 0\}$.

If $\sigma = \emptyset$, then (15) is fulfilled for any β. Define

$$\beta_* = \begin{cases} \max\limits_{i \in \sigma} \dfrac{q^i}{(v_*^d)^i}, & \sigma \neq \emptyset, \\ \alpha > -\infty, & \sigma = \emptyset, \end{cases} \tag{18}$$

where α is an arbitrary scalar. Then, inequality (15) is valid for any $\beta \geq \beta_*$, and one can solve the simplified unconstrained maximization problem (9). Its solution yields, at the same time, a solution to the dual problem (8). Then, using formula (7), one obtains the projection $\hat{x}_*$. Thus, we have proved the following proposition.

Theorem 1. *Let the solution set X_* of problem* (P) *be nonempty. Assume that the matrix A_ℓ corresponding to the nonzero components of the vector $\hat{x}_*$ has rank m. Then, for any $\beta \geq \beta_*$, the projection $\hat{x}_*$ of the point $\hat{x}$ on the solution set X_* of the primal problem* (P) *is given by the formula*

$$\hat{x}_* = [\hat{x} + A^\top p(\beta) - \beta c]_+, \tag{19}$$

where $p(\beta)$ is the solution to the unconstrained maximization problem (9).

This theorem allows us to replace problem (8), containing the a priori unknown scalar f_*, by problem (9), where this scalar is replaced by the interval $[\beta_*, +\infty)$. Computationally, the latter problem is much simpler. The theorem extends the results obtained in [10], which are related to finding the normal solution to a primal LP problem (i.e., the projection of the zero on the solution set of problem (P)). It is obvious that the value β_* found by formula (18) can be negative. A corresponding example of projecting the origin is given in [10].

From the formal point of view, the unconstrained maximization problem (9) has no Lagrangian function; hence, the corresponding dual problem cannot be constructed. However, one can introduce auxiliary variables into problem (9) and use them to generate artificial constraints. This results in an equivalent nonlinear programming problem that already admits the dual problem. This (rather unconventional) approach to constructing the dual problem is based on a two-step representation of problem (9) (see, e.g., [11, 12]).

We define the vector of auxiliary variables $y = \hat{x} + A^\top p - \beta c$. Then, problem (9) reduces to the equivalent constrained minimization problem

$$I_1 = \max_{[p,y]\in G} \left\{ b^\top p - \frac{1}{2}\|y_+\|^2 \right\}, \qquad G = \{[p,y] \in \mathbb{R}^{m+n} : y = \hat{x} + A^\top p - \beta c\}. \tag{20}$$

The Lagrangian function for the quadratic programming problem (20) has the form

$$L(p,y,x) = b^\top p - \frac{1}{2}\|y_+\|^2 - x^\top(\hat{x} + A^\top p - \beta c - y),$$

where $x \in \mathbb{R}^n$ is the vector of Lagrange multipliers. Consider the minimax problem

$$I_2 = \min_{x\in\mathbb{R}^n} \max_{p\in\mathbb{R}^m} \max_{y\in\mathbb{R}^n} L(p,y,x), \tag{21}$$

which is dual to problem (20). From the optimality conditions for the inner maximization problem

$$L_p(p,y,x) = b - Ax = 0_m, \qquad L_y(p,y,x) = -y_+ + x = 0_n$$

we have $x = y_+$, $Ax = b$. Substituting $y_+ = x$ in $L(p,y,x)$ and taking into account the conditions $Ax = b$, $x \geq 0_n$, we obtain the dual function

$$\tilde{L}(x) = \beta c^\top x + \frac{1}{2}\|x - \hat{x}\|^2 - \frac{1}{2}\|\hat{x}\|^2.$$

Thus, the dual problem (21) is reduced to the quadratic programming problem

$$I_2 = \min_{x\in X} \left\{ \beta c^\top x + \frac{1}{2}\|x - \hat{x}\|^2 - \frac{1}{2}\|\hat{x}\|^2 \right\}, \qquad X = \{x \in \mathbb{R}^n : Ax = b, \ \ x \geq 0_n\}. \tag{22}$$

Since the objective function in problem (22) is strictly convex, its solution $x(\beta)$ is unique. Problem (22) is dual to problem (20) and hence to problem (9). By the duality theorem, the optimal values of the objective functions of these problems are the same: $I_1 = I_2$. This can be written as $b^\top p(\beta) = \beta f_* + x(\beta)^\top [x(\beta) - \hat{x}]$, where $p(\beta)$ is an arbitrary solution to problem (9). One can show in the standard way that the unconstrained maximization problem (9) for the piecewise quadratic function $S(p,\beta,\hat{x})$ defined by formula (10) is dual to the quadratic programming problem (22). Thus, the unconstrained maximization problem (9) and the quadratic programming problem (22) can be considered as mutually dual problems. Problem (22) can be considered as a perturbed or regularized problem (P). The solution $x(\beta)$ to problem (22) can be obtained from an arbitrary solution $p(\beta)$ to problem (9) by the formula

$$x(\beta) = [\hat{x} + A^\top p(\beta) - \beta c]_+,$$

which, according to Theorem 1, turns into formula (19) for $\beta \geq \beta_*$; i.e., in this case, we have $x(\beta) = \hat{x}_*$.

If the rank of the matrix A_ℓ is equal to m, then, by Theorem 1, β_* can be arbitrary when $\|v_*\| = 0$; in particular, it can be negative. If $q \leq 0_d$, then $\beta_* \leq 0$. If one sets $\beta = 0$, then the regularized problem (22) turns into the problem of finding the projection $\hat{x}_*$ of a given vector $\hat{x}$ on the feasible set X of problem (P). The vector $\hat{x}_*$ is at the same time a solution to problem (22) for any $\beta \geq \beta_*$ and a solution to problem (3). Therefore, in the case when $\beta_* \leq 0$, the distance from a given vector $\hat{x}$ to the solution set X_* of problem (P) is the same as the distance from $\hat{x}$ to the feasible set X of problem (P).

Consider the particular case of Theorem 1, where $\hat{x} = 0_n$. From (19), we obtain the following formula for the normal solution $\tilde{x}_*$ of the primal problem (P):

$$\tilde{x}_* = [A^\top p(\beta) - \beta c]_+.$$

Here, $p(\beta)$ is a solution to the unconstrained maximization problem

$$\max_{p \in \mathbb{R}^m} \left\{ b^\top p - \frac{1}{2}\|(A^\top p - \beta c)_+\|^2 \right\} \tag{23}$$

for any $\beta \geq \beta_*$. Problem (23) and the quadratic programming problem

$$\min_{x \in X} \left\{ \beta c^\top x + \frac{1}{2}\|x\|^2 \right\}, \qquad X = \{x \in \mathbb{R}^n : Ax = b, \;\; x \geq 0_n\} \tag{24}$$

are mutually dual. For any $\beta \geq \beta_*$, the unique solution $\tilde{x}_*$ to problem (24) is the projection of the zero on the solution set X_* of the primal problem (P); i.e., $\tilde{x}_*$ is the normal solution to problem (P) and can be determined from the problem

$$\min_{x \in X_*} \frac{1}{2}\|x\|^2, \qquad X_* = \{x \in \mathbb{R}^n : Ax = b, \;\; x \geq 0_n, \;\; c^\top x = f_*\}. \tag{25}$$

The following problem is mutually dual to problem (25):

$$\max_{p \,\in\, \mathbb{R}^m} \; \max_{\beta \in \mathbb{R}^1} \left\{ b^\top p - \frac{1}{2}\|(A^\top p - \beta c)_+\|^2 - \beta f_* \right\}. \tag{26}$$

The solution $\tilde{x}_*$ to the quadratic programming problem (25) can be obtained from an arbitrary solution $[p, \beta]$ to the unconstrained maximization problem (26) by the formula

$$\tilde{x}_* = (A^\top p - \beta c)_+.$$

Consider the case $\beta \geq \beta_*$ and $\beta > 0$. We change the variables in problem (23), setting $p = \beta u$. Then, (23) is replaced by the equivalent problem

$$\max_{u \in \mathbb{R}^m} \left\{ b^\top u - \frac{\beta}{2}\|(A^\top u - c)_+\|^2 \right\}, \tag{27}$$

i.e., we arrive at the method of exterior quadratic penalty applied to problem (D). In this case, Theorem 1 implies that the vector $u(\beta)$ obtained as a result of the maximization of the differentiable exterior penalty function in problem (26) determines the normal solution to problem (P) by the formula $\tilde{x}_* = \beta[A^\top u(\beta) - c]_+$ when $\beta > \beta_*$. It follows from the well-known properties of the method of exterior quadratic penalty (see [13]) that $u(\beta) = p(\beta)/\beta$ asymptotically tends to u_* as $\beta \to +\infty$. Problem (23) was introduced in [10, 6]. Since problem (23) is equivalent to (27), it can be considered as a new variant of the exterior-penalty method applied to the dual problem (D).

Thus, formula (18) for β_* yields bounds for the penalty coefficient in the classical method of exterior quadratic penalty applied to problem (D), i.e., an LP problem with inequality constraints. Using the solution $u(\beta)$ to problem (27), we obtain the normal solution to the primal problem (P) for any positive β if $\beta_* \leq 0$ and for any $\beta \geq \beta_*$ if $\beta_* > 0$.

Problem (27) is mutually dual to the following regularized LP problem (see, e.g., [14, 15, 1]):

$$\min_{x \in X} \left\{ c^\top x + \frac{1}{2\beta}\|x\|^2 \right\}.$$

Let β_* be defined as in Theorem 1. If $\beta_* > 0$, then we set $\varepsilon_* = 1/\beta_*$. If $\beta_* \leq 0$, then β_* can be set to any positive number. Thus, we arrive at a bound for the regularization parameter ε in the classical Tikhonov regularization of the LP problem:

$$\min_{x\in X} \left\{ c^\top x + \frac{\varepsilon}{2}\|x\|^2 \right\}. \tag{28}$$

For $0 < \varepsilon \leq \varepsilon_*$, the solution to problem (28) and the normal solution to problem (P) are the same.

The point $p(\beta)$, which is a maximizer of the function $S(p, \beta, \hat{x})$, does not give a solution to the dual problem (D) if $\hat{x} \notin X_*$. However, when $\beta \geq \beta_*$, formula (19) yields the exact solution $\hat{x}_*$ to problem (3) (which is the projection of $\hat{x}$ on the solution set X_* of the primal problem (P)) and the exact normal solution to problem (P) in the case $\hat{x} = 0_n$.

The next theorem asserts that if a point $x_* \in X_*$ is available, then one can obtain a solution to the dual problem (D) after solving the unconstrained maximization problem (9) one time.

Theorem 2. *Let the solution set X_* of problem* (P) *be nonempty. Then, for any $\beta > 0$ and any $\hat{x} = x_* \in X_*$, the exact solution to the dual problem* (D) *is given by the formula $u_* = p(\beta)/\beta$, where $p(\beta)$ is a maximizer of the function $S(p, \beta, x_*)$.*

Proof. The necessary and sufficient optimality conditions for problem (9) have the form

$$b - A(x_* + A^\top p_* - \beta c)_+ = 0_m. \tag{29}$$

Define $x = (x_* + A^\top p_* - \beta c)_+$. This expression is equivalent to the following three vector relations:

$$x - (x_* + A^\top p_* - \beta c) \geq 0_n, \qquad x \geq 0_n, \tag{30}$$

$$D(x)[x - (x_* + A^\top p_* - \beta c)] = 0_n. \tag{31}$$

Take x_* as the vector x and βu_* as the vector p_*. Then, from (29) – (31), we obtain, respectively,

$$Ax_* = b, \qquad c - A^\top u_* \geq 0_n, \qquad x_* \geq 0_n, \qquad c^\top x_* = b^\top u_*.$$

It follows that $[x_*, u_*]$ is a Kuhn–Tucker point for the LP problem and u_* is a solution to the dual problem (D). The theorem is proved. $\square$

Function (10) can be considered as the modified Lagrangian function for the dual problem (D). We define the following iterative process:

$$\begin{aligned} p_{k+1} &\in \arg\max_{p\in\mathbb{R}^m} \left\{ b^\top p - \frac{1}{2}\|(x_k + A^\top p - \beta c)_+\|^2 \right\}, && (32)\\ x_{k+1} &= (x_k + A^\top p_{k+1} - \beta c)_+, && (33) \end{aligned}$$

where x_0 is an arbitrary initial vector.

This is a finite process, which yields both the exact solution x_* to the primal problem (P) and the exact solution u_* to the dual problem (D).

Theorem 3. *Let the solution set X_* of the primal problem* (P) *be nonempty. Then, for any $\beta > 0$ and any initial vector x_0, the iterative process* (32), (33) *converges to $x_* \in X_*$ in a finite number of steps ω. The formula $u_* = p_{\omega+1}/\beta$ yields the exact solution to the dual problem* (D).

Changing the variables to $p = \beta u$ in (32), (33), we arrive at the method of the modified Lagrangian function proposed in [7, 8] for solving the LP problem:

$$u_{k+1} \in \arg\max_{u\in\mathbb{R}^m} \left\{ \beta b^\top u - \frac{1}{2}\left\|[x_k + \beta(A^\top u - c)]_+\right\|^2 \right\}, \tag{34}$$

$$x_{k+1} = [x_k + \beta(A^\top u_{k+1} - c)]_+. \tag{35}$$

The finite convergence of method (34), (35) was proved in [8]. This proof can be adapted to method (32), (33) in an obvious way.

Note that $x_\omega = x_* \in X_*$ is a projection of $x_{\omega-1}$ on the solution set X_* of problem (P).

3. THE GENERALIZED NEWTON'S METHOD

The unconstrained maximization in (9) and (32) can be performed by any method, say, by the conjugate gradient method. However, O. Mangasarian showed that the unconstrained optimization of a piecewise quadratic function can be most efficiently performed by the generalized Newton's method (see [4, 5]). We give a brief description of this method and numerical results.

The objective functions $S(p, \beta, x_k)$ and $S(p, \beta, \hat{x})$ of problems (32) and (9), respectively, are concave, piecewise quadratic, and differentiable. These functions have no conventional Hessian matrices. Indeed, the gradient

$$S_p(p, \beta, x_k) = b - A(x_k + A^\top p - \beta c)_+$$

of $S(p, \beta, x_k)$ is not differentiable. However, for this function, one can define the generalized Hessian matrix which is an m-by-m symmetric negative semidefinite matrix of the form

$$\partial_p^2 S(p, \beta, x_k) = -AD^\sharp(z)A^\top.$$

Here, $D^\sharp(z)$ denotes the n-by-n diagonal matrix whose ith diagonal entry z^i is equal to one if $(x_k + A^\top p - \beta c)^i > 0$ and z^i is equal to zero if $(x_k + A^\top p - \beta c)^i \le 0$ ($i = 1, 2, \ldots, n$). Since the generalized Hessian matrix can be singular, the following modified Newton direction is used:

$$-[\partial_p^2 S(p, \beta, x_k) - \delta I_m]^{-1} S_p(p, \beta, x_k),$$

where δ is a small positive number (in our computations, we typically set $\delta = 10^{-4}$) and I_m is the identity matrix of order m.

In this case, the modified Newton's method has the form

$$p_{s+1} = p_s - [\partial_p^2 S(p_s, \beta, x_k) - \delta I_m]^{-1} S_p(p_s, \beta, x_k). \tag{36}$$

We used the following stopping criterion for this method:

$$\|p_{s+1} - p_s\| \le \text{tol}.$$

Mangasarian studied the convergence of the generalized Newton's method as applied to the unconstrained minimization of a convex piecewise quadratic function of this type with the steplength chosen by the Armijo rule. The proofs of the finite global convergence of the generalized Newton's method can be found in [4] – [6].

4. RESULTS OF NUMERICAL CALCULATIONS

We solved random LP problems with a large number (up to several million) of nonnegative variables and a moderate number (up to several thousand) of equality constraints; i.e., it held that $n \gg m$.

Thus, the prescribed quantities were the integers m and n, determining the number of rows and columns, respectively, in the matrix A, and the density ρ of nonzero entries in A. In particular, $\rho = 1$ means that all the entries in A were generated as random numbers, whereas $\rho = 0.01$ indicates that only one percent of the entries in A were generated randomly and the others were set to zero. The random entries in A were taken from the interval $[-50, 50]$. The solution x_* to the primal problem (P) and the solution u_* to the dual problem (D) were generated as follows. It was assumed that $n - 3m$ components of x_* were zero, while the rest were taken randomly from the interval $[0, 10]$. Half of the components of u_* were set to zero, while the rest were taken randomly from the interval $[-10, 10]$. The solutions x_* and u_* were used to calculate the coefficients in the objective function c and the right-hand sides b of problem (P). The vectors b and c were defined by the formulas

$$b = Ax_*, \qquad c = A^\top u_* + \xi.$$

Here, $\xi^i = 0$ if $x^{*i} > 0$, whereas, if $x^{*i} = 0$, then the component ξ^i was taken randomly from the interval

$$0 \leq \gamma^i \leq \xi^i \leq \theta^i.$$

In the calculations whose results are shown below, we set $\gamma^i = 1$ and $\theta^i = 10$ for all i. Note that, if γ^i is small, then the quantity $\xi^i = (c - A^\top u_*)^i = (v^d_*)^i$ can also be very small. Then, according to formula (18), the a priori unknown quantity β_* can be very large. In this case, the generated LP problem can be hard to solve.

The proposed method for solving the primal and dual LP problems, which combines the iterative process (32), (33) and the generalized Newton's method, was implemented in Matlab 6.5. The calculations were conducted on a 2.6 GHz Pentium-IV computer with memory of 1 Gb. The results of solving the randomly generated LP problems are presented in the table, where m and n indicate the size of the problem, ρ is the density of the nonzero entries in A, T is the run time (in seconds) of solving an LP problem, and It is the number of iteration steps of Newton's method in the first solution of the maximization problem (32). The fourth, fifth, and sixth columns in the table show, respectively, the accuracy to which the constraints of the primal problem are fulfilled, the similar quantity for the dual problem, and the difference between the optimal values of the objective functions of the primal and dual problems. The matrix A was partitioned into B blocks to improve the performance in calculating the matrix product $AD^\sharp(z)A^\top$ or, in the event of an especially large-scale problem, just because of the shortage of memory; then, multiplication was performed blockwise. The corresponding number of blocks B is shown in the last column of the table.

In all the examples, the zero was taken as the initial vector for the iterative process (32), (33): $x_0 = 0_n$. We always set $\beta = 1$, tol $= 10^{-12}$. In all occasions, it turned out that $\beta > \beta_*$. Thus, the normal solution $\hat{x}_*$ to the primal problem (P) was obtained by a single iteration step of process (32), (33), i.e., with $\omega = 1$. The number of iteration steps of the generalized Newton's method is shown in the third column of the table. By Theorem 2, the maximization of the function $S(p, \beta, \hat{x}_*)$ with respect to p yields the vector $p(\beta)$, which is equal to $u_*\beta$. In all the examples, only two iteration steps of the generalized Newton's method (36) were required for this maximization.

Table

$m \times n \times p$	T, c	It	$\|A\hat{x}_* - b\|$	$\|(A^\top u_* - c)_+\|$	$\|c^\top \hat{x}_* - b^\top u_*\|$	B
$100 \times 10^6 \times 0.01$	29.3	17	1.7×10^{-11}	2.0×10^{-13}	9.7×10^{-11}	5
$300 \times 10^6 \times 0.01$	42.0	13	1.0×10^{-10}	7.0×10^{-13}	2.6×10^{-10}	5
$600 \times 10^6 \times 0.01$	68.4	12	3.1×10^{-10}	1.7×10^{-12}	2.8×10^{-10}	5
$1000 \times 10^6 \times 0.01$	95.8	10	9.4×10^{-10}	3.5×10^{-12}	6.9×10^{-10}	8
$3000 \times 10^4 \times 0.01$	81.5	7	2.0×10^{-9}	9.1×10^{-12}	3.7×10^{-9}	2
$4000 \times 10^4 \times 0.01$	196.2	8	2.9×10^{-9}	1.2×10^{-11}	2.6×10^{-8}	2
$500 \times (3 \times 10^6) \times 0.01$	179.1	12	3.2×10^{-10}	1.4×10^{-12}	1.9×10^{-11}	8
$1000 \times (3 \times 10^6) \times 0.01$	309.1	11	1.2×10^{-9}	4.1×10^{-12}	4.9×10^{-9}	10
$500 \times (5 \times 10^6) \times 0.01$	300.8	12	3.8×10^{-10}	1.6×10^{-12}	8.4×10^{-11}	10
$1000 \times (5 \times 10^6) \times 0.01$	412.8	8	7.3×10^{-9}	7.4×10^{-12}	7.0×10^{-8}	100
$500 \times 10^7 \times 0.01$	387.8	8	7.6×10^{-9}	3.6×10^{-12}	1.1×10^{-7}	400
$1000 \times 10^4 \times 1$	117.2	7	1.3×10^{-7}	1.0×10^{-10}	2.9×10^{-7}	2
$1000 \times 10^5 \times 1$	1496.5	5	5.2×10^{-7}	1.9×10^{-10}	8.2×10^{-7}	200
$100 \times 10^6 \times 1$	376.5	9	4.2×10^{-8}	1.2×10^{-11}	3.0×10^{-7}	500

For the most part, the table presents the results of solving problems with a sparse matrix A. As an exception, the last three rows show the results for problems with completely randomly generated matrices.

The results in the table demonstrate the high efficiency of the proposed method. For instance, an LP problem with five million nonnegative variables and one thousand equality constraints was solved to a good accuracy in less than seven minutes (see the tenth row of the table). Note that only Matlab's facilities were exploited for the computer implementation of our method.

For LP problems with a large number of nonnegative variables and a moderate number of equality constraints, the computer implementation of the proposed method outperformed the packages based on the simplex method for LP problems that were available to the authors.

The authors see the possibility for further progress of the proposed method in the use of parallel computation for implementing the generalized Newton's method. This should make possible solving LP problems with a larger number of equality constraints.

ACKNOWLEDGMENTS

This study was supported by the Russian Foundation for Basic Research (project no. 03-01-00465), by the Program of the State Support of Leading Scientific Schools (project no. NSh–1737.2003.1), and by Program no. 17 of the Presidium of the Russian Academy of Sciences. We are indebted to S.S. Arsen'ev–Obraztsov, A.A. Stanevichyus, and A.V. Tyulenev for helpful discussions and assistance in conducting numerical calculations.

REFERENCES

1. I.I. Eremin, *Linear Optimization Theory* (Ural'skoe Otdelenie Ross. Akad. Nauk, Yekaterinburg, 1998) [in Russian].

2. B.T. Polyak, *Introduction to Optimization* (Nauka, Moscow, 1980) [in Russian].

3. B.S. Razumikhin, *Physical Models and Equilibrium Techniques in Programming and Economics* (Nauka, Moscow, 1975) [in Russian].

4. O.L. Mangasarian, "A Finite Newton Method for Classification", Optimizat. Meth. Software **17**, 913–930 (2002).

5. O.L. Mangasarian, *A Newton Method for Linear Programming*, Data Mining Inst., Techn. Report No. 02-02, 2002.

6. C. Kanzow, H. Qi, and L. Qi, "On the Minimum Norm Solution of Linear Programs", J. Optimizat. Theory Appl. **116**, 333–345 (2003).

7. B.T. Polyak and N.V. Tret'yakov, "On an Iterative Linear Programming Method and Its Economic Interpretation", Ekon. Mat. Metody **8** (5), 740–751 (1972).

8. A.S. Antipin, *Nonlinear Programming Methods Based on the Primal and Dual Modification of the Lagrange Function* (Vsesoyuzn. Nauchno-Issledovatel'skii Inst. Sistemnykh Issledovanii, Moscow, 1979) [in Russian].

9. Yu.G. Evtushenko and A.I. Golikov, "The Augmented Lagrangian Function for the Linear Programming Problems", in *Dynamics of Non-homogeneous Systems* (Inst. System Analysis, Russian Academy of Sciences, Moscow, 1999), vol. 2, pp. 63–67.

10. A.I. Golikov and Yu.G. Evtushenko, "Search for Normal Solutions in Linear Programming Problems", Zh. Vychisl. Mat. Mat. Fiz. **40**, 1766–1786 (2000) [Comp. Math. Math. Phys. **40**, 1694–1714 (2000)].

11. I.I. Eremin, "On Quadratic and Fully Quadratic Convex Programming Problems", Izv. Vyssh. Uchebn. Zaved., Mat. No. 12, 22–28 (1998).

12. A.I. Golikov and Yu.G. Evtushenko, "Theorems of the Alternative and Their Applications in Numerical Methods", Zh. Vychisl. Mat. Mat. Fiz. **43**, 354–375 (2003) [Comp. Math. Math. Phys. **43**, 338–358 (2003)].

13. A. Fiacco and G. McCormic, *Nonlinear Programming: Sequential Unconstrained Minimization Techniques* (Wiley, New York, 1968; Mir, Moscow, 1972).

14. A.N. Tikhonov, *Selected Works* (MAKS, Moscow, 2001) [in Russian].

15. F.P. Vasil'ev and A.Yu. Ivanitskii, *Linear Programming* (Faktorial, Moscow, 2003) [in Russian].