Optimization Methods and Software
Vol. 20, Nos. 4–5, August–October 2005, 515–524

Augmented Lagrangian method for large-scale linear programming problems

Yu.G. EVTUSHENKO[1], A.I. GOLIKOV and N. MOLLAVERDY

Dorodnicyn Computing Centre, Russian Academy of Sciences, Vavilov 40, Moscow 119991, Russia

The augmented Lagrangian and Newton methods are used to simultaneously solve the primal and dual linear programming problems. The proposed approach is applied to the primal linear programming problem with a very large number ($\approx 10^6$) of nonnegative variables and a moderate ($\approx 10^3$) number of equality-type constraints. Computation results such as the solution of a linear programme with 10 million primal variables are presented.

Keywords: Linear programming problem; Augmented Lagrangian; Newton method

1. Introduction

To solve the primal linear programming problem (LP), Mangasarian [1] proposed to use the exterior penalty function of its dual. This function is piecewise quadratic, convex, and differentiable, and a generalized Hessian of this function exists everywhere. These properties enabled him to apply the generalized Newton method to unconstrained minimization. The finite global convergence of the generalized Newton method was established in refs. [1]–[3]. The minimization of the exterior penalty function provided an exact least two-norm solution to the primal problem for a finite value of the penalty parameter.

In this paper, we propose to use an approach close to the augmented Lagrangian technique (see, e.g., refs. [4]–[7]). The approach involved has the following main advantage: after a single unconstrained maximization of the dual function which is similar to the augmented Lagrangian, we obtain the exact projection of a point onto the solution set of primal LP problem. The dual function has a parameter (similar to the penalty coefficient) which must exceed or be equal to some threshold value. This value is found under the regularity condition (Theorem 2.1). Using this result, we maximize once more the dual function with changed Lagrangian multipliers and obtain the exact solution of the dual LP problem (Theorem 2.2). Theorem 3.1 states that the exact primal and dual solutions of the LP problem can be obtained in a finite number of iterations with an arbitrary positive value of the parameter. The auxiliary unconstrained maximization problems are solved by the fast generalized Newton method.

The proposed approach was applied to primal LP problems with a very large number ($\approx 10^6$) of nonnegative variables and a moderate ($\approx 10^3$) number of equality type constraints. The results of computational experiments are given subsequently. The effectiveness of the proposed algorithm is demonstrated by comparing it with several LP solvers on a class of synthetically generated large-scale linear programmes.

2. Finding a projection onto the primal solution set

Consider the primal linear programme in the standard form

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

[1]Corresponding author. Email: evt@ccas.ru

together with its dual

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

where $A \in \mathbb{R}^{m\times n}$, $c \in \mathbb{R}^n$, and $b \in \mathbb{R}^m$ are given, x is a primal variable and u is a dual variable; 0_i denotes the ith dimensional zero vector.

Let us assume that the solution set X_* of the primal problem (P) is nonempty, hence, the solution set U_* of the dual problem (D) is also nonempty.

Let $\hat{x} \in \mathbb{R}^n$ be an arbitrary vector. Consider the problem of finding the least two-norm projection $\hat{x}_*$, of the point $\hat{x}$ on X_*

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

Henceforth, $\|a\|$ denotes the Euclidian norm of a vector a.

The solution $\hat{x}_*$ of problem (1) is unique. Let us introduce the Lagrangian function for problem (1)

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

where $p \in \mathbb{R}^m$ and $\beta \in \mathbb{R}^1$ are Lagrangian multipliers, and $\hat{x}$ is considered to be a fixed parameter vector. The dual problem of (1) is

$$\max_{p\,\in\,\mathbb{R}^m}\ \max_{\beta\,\in\,\mathbb{R}^1}\ \min_{x\,\in\,\mathbb{R}^n_+}\ L(x,p,\beta,\hat{x}). \tag{2}$$

The Kuhn–Tucker conditions for problem (1) imply the existence of $p \in \mathbb{R}^m$ and $\beta \in \mathbb{R}^1$ such that

$$\begin{gathered} x - \hat{x} - A^\top p + \beta c \ge 0_n, \qquad D(x)(x - \hat{x} - A^\top p + \beta c) = 0_n, \\ x \ge 0_n, \qquad Ax = b, \qquad c^\top x = f_*, \end{gathered} \tag{3}$$

where $D(z)$ denotes the diagonal matrix whose ith diagonal element is the ith component of the vector z. It is easy to verify that (3) is equivalent to

$$x = (\hat{x} + A^\top p - \beta c)_+, \tag{4}$$

where a_+ denotes the vector a with all the negative components replaced by zeros.

We can say that (4) gives us the solution of the inner minimization problem in (2). By substituting (4) into $L(x,p,\beta,\hat{x})$, we obtain the dual function

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

Hence, problem (2) is reduced to the solution of the exterior maximization problem

$$\max_{p\,\in\,\mathbb{R}^m}\ \max_{\beta\,\in\,\mathbb{R}^1}\ \tilde{L}(p,\beta,\hat{x}). \tag{5}$$

If the solutions p and β of problem (5) are found, then after substitution of p and β into (4), we obtain the projection $\hat{x}_*$ which solves problem (1).

The optimality conditions for problem (5) are the following

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

where x is given by (4). These conditions are satisfied if and only if $x \in X_*$ and $x = \hat{x}_*$.

Unfortunately, the unconstrained optimization problem (5) contains an unknown value f_*. It is possible to simplify this problem and avoid this shortcoming. We show that if the value β is chosen large enough, then the minimization over the variable β can be dispensed with. Instead of problem (5), we propose to solve the following simplified unconstrained maximization problem

$$\max_{p \in \mathbb{R}^m} S(p, \beta, \hat{x}), \tag{6}$$

where $\hat{x}$ and β are fixed and the function $S(p, \beta, \hat{x})$ is given by

$$S(p, \beta, \hat{x}) = b^\top p - \frac{1}{2} \|(\hat{x} + A^\top p - \beta c)_+\|^2. \tag{7}$$

Without loosing generality, one can assume that the first l components of $\hat{x}_*$ are strictly greater than zero. In accordance with this assumption, we represent vectors $\hat{x}_*$, $\hat{x}$, and c, as well as matrix A in the form

$$\hat{x}_*^\top = \left[[\hat{x}_*^l]^\top, [\hat{x}_*^d]^\top\right], \qquad \hat{x}^\top = \left[[\hat{x}^l]^\top, [\hat{x}^d]^\top\right], \qquad c^\top = \left[[c^l]^\top, [c^d]^\top\right], \qquad A = [A_l \mid A_d],$$

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

The necessary and sufficient optimality conditions (the Kuhn–Tucker conditions) for problem (1) can be written in the expanded form

$$\hat{x}_*^l = \hat{x}^l + A_l^\top p - \beta c^l > 0_l, \tag{8}$$

$$\hat{x}_*^d = 0_d, \qquad \hat{x}^d + A_d^\top p - \beta c^d \leq 0_d, \tag{9}$$

$$A_l \hat{x}_*^l = b, \qquad {c^l}^\top \hat{x}_*^l = f_*.$$

The linear system (8) is consistent; therefore, if we assume that the submatrix A_l has a full rank m, then the unique solution p of (8) is given by

$$p = (A_l A_l^\top)^{-1} A_l (\hat{x}_*^l - \hat{x}^l + \beta c^l). \tag{10}$$

Substituting this function into (9), we obtain the inequality

$$q \leq \beta z, \tag{11}$$

where $q = \hat{x}^d + A_d^\top (A_l A_l^\top)^{-1} A_l (\hat{x}_*^l - \hat{x}^l)$, $z = c^d - A_d^\top (A_l A_l^\top)^{-1} A_l c^l$.

If p is given by function (10) and β satisfies inequality (11), then this pair $[p, \beta]$ is the solution of the dual problem (5). So, here we are able to derive an analytical formula for the minimal value β_* which satisfies inequality (11). Consider the Kuhn–Tucker optimality conditions for the primal problem (P). Besides the primal feasibility, we have a complementarity condition $\hat{x}_*^\top v_* = 0$, where the dual slack $v_* \in \mathbb{R}^n_+$ and the following dual feasibility conditions have the form

$$v_*^l = c^l - A_l^\top u_* = 0_l, \tag{12}$$

$$v_*^d = c^d - A_d^\top u_* \geq 0_d. \tag{13}$$

From (12), we find $u_* = (A_l A_l^\top)^{-1} A_l c^l$, and the substitution of this expression in inequality (13) yields $z = v_*^d \geq 0_d$. We define the index set as follows: $\sigma = \{1 \leq i \leq d : (v_*^d)^i > 0\}$. Inequality (11) holds if $\beta \geq \beta_*$, where

$$\beta_* = \begin{cases} \max\limits_{i \in \sigma} \dfrac{q^i}{(v_*^d)^i}, & \text{if } \sigma \neq \emptyset, \\ \alpha > -\infty, & \text{if } \sigma = \emptyset, \end{cases} \tag{14}$$

and α is an arbitrary number.

If $\beta \geq \beta_*$, then we can solve the simplified unconstrained maximization problem (6). Its solution will be simultaneous to that of the dual problem (5). Hence, using formula (4) we obtain $\hat{x}_*$. Therefore, we conclude that the following theorem holds.

Theorem 2.1. *Assume that the solution set X_* for problem* (P) *is nonempty, the rank of a submatrix A_l corresponding to nonzero components of vector $\hat{x}_*$ is m. Then, for all $\beta \geq \beta_*$, the unique least two-norm projection $\hat{x}_*$ of a point $\hat{x}$ onto X_* is given by*

$$\hat{x}_* = [\hat{x} + A^\top p(\beta) - \beta c]_+, \tag{15}$$

where $p(\beta)$ is a point attaining the maximum in problem (6).

Theorem 2.1 generalizes the results obtained in ref. [8] devoted to finding a normal solution to the primal LP problem. It is obvious that the value of β_* defined by (14) may be negative. The corresponding very simple example is given in ref. [8].

The function $S(p, \beta, \hat{x})$, where $\hat{x} = 0_n$, can be considered as a new asymptotic exterior penalty function of the dual linear programme (D) [3, 8]. The point $p(\beta)$ which maximizes $S(p, \beta, \hat{x})$ does not solve the dual LP problem (D) for finite β, but the ratio $p(\beta)/\beta \to u_*$ as $\beta \to \infty$. If $\beta \geq \beta_*$, then formula (15) provides the exact solution $\hat{x}_*$ to problem (1) (the projection of $\hat{x}$ onto the solution set X_* of the original primal linear programme (P)), and if $\hat{x} = 0_n$, then, we obtain the exact normal solution of (P).

Formally, the unconstrained maximization problem (6) has no Lagrangian function, which implies that the corresponding dual problems cannot be constructed directly. Nevertheless, one can introduce additional variables to construct the artificial constraints and obtain the equivalent nonlinear programming problems for which the dual problems are well defined. This assertion is not quite conventional, it is based on the two-step representation of problem (6).

We introduce a vector of additional variables $y = \hat{x} + A^\top p - \beta c$. Then, problem (6) reduces to the equivalent constrained maximization problem

$$\begin{aligned} I_1 &= \max_{[p,y] \in G} \left\{ b^\top p - \frac{1}{2} \|y_+\|^2 \right\}, \\ G &= \left\{ [p, y] \in \mathbb{R}^{m+n} : y = \hat{x} + A^\top p - \beta c \right\}. \end{aligned} \tag{16}$$

The Lagrangian function for the quadratic programming problem (16) is

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

where the multiplier vector $x \in \mathbb{R}^n$. We introduce the corresponding min max problem

$$\min_{x \in \mathbb{R}^n} \max_{p \in \mathbb{R}^m} \max_{y \in \mathbb{R}^n} L(p, y, x).$$

The solution of the interior maximization problem is $x = y_+$, $Ax = b$.

We substitute these results into $L(p, y, x)$ and obtain the following quadratic problem

$$\begin{aligned} I_2 &= -\frac{1}{2} \|\hat{x}\|^2 + \min_{x \in X} \left\{ \beta c^\top x + \frac{1}{2} \|x - \hat{x}\|^2 \right\}, \\ X &= \{x \in \mathbb{R}^n : Ax = b, \ x \geq 0_n\}. \end{aligned} \tag{17}$$

As the objective function of problem (17) is strongly convex, its solution is unique. Problem (17) is dual to problem (16) and to a certain extent (6). Hence, the unconstrained maximization problem (6) and the quadratic problem (17) can be interpreted as being mutually dual. The duality theorem states that the optimal values of the objective functions are equal: $I_1 = I_2$.

The solution of problems (1) and (17) is connected with the solution of problem (6) by formula (15) if $\beta \geq \beta_*$. Problem (17) was analysed by Udzawa, Tikhonov, Eremin, Polyak, Mangasarian and others. The replacement of the primal problem (P) by problem (17) is called "Tikhonov regularization". In fact, Tikhonov used $\varepsilon = 1/\beta$ instead of β. Our variant of regularization is preferable because it provides us the possibility to consider the cases where $\beta_* \leq 0$.

The proposed approach can be applied to the search for a projection $\hat{x}_*$ of vector $\hat{x}$ onto the primal feasible set X. Setting $c = 0_m$ in the previous formulas, we obtain the following maximization problem with parameter-free objective function

$$\max_{p \in \mathbb{R}^m} \left\{ b^\top p - \frac{1}{2} \|(\hat{x} + A^\top p)_+\|^2 \right\}. \tag{18}$$

If p_* is a solution of this problem, then the two-norm projection $\hat{x}_*$ of vector $\hat{x}$ onto the set X is given by

$$\hat{x}_* = (\hat{x} + A^\top p_*)_+. \tag{19}$$

If we set $b = 0_m$ and solve (18), then formula (19) yields the projection of vector $\hat{x}$ onto the intersection of null-space of matrix A and the nonnegative orthant in $\mathbb{R}^n$.

If the rank of the submatrix A_l is equal to m and if $\|v_*\| = 0$ or $q \leq 0_d$, then $\beta_* \leq 0$. In this case, setting $\beta = 0$, we transform the regularized problem (17) into the problem of determining the projection of vector $\hat{x}$ onto the feasible set X of problem (P). At the same time, according to Theorem 2.1, vector $\hat{x}_*$ is a solution to problems (1) and (17) for any $\beta \geq \beta_*$. Hence, if $\beta_* \leq 0$, then the distance between $\hat{x}$ and X_* coincides with that between $\hat{x}$ and X.

The next Theorem tells us that we can get a solution to problem (D) from the single unconstrained maximization problem (6), if a point $x_* \in X_*$ is known.

Theorem 2.2. *Assume that the solution set X_* of problem* (P) *is nonempty. Then, for all $\beta > 0$ and $\hat{x} = x_* \in X_*$, an exact solution of the dual problem* (D) *is given by $u_* = p(\beta)/\beta$, where $p(\beta)$ is a point attaining the maximum of $S(p, \beta, x_*)$.*

Proof. The necessary and sufficient optimality condition for problem (6) is

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

If we denote $x = (x_* + A^\top p_* - \beta c)_+$, then this expression is equivalent to

$$\begin{gathered} x - (x_* + A^\top p_* - \beta c) \geq 0_n, \qquad x \geq 0_n, \\ D(x)[x - (x_* + A^\top p_* - \beta c)] = 0_n. \end{gathered}$$

These conditions imply that if $u_* = p_*/\beta$, then (x_*, u_*) is a Kuhn–Tucker pair for LP problems (P) and (D). Indeed substituting (x_*, u_*) for $(x, p_*/\beta)$ in the above expressions, we obtain

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

which proves the theorem. $\square$

Hence, when Theorem 2.1 is used and the point $\hat{x}_* \in X_*$ is found, then Theorem 2.2 provides a very effective and simple tool for solving the dual problem (D). An exact solution to (D) can be obtained by only one unconstrained maximizing the function $S(p, \beta, \hat{x}_*)$ with arbitrary $\beta > 0$.

3. Iterative process for solving primal and dual LP problems

In this section, we search for the arbitrary solutions $x_* \in X_*$ and $u_* \in U_*$ instead of the projection $\hat{x}_*$. Owing to this simplification, the iterative process proposed subsequently does not require the knowledge of the threshold value β_*.

Function (7) can be considered as an augmented Lagrangian for the linear programme (D) [4]–[6]. Let us introduce the following iterative process (the augmented Lagrangian method for the dual LP problem (D))

$$p_{k+1} \in \arg\max_{p \in \mathbb{R}^m} \left\{ b^\top p - \frac{1}{2} \|(x_k + A^\top p - \beta c)_+\|^2 \right\}, \tag{20}$$

$$x_{k+1} = (x_k + A^\top p_{k+1} - \beta c)_+, \tag{21}$$

where x_0 is an arbitrary starting point.

Theorem 3.1. *Assume that the solution set X_* of problem* (P) *is nonempty. Then, for all $\beta > 0$ and an arbitrary starting point x_0, the iterative processes* (20) *and* (21) *converge to $x_* \in X_*$ in a finite number of iterations ω. The formula $u_* = p^{\omega+1}/\beta$ gives an exact solution of the dual problem* (D).

Let us introduce a new variable $u = p/\beta$. Then formulas (20) and (21) lead to the following iterative process proposed by Antipin in ref. [5]:

$$u_{k+1} \in \arg\max_{u \in \mathbb{R}^m} \left\{ \beta b^\top u - \frac{1}{2} \left\| [x_k + \beta(A^\top u - c)]_+ \right\|^2 \right\}, \tag{22}$$

$$x_{k+1} = [x_k + \beta(A^\top u_{k+1} - c)]_+. \tag{23}$$

The proof of Theorem 3.1 is similar to the one given in ref. [5].

4. Generalized Newton method

Unconstrained maximization in problems (6) and (20) can be carried out by the conjugate gradient method or by other iterative methods. Following refs. [1]–[3], we utilize the generalized Newton method for solving this problem.

The maximized functions $S(p, \beta, x_k)$ in problem (20) and $S(p, \beta, \hat{x})$ in problem (6) are concave, piecewise quadratic, and differentiable. The ordinary Hessians do not exist for this function because the function gradient

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

is not differentiable. However, one can define its generalized Hessian which is the $m \times m$ symmetric negative semidefinite matrix

$$\partial_p^2 S(p, \beta, x_k) = -AD^{\#}(z)A^\top,$$

where $D^{\#}(z)$ denotes an $n \times n$ diagonal matrix where the ith-diagonal element z^i equals to 1, if $(x_k + A^\top p - \beta c)^i > 0$ and equals to 0, if $(x_k + A^\top p - \beta c)^i \leq 0$, $i = 1, \ldots, n$. Since the generalized Hessian may be singular, we used a modified Newton direction

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

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

In this case, the modified Newton method is

$$p_{s+1} = p_s - \lambda_s[\partial_p^2 S(p_s, \beta, x_k) - \delta I_m]^{-1} S_p(p_s, \beta, x_k), \tag{24}$$

where stepsize λ_s is chosen by Armijo rule. The stopping rule is

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

Mangasarian investigated the convergence of the generalized Newton method for unconstrained minimization of similar convex piecewise quadratic function with Armijo stepsize regulation. The proof of the finite global convergence is given in refs. [1, 2].

5. Numerical results

Following the approach proposed by Mangasarian in ref. [1], we introduced the syntetically generated linear test programme. We considered LP problem (P) with a very large number of variables and a moderate number of constraints such that $n \gg m$. Typically $n \approx 10^6$, $m \approx 10^3$.

The test programme generator produced a random matrix A for a given m, n, and density ρ. In particular, $\rho = 1$ means that all the entries in A were generated as random numbers, whereas $\rho = 0.01$ indicated that only 1% of the entries in A was generated randomly and others were set equal to zero. The elements of A were uniformly distributed within the interval $[-50, 50]$. The primal random solution x_* contained the components in the interval $[0, 10]$ and the dual random solution u_* contained the components in the interval $[-10, 10]$. About one-half components of the dual solution were set to zero and nearly $3m$ components of the primal solution were positive. The vectors x_* and u_* satisfied the complementarity condition. The solutions x_* and u_* were used to generate an objective function vector c and a right-hand side vector b for the linear programming (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 $1 \leq \xi^i \leq 10$ for all i. The LP test generator is given below as Code 1.

Code 1. LP MATLAB Test Generator

```
%Code 1: Generator random solvable LP: min c'x s.t.
%Ax=b; x>=0; A:m-by-n
% Input: m,n,d(ensity); Output: A,b,c; (x,u): primal-dual
%solution
pl=inline('(abs(x)+x)/2');  %pl(us) function
A=sprand (m,n,d); A=100*(A-0.5*spones (A));
u=10*spdiags((sign (pl(rand(m,1)-rand(m,1)))),0,m,m)*
(rand(m,1)-rand(m,1));
x=sparse(10*pl(rand(n,1)-(n-3*m)/n)); b=A*x;
xi=spdiags((ones(n,1)-sign(pl(x))),0,n,n)*
(ones(n,1)+9*rand(n,1));
c=A'*u+xi;
```

The algorithm presented was used for solving both primal and dual LP problems. It combined the iterative processes (20), (21), and the generalized Newton method (24) applied to the solution of the maximization problem (20). The proposed method (20), (21), and the generalized Newton method were implemented in MATLAB 6.5 as the EGM algorithm. We used a 2.26 GHz Pentium 4 with 1 Gb RAM. The numerical results obtained by the EGM algorithm are presented in table 1.

The starting point used in all the given examples was $x_0 = 0_n$. We always set $\beta = 0.1$, tol $= 10^{-12}$. In all occasions, it turned out that $\beta \geq \beta_*$. Thus, the normal solution $\hat{x}_*$ was obtained by a single iteration of processes (20), (21), i.e. $\omega = 1$. The number of iterations made by the generalized Newton method used for solving problem $\max\limits_{p \in \mathbb{R}^m} S(p, \beta, x_0)$ is shown in the third column of table 1. In all examples, after the first iteration, we got from formula (21) $x_1 = \hat{x}_*$, which was a normal solution of the primal problem (P).

According to Theorem 2.2, the maximization of the function $S(p, \beta, x_1)$ (where $x_1 = \hat{x}_*$) with respect to p yielded the maximal value $p(\beta)$, which gave the dual solution $u_* = p(\beta)/\beta$. In all examples, only two iterations of the generalized Newton method were required for this maximization. The total time of computation is given in the second column. The fourth and fifth columns give us the Euclidian norms of the residual vectors. The last column contains the difference between the optimal values of the goal functions for problems (P) and (D).

Table 1. Performance of the EGM algorithm.

$m \times n \times \rho$	$T(s)$	Newton iterations	$\|Ax - b\|$	$\|(A^\top u - c)_+\|$	$\lvert c^\top x - b^\top u \rvert$
$100 \times 10^6 \times 0.01$	43.7	17	1.7×10^{-11}	2.0×10^{-13}	2.8×10^{-11}
$300 \times 10^6 \times 0.01$	61.6	13	8.8×10^{-11}	5.4×10^{-13}	2.7×10^{-10}
$600 \times 10^6 \times 0.01$	98.7	12	2.8×10^{-10}	1.5×10^{-12}	1.2×10^{-9}
$1000 \times 10^6 \times 0.01$	136.0	10	1.2×10^{-9}	4.2×10^{-12}	1.7×10^{-9}
$500 \times 10^4 \times 1$	38.1	8	3.2×10^{-8}	3.5×10^{-11}	8.3×10^{-8}
$1000 \times 10^4 \times 1$	147.2	7	1.2×10^{-7}	1.1×10^{-10}	1.8×10^{-7}
$3000 \times 10^4 \times 0.01$	104.9	7	2.1×10^{-9}	9.6×10^{-12}	5.8×10^{-10}
$4000 \times 10^4 \times 0.01$	308.6	7	3.1×10^{-9}	1.3×10^{-11}	8.8×10^{-9}
$500 \times (3 \times 10^6) \times 0.01$	257.0	12	3.3×10^{-10}	1.5×10^{-12}	2.4×10^{-10}
$1000 \times (3 \times 10^6) \times 0.01$	552.6	15	1.1×10^{-9}	3.5×10^{-12}	1.2×10^{-9}
$1000 \times (5 \times 10^6) \times 0.01$	1167.8	13	5.9×10^{-9}	2.2×10^{-12}	6.6×10^{-8}
$500 \times 10^7 \times 0.01$	1443.7	12	7.4×10^{-9}	2.3×10^{-12}	1.3×10^{-8}

To compare the proposed EGM algorithm with the available solvers, we present the comparative tests. To simplify our task, we used the integer test generator which provided integer entries in A, b, c, x, u.

Code 2. Integer LP MATLAB Test Generator

```
%Code 2: Integer generate random solvable LP: min c'x s.t.
%Ax = b, x>=0; A:m-by-n
%Input: m,n,d(ensity); Output: A,b,c; (x,u): primal-dual
%solution
pl=inline('(abs(x)+x)/2');%pl(us) function
tic;A=sprand(m,n,d);A=fix(100*(A-0.5*spones (A)));
uu=10*spdiags((sign(pl(rand(m,1)-rand(m,1)))),0,m,m)*
(rand(m,1)-rand(m,1));
u=fix(uu); x=fix(sparse(10*pl(rand(n,1)-(n-3*m)/n))); b=A*x;
```

```
xi=fix(spdiags((ones(n,1)-sign(pl(x))),0,n,n)*
(ones(n,1)+9*rand(n,1)));
c=A'*u+xi;
```

The termination criterion of EGM algorithm was

$$\Delta_1 = \|Ax - b\|_\infty \le \mathrm{tol}, \qquad \Delta_2 = \|(A^\top u - c)_+\|_\infty \le \mathrm{tol}, \qquad \Delta_3 = |c^\top x - b^\top u| \le \mathrm{tol},$$

where $\|a\|_\infty$ is the Chebyshev norm of a vector a, tol is tolerance, usually tol $= 10^{-6}$.

To compare the effectiveness of different algorithms, we solved four linear programmes generated by Code 2 on a Celeron 2.02 GHz computer with 1.0 Gb of memory running Windows XP. We compared EGM algorithm with BPMPD v.2.3 (interior point method) [9], MOSEK v.2.0 (interior point method) [10], CPLEX (interior point method), and CPLEX (simplex method) v.6.0.1. The results are presented in table 2.

It should be stressed that neither the linear programme with 5 million variables, $\rho = 0.01$, $m = 1000$, nor those with 10^5 variables, $\rho = 1$, $m = 1000$, could be successfully solved by any of the codes listed previously, excepting the proposed EGM algorithm which solved these problems in 1021 and 2663 s, respectively, and gave very high accuracy (less than 10^{-7}). Nevertheless, our algorithm did not give the best results for small and moderate size problems. As is evident from table 2, the BPMPD solver was superior to the others as applied for the first test problem; the MOSEK solver showed the best results in the second test problem.

Table 2. Comparative results.

$m \times n \times \rho$	Solver	$T(s)$	Iterations	Δ_1	Δ_2	Δ_3
$500 \times 10^4 \times 1$	EGM (NATLAB)	55.0	12	1.5×10^{-8}	1.8×10^{-12}	1.2×10^{-7}
	BPMPD (interior point)	37.4	23	2.3×10^{-10}	1.8×10^{-11}	1.1×10^{-10}
	MOSEK (interior point)	87.2	6	9.7×10^{-8}	3.8×10^{-9}	1.6×10^{-6}
	CPLEX (interior point)	80.3	11	1.8×10^{-8}	1.1×10^{-7}	0.0
	CPLEX (simplex)	61.8	8308	8.6×10^{-4}	1.9×10^{-10}	7.2×10^{-3}
$3000 \times 10^4 \times 0.01$	EGM (NATLAB)	155.4	11	6.1×10^{-10}	3.4×10^{-13}	3.6×10^{-8}
	BPMPD (interior point)	223.5	14	4.6×10^{-9}	2.9×10^{-10}	3.9×10^{-9}
	MOSEK (interior point)	42.6	4	3.1×10^{-8}	1.2×10^{-8}	3.7×10^{-8}
	CPLEX (interior point)	69.9	5	1.1×10^{-6}	1.3×10^{-7}	0.0
	CPLEX (simplex)	1764.9	6904	3.0×10^{-3}	8.1×10^{-9}	9.3×10^{-2}
$1000 \times (5 \times 10^6) \times 0.01$	EGM (NATLAB)	1007.5	10	3.9×10^{-8}	1.4×10^{-13}	6.1×10^{-7}
$1000 \times 10^5 \times 1$	EGM (NATLAB)	2660.8	8	2.1×10^{-7}	1.4×10^{-12}	7.1×10^{-7}

Thus, the examples presented previously have clearly demonstrated the principal effectiveness of the method proposed in solving the problems with a very large number ($\approx 10^6$) of nonnegative variables and a moderate ($\approx 10^3$) number of equality type constraints.

It is also evident that the results of the EGM algorithm are highly comparative with those of simplex and interior point methods, at the same time outperforming them in solving large LP problems.

Acknowledgements

This research was supported by the programme 'Leading Scientific Schools' (project no. NSh-1737.2003.1) and the Russian Foundation for Basic Research (project no. 03-01-00465). We are grateful to Ustav Malkov who enabled us to use LP solvers for comparison purposes. We would like to thank Vlasta Malkova for the assistance in conducting the numerical calculations results presented in table 2. We are also indebted to our anonymous referees for their valuable comments and especially for the request to present the results of our computations, which helped to justify the superiority of our algorithm.

REFERENCES

[1] Mangasarian, O.L., 2004, A Newton method for linear programming. *Journal of Optimization Theory and Applications*, **121**, 1–18.

[2] Mangasarian, O.L., 2002, A finite Newton method for classification. *Optimization Methods and Software*, **17**, 913–930.

[3] Kanzow, C., Qi, H. and Qi, L., 2003, On the minimum norm solution of linear programs. *Journal of Optimization Theory and Applications*, **116**, 333–345.

[4] Bertsekas, D., 1982, *Constrained Optimization and Lagrange Multiplier Methods*, Academic Press.

[5] Antipin, A.S., 1979, *Nonlinear Programming Methods Based on Primal and Dual Augmented Lagrangian* (Moscow: Institute for Systems Studies) (in Russian).

[6] Di Pillo, G. and Grippo, L., 1989, Exact penalty functions in constrained optimization. *SIAM Journal of Control and Optimization*, **27**, 1333–1360.

[7] Evtushenko, Yu.G. and Golikov, A.I., 1999, The augmented Lagrangian function for the linear programming problems. In: Yu.S. Popkov (Ed.) *Dynamics of Non-Homogeneous Systems* (Moscow: Institute for System Analysis), vol. 2, pp. 63–67.

[8] Golikov, A.I. and Evtushenko, Yu.G., 2000, Search for normal solutions in linear programming problems. *Computational Mathematics and Mathematical Physics*, **40**, 1694–1714.

[9] Meszaros, Cs., 1999, The BPMPD interior point solver for convex quadratic programming problems. *Optimization Methods and Software*, **11&12**, 431–449.

[10] Andersen, E.D. and Andersen, K.D., 2000, The MOSEK interior point optimizer for linear programming: an implementation of homogeneous algorithm, In: H. Frenk, K. Roos, T. Terlaky and S. Zhang (Eds) *High Performance Optimization* (New York: Kluwer Academic Publishers), pp. 197–232.