Efficient Algorithm for Large Sparse Equality Constrained Nonlinear Programming Problems

Lukšan, Ladislav
1996

INSTITUTE OF COMPUTER SCIENCE

ACADEMY OF SCIENCES OF THE CZECH REPUBLIC

Efficient Algorithm for Large Sparse Equality Constrained Nonlinear Programming Problems

Ladislav Lukšan, Jan Vlček

Technical report No. 652

February 1996

Institute of Computer Science, Academy of Sciences of the Czech Republic
Pod vodárenskou věží 2, 182 07 Prague 8, Czech Republic
phone: (+4202) 6884244 fax: (+4202) 8585789
e-mail: luksan@uivt.cas.cz

INSTITUTE OF COMPUTER SCIENCE

ACADEMY OF SCIENCES OF THE CZECH REPUBLIC

Efficient Algorithm for Large Sparse Equality Constrained Nonlinear Programming Problems[1]

Ladislav Lukšan, Jan Vlček

Technical report No. 652
February 1996

Abstract

An efficient method for large sparse equality constrained nonlinear programming problems is proposed. This method is based on partial elimination of variables in indefinite KKT system. The direction vector is determined directly, using sparse Gill-Murray decomposition. The Lagrange multiplier vector is obtained iteratively, using smoothed conjugate gradient method. The KKT system is preliminarily transformed, which leads to a special merit function. The efficiency of our algorithm is demonstrated by extensive numerical experiments.

Keywords
Nonlinear programming, sparse problem, equality constraints, truncated Newton method, range space method, special merit function, indefinite system, conjugate gradient method, residual smoothing.

[1]This work was supported under grant No. 201/96/0918 given by the Czech Republic Grant Agency

1 Introduction

Consider the problem of finding a point $x^* \in R^n$, such that

$$x^* = \arg \min_{x \in \mathcal{F}} F(x), \tag{1.1}$$

where $\mathcal{F} \subset R^n$ is a feasible set defined by the system of equations

$$\mathcal{F} = \{x \in R^n : c_k(x) = 0, 1 \le k \le m\}. \tag{1.2}$$

where $m \le n$ (in fact we consider only local minimum). Here $F : R^n \to R$ and $c_k : R^n \to R$, $1 \le k \le m$, are twice continuously differentiable functions, whose gradients and Hessian matrices will be denoted by $\nabla F(x)$, $\nabla c_k(x)$, $1 \le k \le m$, and $\nabla^2 F(x)$, $\nabla^2 c_k(x)$, $1 \le k \le m$, respectively. Furthermore, we use the notation $c(x) = [c_1(x), \dots, c_m(x)]^T$ and $A(x) = [a_1(x), \dots, a_m(x)] = [\nabla c_1(x), \dots, \nabla c_m(x)]$ and we suppose that the matrix $A(x)$ has a full column rank. Then the solution $x^* \in R^n$ of the problem (1.1)-(1.2) satisfies the Karush-Kuhn-Tucker (KKT) conditions, i.e. there exists a vector $u^* \in R^m$, such that

$$\begin{aligned} \nabla_x L(x^*, u^*) &= \nabla F(x^*) + A(x^*)u^* = 0, && (1.3)\\ \nabla_u L(x^*, u^*) &= c(x^*) = 0, && (1.4) \end{aligned}$$

where

$$L(x, u) = F(x) + u^T c(x)$$

is the Lagrangian function, whose gradient and Hessian matrix will be denoted by

$$\begin{aligned} g(x,u) &= \nabla_x L(x,u) = \nabla F(x) + \sum_{k=1}^{m} u_k \nabla c_k(x),\\ G(x,u) &= \nabla_x^2 L(x,u) = \nabla^2 F(x) + \sum_{k=1}^{m} u_k \nabla^2 c_k(x), \end{aligned}$$

and $(x^*, u^*) \in R^{n+m}$ is the KKT pair (first order necessary conditions). Let $Z(x)$ be the matrix whose columns form an orthonormal basis in the null space of $A^T(x)$ so that $A^T(x)Z(x) = 0$ and $Z^T(x)Z(x) = I$. If, in addition to (1.3)-(1.4), the matrix $Z^T(x^*)G(x^*, u^*)Z(x^*)$ is positive definite, then the point $x^* \in R^n$ is a solution of the problem (1.1)-(1.2) (second order sufficient conditions).

Basic methods for a solution of the problem (1.1)-(1.2) are iterative and their iteration step has the form

$$\begin{aligned} x^+ &= x + \alpha d, && (1.5)\\ u^+ &= u + \alpha v, && (1.6) \end{aligned}$$

where $(d, v) \in R^{n+m}$ is a direction pair ($d \in R^n$ is a direction vector) and $\alpha > 0$ is a stepsize. In this contribution, we confine our attention to methods derived from the

Newton method used for a solution of the KKT system (1.3)-(1.4). The iteration step of the Newton method has the form (1.5)-(1.6), where $\alpha = 1$ and

$$\begin{bmatrix} G(x,u) & A(x) \\ A^T(x) & 0 \end{bmatrix} \begin{bmatrix} d \\ v \end{bmatrix} = - \begin{bmatrix} g(x,u) \\ c(x) \end{bmatrix}. \tag{1.7}$$

This is a system of $n+m$ linear equations with $n+m$ unknowns $(d,v) \in R^{n+m}$. The matrix

$$K = \begin{bmatrix} G & A \\ A^T & 0 \end{bmatrix} \tag{1.8}$$

is always indefinite (cf. Theorem 1). Moreover, the matrix G is not positive definite in general even if the matrix $Z^T G Z$ is. This fact can lead to some difficulties. Therefore, it is advantageous to transform the system (1.7) in such a way as to contain, if possible, a positive definite matrix in the left-upper corner. This can often be done by addition of the second equation, multiplied by ρA, to the first equation (cf. Theorem 2), which yields

$$\begin{bmatrix} B & A \\ A^T & 0 \end{bmatrix} \begin{bmatrix} d \\ v \end{bmatrix} = - \begin{bmatrix} b \\ c \end{bmatrix}. \tag{1.9}$$

where

$$B = G + \rho A A^T,$$

$$b = g + \rho A c = \nabla F + Au + \rho A c.$$

Using a partial elimination of variables, (1.9) can be transformed to the form

$$\begin{aligned} Bd &= -(b + Av), && (1.10) \\ A^T B^{-1} A v &= c - A^T B^{-1} b. && (1.11) \end{aligned}$$

If the matrices A and B are dense, then we can construct matrix $A^T B^{-1} A$, determine vector v using (1.11) and compute vector d by substituting v into (1.10). If the matrices A and B are large and sparse, then matrices B^{-1} and, especially, $A^T B^{-1} A$ are usually dense, and we cannot use this way. In this case, we can solve the system (1.9) either iteratively using the smoothed conjugate gradient method [10] or directly using the sparse Bunch-Parlett [1] decomposition. However, the matrix K has relatively large dimension $n+m$ and its nonzero elements, derived from the matrix A, are usually far from the main diagonal, which can lead to considerable fill-in. Therefore, it is useful to find another possibility which removes these insufficiences.

In this contribution, we will concentrate on a combined direct and iterative method, which is based on the equations (1.10)-(1.11). Equation (1.10) will be solved directly using the sparse Gill-Murray [5] decomposition

$$LDL^T = B + E, \tag{1.12}$$

where L is a nonsingular lower triangular matrix, D is a positive definite diagonal matrix and E is a positive semidefinite diagonal matrix. The matrix E is determined in such a way as to guarantee positive definiteness of the matrix $B+E$ (if B is sufficiently positive definite, then $E = 0$). Equation (1.11) will be solved iteratively using the smoothed conjugate gradient method. An advantage of this approach consists in the fact that the matrix B has a lower dimension n and its elements are not usually too far from the main diagonal, which leads to a lower fill-in. Moreover, equation (1.11) can be solved approximately, like the truncated Newton method for unconstrained optimization [3]. However, this procedure lays a higher emphasis on the determination of the parameter α in (1.5)-(1.6), i.e. on the choice of a merit function for the stepsize selection.

The contribution is organized as follows. In Section 2, we propose some results concerning system (1.9), define the special merit function (2.5) suitable for inexact solution to the system (1.9) and show a correctness of the Armijo type line search procedure. Section 3 contains a detailed description of our algorithm for large sparse equality constrained nonlinear programming problems together with results obtained by extensive numerical experiments.

In this contribution, we denote by $\|.\|$ the Euclidean (or spectral) norm and by $\|.\|_1$ the l_1 norm (sum of absolute values).

2 Direction determination and stepsize selection

For the solution of equations (1.10)-(1.11), we can use the Gill-Murray decomposition (1.12). If the matrix B is indefinite, which is a frequent situation when $\rho = 0$, then the matrix $B + E$ can be different enough from the matrix B, and good convergence properties of the Newton method can be lost. Therefore, it can be advantageous to use the value $\rho > 0$. The following theorems hold for the matrix K defined by (1.8) and the matrix $B = G + \rho AA^T$:

Theorem 1. *Let k_+, k_- and k_0 be the number of positive, negative and zero eigenvalues of the matrix K and let l_+, l_- and l_0 be the number of positive, negative and zero eigenvalues of the matrix Z^TGZ. Then $k_- = l_- + m$, $k_+ = l_+ + m$ and $k_0 = l_0$.*

Proof. See [6].

Theorem 2. *Let the matrix Z^TGZ be positive definite. Then there exists a number $\overline{\rho} > 0$, such that the matrix B is positive definite whenever $\rho \geq \overline{\rho}$.*

Proof. See [4].

Theorem 3. *Let the matrix K be nonsingular. Then there exists a number $\overline{\rho} > 0$, such that the matrix $A^TB^{-1}A$ is positive definite whenever $\rho \geq \overline{\rho}$.*

Proof. (a) First we prove that there exists a number $\rho_0 > 0$ such that the matrix $G + \rho AA^T$ is nonsingular whenever $\rho \geq \rho_0$. From Theorem 1, we can deduce that nonsingularity of the matrix K implies nonsingularity of the matrix Z^TGZ. Therefore, there exists a number $\underline{G} > 0$ such that $\|Z^TGZz\| \geq \underline{G}\|z\|$ $\forall z \in R^{n-m}$. Denote $Y = A(A^TA)^{-1}$ so that $A^TY = I$, $Z^TY = 0$ and $\|Y\| \leq \overline{A}/\underline{A}^2$, where $\overline{A} = \|A\|$ and $\underline{A}$ is the lowest singular value of the matrix A. Then every vector $x \in R^n$ can be uniquely expressed in the form $x = Yy + Zz$, where $y \in R^m$ and $z \in R^{n-m}$. Suppose that

$$(G + \rho AA^T)x = GYy + GZz + \rho Ay = 0$$

for some nonzero vector $x \in R^n$. Then necessarily

$$Z^T G Y y + Z^T G Z z = 0 \tag{2.1}$$

and

$$y^T Y^T G Y y + y^T Y^T G Z z + \rho y^T y = 0. \tag{2.2}$$

From (2.1) we obtain

$$\frac{\overline{GA}}{\underline{A}^2}\|y\| \geq \|Z^T G Y y\| = \|Z^T G Z z\| \geq \underline{G}\|z\|,$$

where $\overline{G} = \|G\|$, so that $\|z\| \leq (\overline{GA})/(\underline{GA}^2)\|y\|$. On the other hand, we can write

$$\begin{aligned} y^T Y^T G Y y + y^T Y^T G Z z + \rho y^T y &\geq \rho\|y\|^2 - \frac{\overline{GA}^2}{\underline{A}^4}\|y\|^2 - \frac{\overline{GA}}{\underline{A}^2}\|y\|\|z\| \\ &\geq \left[\rho - \frac{\overline{GA}^2}{\underline{A}^4}\left(1 + \frac{\overline{G}}{\underline{G}}\right)\right]\|y\|^2, \end{aligned}$$

so that (2.2) cannot be satisfied if $\|x\| > 0$ and $\rho \geq \rho_0$, where $\rho_0 > (\overline{GA}^2/\underline{A}^4)(1 + (\overline{G}/\underline{G}))$, which is a contradiction.
(b) Denote $B_0 = G + \rho_0 A A^T$. Since the matrix B_0 is nonsingular by (a), its Schur complement $A^T B_0^{-1} A$ in the matrix K is also nonsingular. Let μ be an eigenvalue of the matrix $A^T B_0^{-1} A$ and w be a corresponding eigenvector. Then we obtain successively

$$\begin{aligned} A^T B_0^{-1} A w &= \mu w \\ B_0^{-1} A A^T B_0^{-1} A w &= \mu B_0^{-1} A w \\ (I + (\rho - \rho_0) B_0^{-1} A A^T) B_0^{-1} A w &= (1 + (\rho - \rho_0)\mu) B_0^{-1} A w \\ (1 + (\rho - \rho_0)\mu)^{-1} B_0^{-1} A w &= (I + (\rho - \rho_0) B_0^{-1} A A^T)^{-1} B_0^{-1} A w \\ (1 + (\rho - \rho_0)\mu)^{-1} A^T B_0^{-1} A w &= A^T (B_0 + (\rho - \rho_0) A A^T)^{-1} A w \\ \mu (1 + (\rho - \rho_0)\mu)^{-1} w &= A^T B^{-1} A w \end{aligned}$$

provided $\rho - \rho_0 \neq -1/\mu$. Consider the function $\lambda(\mu) = \mu/(1 + (\rho - \rho_0)\mu)$ for a given $\rho \geq \rho_0$. If $\mu > 0$, then $\lambda(\mu) > 0$ for an arbitrary $\rho \geq \rho_0$. If $\mu < -1/(\rho - \rho_0) < 0$ then again $\lambda(\mu) > 0$. Therefore, if either $\mu > 0$ or $\mu < -1/(\rho - \rho_0) < 0$ for all eigenvalues of the matrix $A^T B_0^{-1} A$, then all eigenvalues $\lambda(\mu)$ of the matrix $A^T B^{-1} A$ are positive. This situation appears if $\rho \geq \overline{\rho} > \rho_0 - 1/\mu_0$, where $\mu_0 < 0$ is the greatest negative eigenvalue of the matrix $A^T B_0^{-1} A$. □

Theorem 3 has a practical corolary. It shows that there exists a transformation of system (1.7), such that the system (1.11) has positive definite matrix. This fact is very advantageous for application of the conjugate gradient method to (1.11).

In the subsequent considerations, we will suppose that $LDL^T = B + E$ is the Gill-Murray decomposition such that $\underline{B}\|d\|^2 \leq d^T LDL^T d \leq \overline{B}\|d\|^2 \; \forall d \in R^n$, where $\underline{B}$ and $\overline{B}$ are some constants independent on the current iteration. The left inequality is a consequence of the Gill-Murray decomposition. If the right inequality is not satisfied, then the matrix B has to be modified before decomposition.

Using partial elimination of variables, we can transform (1.9) (with LDL^T instead of B) to the form

$$\begin{aligned} LDL^T d &= -(b + Av), & (2.3) \\ A^T(L^{-1})^T D^{-1} L^{-1} A v &= c - A^T (L^{-1})^T D^{-1} L^{-1} b. & (2.4) \end{aligned}$$

We will use the following merit function

$$P(\alpha) = F(x + \alpha d) + (u + v)^T c(x + \alpha d) + \frac{\rho}{2} \|c(x + \alpha d)\|^2 + \sigma \|c(x + \alpha d)\|_1 \qquad (2.5)$$

for the stepsize selection ($\sigma > 0$ is an additional penalty parameter). Together with this merit function we also use its piecewise linear approximation

$$\overline{P}(\alpha) = P(0) + \alpha d^T (b + Av) + \sigma(\|c + \alpha A^T d)\|_1 - \|c\|_1) \qquad (2.6)$$

and we denote by $dP_+(0)/d\alpha = \lim_{\alpha \downarrow 0} (P(\alpha) - P(0))/\alpha$ the corresponding directional derivative. The main advantage of the merit function (2.5) is the fact that it implies a good descent property of an inexact solution to the system (1.9). The following theorem holds:

Theorem 4. *Suppose that* $\underline{B}\|d\|^2 \le d^T LDL^T d \le \overline{B}\|d\|^2 \ \forall d \in R^n$. *Let* $v \in R^m$ *be an inexact solution of the equation (2.4) such that* $\|r\|_1 \le \|c\|_1$, *where* $r \in R^m$ *is the residual vector determined by the formula*

$$r = c - A^T (L^{-1})^T D^{-1} L^{-1} b - A^T (L^{-1})^T D^{-1} L^{-1} A v = c + A^T d \qquad (2.7)$$

and $d \in R^n$ *is a solution of the equation (2.3). Then* $dP_+(0)/d\alpha \le \overline{P}(1) - \overline{P}(0) \le -\underline{B}\|d\|^2$.

Proof. Differentiating (2.5) or (2.6) we get

$$\begin{aligned} dP_+(0)/d\alpha &= d^T(b + Av) + \sigma\Big(\sum_{c_k = 0} |a_k^T d| + \sum_{c_k > 0} a_k^T d - \sum_{c_k < 0} a_k^T d\Big) \\ &= d^T(b + Av) + \sigma\Big(\sum_{c_k = 0} (|c_k + a_k^T d| - |c_k|) + \sum_{c_k > 0} (c_k + a_k^T d - |c_k|) \\ &\quad - \sum_{c_k < 0} (c_k + a_k^T d + |c_k|)\Big) \\ &\le d^T(b + Av) + \sigma(\|c + A^T d\|_1 - \|c\|_1) = \overline{P}(1) - \overline{P}(0). \end{aligned}$$

On the other hand (2.3) and (2.7) imply that

$$\overline{P}(1) - \overline{P}(0) = d^T(b + Av) + \sigma(\|c + A^T d\|_1 - \|c\|_1) = -d^T LDL^T d + \sigma(\|r\|_1 - \|c\|_1)$$

which together with the assumptions $d^T LDL^T d \ge \underline{B}\|d\|^2$ and $\|r\|_1 \le \|c\|_1$ gives assertion of the theorem. $\square$

Note that the main reason for use of the Gill-Murray decomposition (1.12) is a required positive definiteness of the matrix LDL^T which is essential for proof of Theorem 4.

Let $v \in R^m$ be an inexact solution of the equation (2.4) satisfying assumptions of Theorem 4 and $d \in R^n$ be the corresponding solution of the equation (2.3). Then we can use the standard Armijo rule for steplength determination i.e. $\alpha > 0$ in (1.5)-(1.6) is chosen so that it is the first member of the sequence $\underline{\beta}^j$, $j = 0, 1, 2, \ldots$, $0 < \underline{\beta} < 1$, such that

$$P(\alpha) - P(0) \leq \underline{\varepsilon}\alpha(\overline{P}(1) - \overline{P}(0)), \tag{2.8}$$

where $0 < \underline{\varepsilon} < 1$. In the subsequent considerations, we will assume that there exist constants $\overline{g}$, $\overline{G}$, $\overline{c}$, $\overline{A}$, $\underline{A}$, independent of the current iteration, such that $\|\nabla F(x+\alpha d)\| \leq \overline{g}$, $\|\nabla^2 F(x+\alpha d)\| \leq \overline{G}$, $\|c(x+\alpha d)\| \leq \overline{c}$, $\|A(x+\alpha d)\| \leq \overline{A}$, $\|\nabla^2 c_k(x+\alpha d)\| \leq \overline{G}$, $1 \leq k \leq m$, $\|A(x+\alpha d)w\| \geq \underline{A}\|w\|$ $\forall w \in R^n$ hold, respectively, for all $0 \leq \alpha \leq 1$.

Lemma 1. *Let assumptions of Theorem 4 be satisfied (together with assumptions of boundedness given above). Then there exists a constant $\overline{K}$, independent of the current iteration, such that*

$$P(\alpha) \leq \overline{P}(\alpha) + \alpha^2 \overline{K}\|d\|^2 \tag{2.9}$$

$\forall 0 \leq \alpha \leq 1$.

Proof. Since (2.7) implies

$$r = c - A^T(L^{-1})^T D^{-1} L^{-1}(\nabla F + \rho A c) - A^T(L^{-1})^T D^{-1} L^{-1} A(u+v)$$

and since $\|r\| \leq \|r\|_1 \leq \|c\|_1 \leq \sqrt{m}\|c\|$ and

$$w^T A^T (L^{-1})^T D^{-1} L^{-1} A w \geq \frac{1}{\overline{B}}\|Aw\|^2 \geq \frac{\underline{A}^2}{\overline{B}}\|w\|^2$$

$\forall w \in R^m$ hold by assumptions, we can write

$$\frac{\underline{A}^2}{\overline{B}}\|u+v\| \leq \|A^T(L^{-1})^T D^{-1} L^{-1} A(u+v)\| \leq \overline{c}(1+\sqrt{m}) + \frac{\overline{A}}{\underline{B}}(\overline{g} + \rho\overline{A}\overline{c}),$$

so that

$$\|u+v\| \leq \frac{\overline{B}}{\underline{A}^2}\left[\overline{c}(1+\sqrt{m}) + \frac{\overline{A}}{\underline{B}}(\overline{g} + \rho\overline{A}\overline{c})\right] \triangleq \overline{U}.$$

Applying the Taylor expansion to every term of (2.5) and using (2.6), we get

$$\begin{aligned} P(\alpha) &\leq \overline{P}(\alpha) + \frac{1}{2}\alpha^2\overline{G}\|d\|^2 + \frac{1}{2}\alpha^2\sum_{k=1}^m |u_k + v_k|\overline{G}\|d\|^2 \\ &\quad + \frac{1}{2}\rho\alpha^2\overline{A}^2\|d\|^2 + \frac{1}{2}\rho\alpha^2\sum_{k=1}^m |c_k|\overline{G}\|d\|^2 + \frac{1}{2}\sigma\alpha^2\sum_{k=1}^m \overline{G}\|d\|^2 \\ &\leq \overline{P}(\alpha) + \frac{1}{2}\alpha^2\left[(1+\overline{U}\sqrt{m} + \rho\overline{c}\sqrt{m} + \sigma m)\overline{G} + \rho\overline{A}^2\right]\|d\|^2 \triangleq \overline{P}(\alpha) + \alpha^2\overline{K}\|d\|^2 \end{aligned}$$

$\forall 0 \leq \alpha \leq 1$ (ρ and σ are assumed to be constants). $\square$

Theorem 5. *Let the assumptions of Lemma 1 hold and let $d \neq 0$. Then there exist an integer $k \geq 0$ and a number $\underline{\alpha} > 0$, independent of the current iteration, such that the Armijo rule gives the value $\alpha = \underline{\beta}^j$, satisfying (2.8), with $j \leq k$ and $\alpha \geq \underline{\alpha}$. Moreover*

$$P(\alpha) - P(0) \leq -\underline{\alpha \varepsilon B}\|d\|^2. \tag{2.10}$$

Proof. Since

$$\begin{aligned}
\overline{P}(\alpha) - \overline{P}(0) - \alpha(\overline{P}(1) - \overline{P}(0)) &= \sigma(\|c + \alpha A^T d\|_1 - \|c\|_1) - \alpha\sigma(\|c + A^T d\|_1 - \|c\|_1) \\
&\leq \sigma(\alpha\|c + A^T d\|_1 + (1-\alpha)\|c\|_1 - \|c\|_1 \\
&\quad -\alpha\|c + A^T d\|_1 + \alpha\|c\|_1) = 0
\end{aligned}$$

$\forall 0 \leq \alpha \leq 1$, we can write

$$\begin{aligned}
P(\alpha) - P(0) &\leq \overline{P}(\alpha) - \overline{P}(0) + \alpha^2 \overline{K}\|d\|^2 \leq \alpha(\overline{P}(1) - \overline{P}(0) + \alpha\overline{K}\|d\|^2) \\
&\leq \alpha(\overline{P}(1) - \overline{P}(0))(1 - \alpha\frac{\overline{K}}{\underline{B}})
\end{aligned}$$

by Lemma 1 and Theorem 4, so that (2.8) holds whenever $\alpha \leq (\underline{B}/\overline{K})(1 - \underline{\varepsilon})$. Let $k \geq 0$ be chosen so that it is the lowest integer such that $\underline{\beta}^k \leq (\underline{B}/\overline{K})(1 - \underline{\varepsilon})$ and let $\alpha = \underline{\beta}^j$ be given by the Armijo rule to satisfy (2.8). Then

$$\alpha = \underline{\beta}^j \geq \underline{\beta}^k \geq \underline{\beta}\frac{\underline{B}}{\overline{K}}(1 - \underline{\varepsilon}) \triangleq \underline{\alpha}. \tag{2.11}$$

Using (2.11) and Theorem 4, we get

$$P(\alpha) - P(0) \leq \alpha\underline{\varepsilon}(\overline{P}(1) - \overline{P}(0)) \leq -\underline{\alpha\varepsilon B}\|d\|^2.$$

□

Now we focus our attention on the inexact solution of equations (2.3)-(2.4). The matrix $A^T(L^{-1})^T D^{-1} L^{-1} A$ is positive definite (since A has a full column rank and LDL^T is positive definite), so that the equation (2.4) can be solved by the smoothed conjugate gradient method [10]. The iterative process is terminated if a sufficient accuracy, guaranteeing superlinear rate of convergence (see [3]), is reached and, at the same time, the condition $\|r\|_1 \leq \|c\|_1$ is satisfied. These facts imply the following algorithm for the direction determination:

Algorithm 1. Direction determination.

Data: $0 < \omega < 1$.

Step 1: Initiation. Set $\tilde{v}_0 := 0$, $\tilde{r}_0 := c - A^T(L^{-1})^T D^{-1} L^{-1} b$, $v_0 := \tilde{v}_0$, $r_0 := \tilde{r}_0$ $\omega := \min(\omega, \|r_0\|)$, and $j := 0$.

Step 2: CG iteration. If $j \geq n+3$, then go to Step 6, otherwise set $j := j+1$. Compute $\beta_{j-1} := \|\tilde{r}_{j-1}\|^2$. If $j = 1$, then set $p_{j-1} := \tilde{r}_{j-1}$, otherwise set $p_{j-1} := \tilde{r}_{j-1} + (\beta_{j-1}/\beta_{j-2})p_{j-2}$. Compute $q_{j-1} := A^T(L^{-1})^T D^{-1} L^{-1} A p_{j-1}$ and $\gamma_{j-1} := \beta_{j-1}/p_{j-1}^T q_{j-1}$ and set $\tilde{v}_j := \tilde{v}_{j-1} + \gamma_{j-1} p_{j-1}$, $\tilde{r}_j := \tilde{r}_{j-1} - \gamma_{j-1} q_{j-1}$.

Step 3: Residual smoothing. Compute $\lambda_j := -(r_{j-1} - \tilde{r}_j)^T \tilde{r}_j / \|r_{j-1} - \tilde{r}_j\|^2$ and set $v_j := \tilde{v}_j + \lambda_j (v_{j-1} - \tilde{v}_j)$, $r_j := \tilde{r}_j + \lambda_j (r_{j-1} - \tilde{r}_j)$.

Step 4: Test for sufficient precision. If $\|r_j\|_1 > \omega \|r_0\|_1$, then go to Step 2.

Step 5: Test for sufficient descent. If $\|r_j\|_1 > \|c\|_1$, then go to Step 2.

Step 6: Termination. Set $v := v_j$, compute the direction vector $d := -(L^{-1})^T D^{-1} L^{-1}(b + Av)$ and terminate the computation.

Note that the main reasons for residual smoothing in Step 3 are requirements $\|r\|_1 \le \omega \|r_0\|_1$ and $\|r\|_1 \le \|c\|_1$, so that the norm $\|r\|_1$ should always be as small as possible.

3 Numerical experiments

Now we summarize results from the previous section and give a detailed description of our algorithm. This algorithm uses the sparse Gill-Murray decomposition together with smoothed conjugate gradient method for direction determination and the classical Armijo rule for stepsize selection.

Algorithm 2. Equality constrained optimization (GM+CG).

Data: $\rho \ge 0$, $\sigma \ge 0$, $0 < \underline{\beta} < 1$, $0 < \underline{\varepsilon} < 1$, $0 < \overline{\omega} < 1$, $\overline{\delta} > 0$.

Input: Sparsity pattern of the matrices $\nabla^2 F$ and A. Initial choice of the vextor x.

Step 1: Initiation. Determine sparsity pattern of the matrix B and carry out its symbolic Gill-Murray decomposition. Compute the value $F := F(x)$ and the vector $c := c(x)$. Set $u := 0$ and $i := 0$.

Step 2: Termination. Compute the matrix $A := A(x)$ and the vector $g := g(x, u)$. If $\|c\| \le \overline{\delta}$ and $\|g\| \le \overline{\delta}$, then terminate the computation (the solution is found). Otherwise set $i = i + 1$.

Step 3: Approximation of the Hessian matrix. Compute an approximation G of the Hessian matrix $G(x, u)$, using differences of gradient $g(x, u)$ as in [2]. Compute the matrix $B := G + \rho A A^T$ and carry out its numerical Gill-Murray decomposition.

Step 4: Direction determination. Set $\omega = \min(1/i, \overline{\omega})$. Determine the direction pair (d, v) using Algorithm 1. Set $\alpha := 1$ and compute values of the merit function $P(\alpha)$ and its piecewise linear approximation $\overline{P}(\alpha)$.

Step 5: Termination of the stepsize selection. If $P(\alpha) - P(0) \le \underline{\varepsilon} \alpha (\overline{P}(1) - \overline{P}(0))$, then set $x := x + \alpha d$, $u := u + \alpha v$ and go to Step 2.

Step 6: Continuation of the stepsize selection. Set $\alpha := \underline{\beta} \alpha$, compute value of the merit function $P(\alpha)$ and go to Step 5.

Computational efficiency of Algorithm 2 was tested using 18 sparse problems, listed in the Appendix, which had either 50 or 100 variables. We used parameters $\sigma = 0.15$, $\underline{\beta} = 0.5$, $\underline{\varepsilon} = 10^{-4}$, $\overline{\omega} = 0.9$, $\overline{\delta} = 10^{-6}$, in all numerical experiments. Values of the parameter ρ depended on the problem solved as will be shown below. They were selected to give good results.

The summary of results for all 18 problems is given in Table 1. This table contains the total number of iterations NIT, the total number of function evaluations NFV, the total number of gradient evaluations NGR, the total number of conjugate gradient iterations NCG and the total CPU time on Pentium PC (90 MHz) for double precision arithmetic implementation. The rows correspond to the direct method with the Bunch-Parlett (BP) decomposition of the matrix B, our method (GM+CG) realized by Algorithm 2 and the smoothed conjugate gradient method (CG) applied directly to indefinite system (1.9) and preconditioned using the positive definite matrix

$$C = \begin{bmatrix} LDL^T & A \\ A^T & A^T(L^{-1})^T D^{-1} L^{-1} A + I \end{bmatrix}, \tag{3.1}$$

where LDL^T is an incomplete Gill-Murray decomposition of the matrix B (more details about preconditioner (3.1) can be found in [9]). For (BP) method we used the values $\rho = 0.1$, $\rho = 0.001$ for problems 5,9, respectively, and the value $\rho = 0.0$ in the other cases. For (GM+CG) method we used the values $\rho = 5000.0$, $\rho = 0.0$, $\rho = 0.1$, $\rho = 0.1$, $\rho = 1.0$ for problems 8,9,13,15,16, respectively, and the value $\rho = 50.0$ in the other cases. For (CG) method we used the values $\rho = 0.1$, $\rho = 50000.0$, $\rho = 0.005$, $\rho = 0.01$, $\rho = 100.0$, $\rho = 0.1$, $\rho = 100.0$ for problems 5,8,9,10,13,14,15, respectively, and the value $\rho = 10.0$ in the other cases. All methods presented in Table 1 were implemented using the modular interactive system for universal functional optimization UFO [8].

$n = 50$	NIT	NFV	NGR	NCG	CPU
BP	245	304	1732	0	6.86
GM+CG	258	298	1847	1089	5.60
CG	275	407	2027	3835	13.30
$n = 100$	NIT	NFV	NGR	NCG	CPU
BP	274	326	1909	0	15.60
GM+CG	265	293	1889	1370	12.52
CG	331	622	2376	4720	34.44

Table 1

From the results presented in Table 1, we can draw several conclusions. First, our algorithm (GM+CG) is faster and has much lower storage requirements than the direct method (BP). It is also much faster than the pure iterative method (CG) with the preconditioner (3.1). Second, efficiency of our algorithm depends on the parameter ρ which has sometimes to be adjusted according to the problem to be solved. The pure iterative method (CG) with the preconditioner (3.1) also has this property. Third, we have also tested two preconditioners $C = A^T \tilde{D} A$, with $\tilde{D}$ a positive definite diagonal approximation to the matrix LDL^T, and $C^{-1} = (A^T A)^{-1} A^T LDL^T A (A^T A)^{-1}$, applied to the system (2.4). Efficiency obtained in both these cases was worse than that without preconditioning.

Appendix

This Appendix contains 18 original sparse problems for equality constrained optimization. We use, for prime k and l, the notation $\mathrm{div}(k,l)$ for integer division, i.e., maximum integer not greater than k/l, and $\mathrm{mod}(k,l)$ for remainder after integer division, i.e., $\mathrm{mod}(k,l) = l(k/l - \mathrm{div}(k,l))$. The starting point is $\bar{x}$. Dense problems HS46 - HS53 can be found in [7].

Problem 1. Chained Rosenbrock function with trigonometric-exponential constrains.

$$
\begin{aligned}
F(x) &= \sum_{i=1}^{n-1} [100(x_i^2 - x_{i+1})^2 + (x_i - 1)^2] \\
c_k(x) &= 3x_{k+1}^3 + 2x_{k+2} - 5 + \sin(x_{k+1} - x_{k+2})\sin(x_{k+1} + x_{k+2}) + 4x_{k+1} \\
&- x_k \exp(x_k - x_{k+1}) - 3
\end{aligned}
$$

$1 \le k \le m = n - 2$

$\overline{x}_i = -1.2, \quad \mathrm{mod}(i,2) = 1$

$\overline{x}_i = 1.0, \quad \mathrm{mod}(i,2) = 0$

Problem 2. Chained Wood function with Broyden banded constraints.

$$
\begin{aligned}
F(x) &= \sum_{i=1}^{n/2} [100(x_{2i-1}^2 - x_{2i})^2 + (x_{2i-1} - 1)^2 + 90(x_{2i+1}^2 - x_{2i+2})^2 + (x_{2i+1} - 1)^2 \\
&+ 10(x_{2i} + x_{2i+2} - 2)^2 + (x_{2i} - x_{2i+2})^2/10] \\
c_k(x) &= (2 + 5x_{k+5}^2)x_{k+5} + 1 + \sum_{i=k-5}^{k+1} x_i(1 + x_i)
\end{aligned}
$$

$1 \le k \le m = n - 7$

$\overline{x}_i = -2, \quad \mathrm{mod}(i,2) = 1$

$\overline{x}_i = 1, \quad \mathrm{mod}(i,2) = 0$

Problem 3. Chained Powell singular function with simplified trigonometric-exponential constraints.

$$
\begin{aligned}
F(x) &= \sum_{i=1}^{n/2} [(x_{2i-1} + 10x_{2i})^2 + 5(x_{2i+1} - x_{2i+2})^2 + (x_{2i} - 2x_{2i+1})^4 + 10(x_{2i-1} - x_{2i+2})^4] \\
c_1(x) &= 3x_1^3 + 2x_2 - 5 + \sin(x_1 - x_2)\sin(x_1 + x_2) \\
c_2(x) &= 4x_n - x_{n-1}\exp(x_{n-1} - x_n) - 3
\end{aligned}
$$

$\overline{x}_i = 3, \quad \mathrm{mod}(i,4) = 1$

$\overline{x}_i = -1, \quad \mathrm{mod}(i,4) = 2$

$\overline{x}_i = \quad 0, \quad \mathrm{mod}(i,4) = 3$

$\overline{x}_i = \quad 1, \quad \mathrm{mod}(i,4) = 0$

Problem 4. Chained Cragg-Levy function with tridiagonal constraints.

$$
\begin{aligned}
F(x) &= \sum_{i=1}^{n/2} [(\exp(x_{2i-1}) - x_{2i})^4 + 100(x_{2i} - x_{2i+1})^6 + \tan^4(x_{2i+1} - x_{2i+2}) + x_{2i-1}^8 \\
&+ (x_{2i+2} - 1)^2] \\
c_k(x) &= 8x_{k+1}(x_{k+1}^2 - x_k) - 2(1 - x_{k+1}) + 4(x_{k+1} - x_{k+2}^2)
\end{aligned}
$$

$1 \le k \le m = n - 2$

$\overline{x}_i = 1, \quad i = 1$
$\overline{x}_i = 2, \quad i > 1$

Problem 5. Generalized Broyden tridiagonal function with five diagonal constraints.

$$
\begin{aligned}
F(x) &= \sum_{i=1}^{n} |(3 - 2x_i)x_i - x_{i-1} - x_{i+1} + 1|^p \\
c_k(x) &= 8x_{k+2}(x_{k+2}^2 - x_{k+1}) - 2(1 - x_{k+2}) + 4(x_{k+2} - x_{k+3}^2) + x_{k+1}^2 - x_k \\
&+ x_{k+3} - x_{k+4}^2
\end{aligned}
$$

$p = 7/3, \quad x_0 = x_{n+1} = 0, \quad 1 \le k \le m = n - 4$

$\overline{x}_i = -1, \quad \forall i$

Problem 6. Generalized Broyden banded function with exponential constraints.

$$
\begin{aligned}
F(x) &= \sum_{i=1}^{n} |(2 + 5x_i^2)x_i + 1 + \sum_{j=\max(1,i-5)}^{\min(n,i+1)} x_j(1 + x_j)|^p \\
c_k(x) &= 4x_{2k} - (x_{2k-1} - x_{2k+1}) \exp(x_{2k-1} - x_{2k} - x_{2k+1}) - 3
\end{aligned}
$$

$p = 7/3, \quad 1 \le k \le m = n/2$

$\overline{x}_i = -1, \quad \forall i$

Problem 7. Trigonometric tridiagonal function with simplified five-diagonal constraints.

$$
F(x) = \sum_{i=1}^{n} |n + i(1 - \cos x_i) - \sin x_{i+1} + \sin x_{i-1}|
$$

$$
\begin{aligned}
c_1(x) &= 4(x_1 - x_2^2) + x_2 - x_3^2 \\
c_2(x) &= 8x_2(x_2^2 - x_1) - 2(1 - x_2) + 4(x_2 - x_3^2) + x_3 - x_4^2 \\
c_3(x) &= 8x_{n-1}(x_{n-1}^2 - x_{n-2}) - 2(1 - x_{n-1}) + 4(x_{n-1} - x_n^2) + x_{n-2}^2 - x_{n-3} \\
c_4(x) &= 8x_n(x_n^2 - x_{n-1}) - 2(1 - x_n) + x_{n-1}^2 - x_{n-2}
\end{aligned}
$$

$\overline{x}_i = 1, \ \forall i$

Problem 8. Augmented Lagrangian function with discrete boundary value constraints.

$$
\begin{aligned}
F(x) &= \sum_{i=1}^{n/5}[\exp(\prod_{j=1}^{5} x_{5i+1-j}) + 10((\sum_{j=1}^{5} x_{5i+1-j}^2 - 10 - \lambda_1)^2 \\
&+ (x_{5i-3}x_{5i-2} - 5x_{5i-1}x_{5i} - \lambda_2)^2 + (x_{5i-4}^3 + x_{5i-3}^3 + 1 - \lambda_3)^2)] \\
c_k(x) &= 2x_{k+1} + h^2(x_{k+1} + h(k+1) + 1)^3/2 - x_k - x_{k+2}
\end{aligned}
$$

$\lambda_1 = -0.002008, \quad \lambda_2 = -0.001900, \quad \lambda_3 = -0.000261, \quad h = 1/(n+1), \quad 1 \le k \le m = n-2$

$\overline{x}_i = -1, \quad \mathrm{mod}(i,2) = 1$

$\overline{x}_i = \quad 2, \quad \mathrm{mod}(i,2) = 0$

Problem 9. Modified Brown function with simplified seven-diagonal constraints.

$$
\begin{aligned}
F(x) &= \sum_{i=1}^{n/2}[(x_{2i-1} - 3)^2/1000 - (x_{2i-1} - x_{2i}) + \exp(20(x_{2i-1} - x_{2i}))] \\
c_1(x) &= 4(x_1 - x_2^2) + x_2 - x_3^2 + x_3 - x_4^2 \\
c_2(x) &= 8x_2(x_2^2 - x_1) - 2(1 - x_2) + 4(x_2 - x_3^2) + x_1^2 + x_3 - x_4^2 + x_4 - x_5^2 \\
c_3(x) &= 8x_3(x_3^2 - x_2) - 2(1 - x_3) + 4(x_3 - x_4^2) + x_2^2 - x_1 + x_4 - x_5^2 + x_1^2 + x_5 - x_6^2 \\
c_4(x) &= 8x_{n-2}(x_{n-2}^2 - x_{n-3}) - 2(1 - x_{n-2}) + 4(x_{n-2} - x_{k+1}^2) + x_{n-3}^2 - x_{n-4} \\
&+ x_{n-1} - x_n^2 + x_{n-4}^2 + x_n - x_{n-5} \\
c_5(x) &= 8x_{n-1}(x_{n-1}^2 - x_{n-2}) - 2(1 - x_{n-1}) + 4(x_{n-1} - x_n^2) + x_{n-2}^2 - x_{n-3} \\
&+ x_n + x_{k-2}^2 - x_{k-3} \\
c_6(x) &= 8x_n(x_n^2 - x_{n-1}) - 2(1 - x_n) + x_{n-1}^2 - x_{n-2} + x_{n-2}^2 - x_{n-3}
\end{aligned}
$$

$\overline{x}_i = -1, \ \forall i$

Problem 10. Generalized Brown function with Broyden tridiagonal constraints.

$$
\begin{aligned}
F(x) &= \sum_{i=1}^{n/2}[(x_{2i-1}^2)^{(x_{2i}^2+1)} + (x_{2i}^2)^{(x_{2i-1}^2+1)}] \\
c_k(x) &= (3 - 2x_{k+1})x_{k+1} + 1 - x_k - 2x_{k+2}
\end{aligned}
$$

$1 \le k \le m = n - 2$

$\overline{x}_i = -1, \quad \mathrm{mod}(i,2) = 1$

$\overline{x}_i = \quad 1, \quad \mathrm{mod}(i,2) = 0$

Problem 11. Chained HS46 problem.

$$
\begin{aligned}
F(x) &= \sum_{i=1}^{(n-2)/3} [(x_{j+1} - x_{j+2})^2 + (x_{j+3} - 1)^2 + (x_{j+4} - 1)^4 + (x_{j+5} - 1)^6] \\
c_k(x) &= x_{l+1}^2 x_{l+4} + \sin(x_{l+4} - x_{l+5}) - 1 \ , \ \mathrm{mod}\ (k,2) = 1 \\
c_k(x) &= x_{l+2} + x_{l+3}^4 x_{l+4}^2 - 2 \qquad\qquad , \ \mathrm{mod}(k,2) = 0
\end{aligned}
$$

$j = 3(i-1), \quad l = 3\ \mathrm{div}(k-1,2), \quad 1 \le k \le m = 2(n-2)/3$

$\overline{x}_i = \quad 2.0, \quad \mathrm{mod}(i,3) = 1$

$\overline{x}_i = \quad 1.5, \quad \mathrm{mod}(i,3) = 2$

$\overline{x}_i = \quad 0.5, \quad \mathrm{mod}(i,3) = 0$

Problem 12. Chained HS47 problem.

$$
\begin{aligned}
F(x) &= \sum_{i=1}^{(n-1)/4} [(x_{j+1} - x_{j+2})^2 + (x_{j+2} - x_{j+3})^2 + (x_{j+3} - x_{j+4})^4 + (x_{j+4} - x_{j+5})^4] \\
c_k(x) &= x_{l+1} + x_{l+2}^2 + x_{l+3}^2 - 3 \ , \ \mathrm{mod}(k,3) = 1 \\
c_k(x) &= x_{l+2} + x_{l+3}^2 + x_{l+4} - 1 \ , \ \mathrm{mod}(k,3) = 2 \\
c_k(x) &= x_{l+1} x_{l+5} - 1 \qquad\qquad , \ \mathrm{mod}(k,3) = 0
\end{aligned}
$$

$j = 4(i-1), \quad l = 4\ \mathrm{div}(k-1,3), \quad 1 \le k \le m = 3(n-1)/4$

$\overline{x}_i = \quad 2.0, \quad \mathrm{mod}(i,4) = 1$

$\overline{x}_i = \quad 1.5, \quad \mathrm{mod}(i,4) = 2$

$\overline{x}_i = -1.0, \quad \mathrm{mod}(i,4) = 3$

$\overline{x}_i = \quad 0.5, \quad \mathrm{mod}(i,4) = 0$

Problem 13. Chained modified HS48 problem.

$$
\begin{aligned}
F(x) &= \sum_{i=1}^{(n-2)/3} [(x_{j+1} - 1)^2 + (x_{j+2} - x_{j+3})^2 + (x_{j+4} - x_{j+5})^4] \\
c_k(x) &= x_{l+1} + x_{l+2}^2 + x_{l+3} + x_{l+4} + x_{l+5} - 5 \ , \ \mathrm{mod}(k,2) = 1 \\
c_k(x) &= x_{l+3}^2 - 2(x_{l+4} + x_{l+5}) - 3 \qquad\qquad , \ \mathrm{mod}(k,3) = 0
\end{aligned}
$$

$j = 3(i-1), \quad l = 3\ \mathrm{div}(k-1,2), \quad 1 \le k \le m = 2(n-2)/3$

$$\overline{x}_i = 3.0, \quad \mathrm{mod}(i,3) = 1$$
$$\overline{x}_i = 5.0, \quad \mathrm{mod}(i,3) = 2$$
$$\overline{x}_i = -3.0, \quad \mathrm{mod}(i,3) = 0$$

Problem 14. Chained modified HS49 problem.

$$\begin{aligned}
F(x) &= \sum_{i=1}^{(n-2)/3} [(x_{j+1} - x_{j+2})^2 + (x_{j+3} - 1)^2 + (x_{j+4} - 1)^4 + (x_{j+5} - 1)^6] \\
c_k(x) &= x_{l+1}^2 + x_{l+2} + x_{l+3} + 4x_{l+4} - 7 \quad , \ \mathrm{mod}(k,2) = 1 \\
c_k(x) &= x_{l+3}^2 - 5x_{l+5} - 6 \quad , \ \mathrm{mod}(k,3) = 0
\end{aligned}$$

$$j = 3(i-1), \quad l = 3 \ \mathrm{div}(k-1,2), \quad 1 \le k \le m = 2(n-2)/3$$
$$\overline{x}_i = 10., \quad \mathrm{mod}(i,3) = 1$$
$$\overline{x}_i = 7.0, \quad \mathrm{mod}(i,3) = 2$$
$$\overline{x}_i = -3.0, \quad \mathrm{mod}(i,3) = 0$$

Problem 15. Chained modified HS50 problem.

$$\begin{aligned}
F(x) &= \sum_{i=1}^{(n-1)/4} [(x_{j+1} - x_{j+2})^2 + (x_{j+2} - x_{j+3})^2 + (x_{j+3} - x_{j+4})^4 + (x_{j+4} - x_{j+5})^4] \\
c_k(x) &= x_{l+1}^2 + 2x_{l+2} + 3x_{l+3} - 6 \quad , \ \mathrm{mod}(k,3) = 1 \\
c_k(x) &= x_{l+2}^2 + 2x_{l+3} + 3x_{l+4} - 6 \quad , \ \mathrm{mod}(k,3) = 2 \\
c_k(x) &= x_{l+3}^2 + 2x_{l+4} + 3x_{l+5} - 6 \quad , \ \mathrm{mod}(k,3) = 0
\end{aligned}$$

$$j = 4(i-1), \quad l = 4 \ \mathrm{div}(k-1,3), \quad 1 \le k \le m = 3(n-1)/4$$
$$\overline{x}_i = 35., \quad \mathrm{mod}(i,4) = 1$$
$$\overline{x}_i = -31., \quad \mathrm{mod}(i,4) = 2$$
$$\overline{x}_i = 11., \quad \mathrm{mod}(i,4) = 3$$
$$\overline{x}_i = -5.0, \quad \mathrm{mod}(i,4) = 0$$

Problem 16. Chained modified HS51 problem.

$$\begin{aligned}
F(x) &= \sum_{i=1}^{(n-1)/4} [(x_{j+1} - x_{j+2})^4 + (x_{j+2} + x_{j+3} - 2)^2 + (x_{j+4} - 1)^2 + (x_{j+5} - 1)^2] \\
c_k(x) &= x_{l+1}^2 + 3x_{l+2} - 4 \quad , \ \mathrm{mod}(k,3) = 1
\end{aligned}$$

$$
\begin{aligned}
c_k(x) &= x_{l+3}^2 + x_{l+4} - 2x_{l+5} &&, \ \mathrm{mod}(k,3) = 2 \\
c_k(x) &= x_{l+2}^2 - x_{l+5} &&, \ \mathrm{mod}(k,3) = 0
\end{aligned}
$$

$j = 4(i-1), \ \ l = 4 \ \mathrm{div}(k-1,3), \ \ 1 \le k \le m = 3(n-1)/4$

$\overline{x}_i = \ \ 2.5, \ \ \mathrm{mod}(i,4) = 1$

$\overline{x}_i = \ \ 0.5, \ \ \mathrm{mod}(i,4) = 2$

$\overline{x}_i = \ \ 2.0, \ \ \mathrm{mod}(i,4) = 3$

$\overline{x}_i = -1.0, \ \ \mathrm{mod}(i,4) = 0$

Problem 17. Chained modified HS52 problem.

$$
\begin{aligned}
F(x) &= \sum_{i=1}^{(n-1)/4} [(4x_{j+1} - x_{j+2})^2 + (x_{j+2} + x_{j+3} - 2)^4 + (x_{j+4} - 1)^2 + (x_{j+5} - 1)^2] \\
c_k(x) &= x_{l+1}^2 + 3x_{l+2} &&, \ \mathrm{mod}(k,3) = 1 \\
c_k(x) &= x_{l+3}^2 + x_{l+4} - 2x_{l+5} &&, \ \mathrm{mod}(k,3) = 2 \\
c_k(x) &= x_{l+2}^2 - x_{l+5} &&, \ \mathrm{mod}(k,3) = 0
\end{aligned}
$$

$j = 4(i-1), \ \ l = 4 \ \mathrm{div}(k-1,3), \ \ 1 \le k \le m = 3(n-1)/4$

$\overline{x}_i \ = 2, \ \ \forall i$

Problem 18. Chained modified HS53 problem.

$$
\begin{aligned}
F(x) &= \sum_{i=1}^{(n-1)/4} [(x_{j+1} - x_{j+2})^4 + (x_{j+2} + x_{j+3} - 2)^2 + (x_{j+4} - 1)^2 + (x_{j+5} - 1)^2] \\
c_k(x) &= x_{l+1}^2 + 3x_{l+2} &&, \ \mathrm{mod}(k,3) = 1 \\
c_k(x) &= x_{l+3}^2 + x_{l+4} - 2x_{l+5} &&, \ \mathrm{mod}(k,3) = 2 \\
c_k(x) &= x_{l+2}^2 - x_{l+5} &&, \ \mathrm{mod}(k,3) = 0
\end{aligned}
$$

$j = 4(i-1), \ \ l = 4 \ \mathrm{div}(k-1,3), \ \ 1 \le k \le m = 3(n-1)/4$

$\overline{x}_i \ = 2, \ \ \forall i$

Bibliography

[1] J.R.Bunch, B.N.Parlett: Direct methods for solving symmetric indefinite systems of linear equations. SIAM J. Numerical Analysis 8 (1971) 639-655.

[2] A.R.Curtis, M.J.D.Powell, J.K.Reid: On the estimation of sparse Jacobian matrices. IMA Journal of Aplied Mathematics 13 (1974) 117-119.

[3] R.S.Dembo, S.C.Eisenstat, T.Steihaug: Inexact Newton methods. SIAM J. Numerical Analysis 19 (1982) 400-408.

[4] R.Fletcher: Practical methods of optimization. Wiley, Chichester, 1987.

[5] P.E.Gill, W.Murray: Newton type methods for nconstrained and linearly constrained optimization. Mathematical Programming 7 (1974) 311-350.

[6] N.I.M.Gould: On practical conditions for the existence and uniqueness of solutions to the general equality quadratic programming problem. Mathematical Programming 32 (1985) 90-99.

[7] W.Hock, K.Schitkowski: Test Examples for Nonlinear Programming Codes. Lecture Notes in Economics and Mathematical Systems, Vol. 187. Springer-Verlag, Berlin, 1981.

[8] Lukšan, L., Šiška, M., Tůma, M., Vlček, J., and Ramešová, N., *Interactive System for Universal Functional Optimization (UFO), Version 1995,* Research Report No. V-662, Institute of Computer Science, Academy of Sciences of the Czech Republic, Prague, Czech Republic, 1995.

[9] P.S.Vassilevski, D.Lazarov: Preconditioning mixed finite element saddle-point elliptic problems. Numerical Linear Algebra with Applications 3 (1996) 1-20.

[10] R.Weiss: Convergence behavior of generalized conjugate gradient methods. Thesis. University of Karlsruhe, 1990.