Computer Science Journal of Moldova, vol.**3**, no.3(9), 1995

Quasi-Newton Methods for Solving Nonlinear Programming Problems

V.Moraru

Abstract

In the present paper the problem of constrained equality optimization is reduced to sequential solving a series of problems of quadratic programming. The Hessian of the Lagrangian is approximated by a sequence of symmetric positive definite matrices. The matrix approximation is updated at every iteration by a Gram-Schmidt modified algorithm. We establish that methods is locally convergent and the sequence $\{x_k\}$ converges to the solution a two-step superlinear rate.

Key words: Quasi-Newton methods, Constrained Optimization, Superlinear convergence.

1 Introduction

This paper considers the methods of the finding of a vector $x_* \in R^n$, which is the solution of the nonlinear programming problem:

$$minimize f(x)$$

$$subject \quad to \quad h_i(x) = 0, \quad 1 \le i \le r. \qquad (1)$$

It is supposed that the following conditions are fulfiled:

1) the functions $f(x) : R^n \to R^1, \; h_i(x) : R^n \to R^1$ are continuously differentiated twice in some neighbourhood of the point x_*;

2) the gradients $\nabla h_1(x_*), \nabla h_2(x_*), \ldots, \nabla h_r(x_*)$ form a linear independent sistem;

3) $\nabla_x L(x_*, \lambda_*) = 0$, where $\nabla_x L(x, \lambda)$ is a vector of the first derivatives by x from the Lagrangian function

$$L(x, \lambda) = f(x) + \sum_{i=1}^{r} \lambda^i h_i(x);$$

4) the matrix $\nabla^2_{xx} L(x_*, \lambda_*)$ is positive definite on the null spase of the constraint gradients, i.e.

$$\eta^T \nabla^2_{xx} L(x_*, \lambda_*)\eta \geq m\|\eta\|^2, 0 < m < \infty,$$

at any vector $\eta \in R^n, \eta \neq 0$, such as $\nabla h(x_*)\eta = 0$, where $\nabla^2_{xx} L(x, \lambda)$ is a matrix of the second derivatives of the Lagrangian function $L(x, \lambda)$ by $x; \nabla h(x)$ is a matrix, the lines of which are the gradients $\nabla h_i(x), 1 \leq i \leq r$. The symbol "$T$" denotes the transpose of a vector or a matrix.

If the conditions 1) to 4) are fulfiled x_* is the point of the strict local minimum of the function $f(x)$ with the restrictions (1) (see for example [1], p.42–49). Besides that, let us apply the Newton's method to the sistem of the nonlinear equations (see [2], p.275):

$$\left.\begin{array}{l} \nabla_x L(z) = 0, \\ h_1(x) = 0, \\ \cdots\cdots\cdots\cdots \\ h_r(x) = 0. \end{array}\right\} \tag{2}$$

Here $z = (x, \lambda)^T$. Throughout the paper we will denote by $z_k = (x_k, \lambda_k)^T$ and by $z_* = (x_*, \lambda_*)^T$ the Kuhn-Tucker pair for problem (1).

The process, caused by the Newton's method, while applying it to the sistem (2) is equivalent (see [3–8] to the solution of the quadratic programming problem of minimization a the function

$$Q(x, \lambda_k) = \frac{1}{2}(x - x_k)^T \nabla^2_{xx} L(z_k)(x - x_k) + \nabla f(x_k)^T (x - x_k) \tag{3}$$

subject to the linear constraints

$$h_i(x_k) + \nabla h_i(x_k)^T (x - x_k) = 0, \quad 1 \le i \le r. \tag{4}$$

Here λ_{k+1} will be the vector of the Lagrange's multipliers in the extreme point x_{k+1} of the problem $(3),(4)$, i. e. that is x_{k+1} and λ_{k+1} are calculated by x_k and λ_k as a solution and a vector of the Lagrange's multipliers of the problem $(3),(4)$ connected with it. If the initial approximation z_0 is chosen from the sufficiently small neighbourhood of the point z_*, the sequence $x_0, x_1, \ldots, x_k, \ldots$ converges to x_* and besides that it converges with a superlinear rate (see more precise result in [6,9,10]).

One of the main disadvantages of this method is that the function's $Q(x, \lambda_k)$ convexity is not guaranteed. Another one is that while determing every new approximation it is nessesary to calculate the Hesse matrix of the Lagrangian function $L(x, \lambda)$ at the point which coresponds to the precedent approximation. That is why a question appears how to construct the methods, which do not require the calculation of the matrix $\nabla^2_{xx} L(x_k, \lambda_k)$ and provide the convexity of the auxiliary problems $(3),(4)$ keeping the high velocity of convergency. At present, such a method is elaborated, based on using of the variable metric algorithms (see [6,11–15]). In this papers the matrices $\nabla^2_{xx} L(x_k, \lambda_k), k = 0, 1, 2, \ldots,$ are replaced by some other matrices A_k, which are positively definite on the space R^n and approximate the matrix $\nabla^2_{xx} L(x_*, \lambda_*)$ only on the tangent subspace to the constraints of the problem (1) in the point x_*. The construction of the matrices $A_k, k = 0, 1, 2, \ldots,$ are accomplished with help of the variable metric algorithms.

In our work we present methods, in which on each iterates the convex problems of the quadratic programming are solved and the approximation of the requiered matrices is accomplished by the methods, examined in the papers [16–18] will be studied.

2 The method and its proprieties

Consider a matrix

$$M(z) = \nabla^2_{xx} L(z) + \nabla h(x)^T D \nabla h(x), \tag{5}$$

where D is some bounded symmetric and positively definite matrix of the dimension $r \times r$.

The matrix $\nabla h(x)^T D \nabla h(x)$ is positively definite on the set

$$\{y \in R^n : \nabla h(x) y \neq 0\}.$$

According to the assumptions done in Section 1 it is possible to select a matrix D such, that the matrix $M(z)$, which is defined by formula (5) would be positively definite on R^n in some neighbourhood of the point z_*. More detailed about the selection of the matrix D will be said laver in Section 3.

Further, every where will consider, that the matrix D is selected such that

$$y^T M(z) y \geq m\|y\|^2, \forall y \in R^n, m > 0, \tag{6}$$

for all z from the small'enough neighbourhood of the point z_*, marked by Ω.

Let $z_0 \in \Omega$. Consider an iteration process in which z_{k+1} is defined according to z_k from the solution of the problem of minimization of the quadratic function

$$\tilde{Q}(x, \lambda_k) = \frac{1}{2}(x - x_k)^T M(z_k)(x - x_k) + \nabla f(x_k)^T (x - x_k) \tag{7}$$

with the constraints (4).

By the solution of the quadratic programming problem $(7), (4)$ it is understood the same as for the problem $(3), (4)$. This process of the construction of the sequence $\{z_k\}$ will name further as the method $(7), (4)$.

Lemma 1 *The sequence* $\{z_k\}$ *constructed according to the method* $(7), (4)$ *converges to* z_* *at superlinear rate.*

Proof is done analogically to the proof of the Newton's method of solving the systems of equation [19]. In the considered case

$$z_{k+1} = z_k - [B(z_k)]^{-1} F(z_k), \tag{8}$$

where

$$B(z_k) = \begin{bmatrix} M(z_k) & \nabla h(x_k)^T \\ \nabla h(x_k) & 0 \end{bmatrix}$$

and

$$F(z_k) = \begin{bmatrix} \nabla_x L(z_k) \\ h_1(x_k) \\ \cdots\cdots\cdots \\ h_r(x_k) \end{bmatrix}$$

The matrix $B(z_k)$ is nonsingular accorrding to the assumptions 1)–4) from Section 1. It is possible to prove that in the same way as it was done in [2], p.275. Hence, on Ω is correctly defined the operator

$$G(z) = z - [B(z)]^{-1} F(z) \tag{9}$$

and besides

$$G'(z_*) = I - [B(z_*)]^{-1} \Im(z_*),$$

where I is unit matrix of the dimension $(n+r) \times (n+r)$ and $\Im$ is the $(n+r) \times (n+r)$ Jacobian matrix of F.

It is easy to convinced (see [20] p. 252)

$$[B(z)]^{-1} = [\Im(z)]^{-1} - \begin{bmatrix} 0 & 0 \\ 0 & D \end{bmatrix} \tag{10}$$

Taking this in consideration we'll obtain:

$$G'(z_*) = \begin{bmatrix} 0 & 0 \\ D\nabla h(x_*) & 0 \end{bmatrix}$$

that is the matrix $G'(z_*)$ is a triangular matrix, and it means that it's eigenvalues $\mu_1, \mu_2, \ldots, \mu_{n+r}$ exactly coincide with the diagonal elements. But $r \leq n$, because in the opposite case the system of vector $\{\nabla h_i(x_*)\}_{i=1}^r$ will not be linearly independent. Hence it follows, that

$$\mu_1 = \mu_2 = \ldots = \mu_{n+r} = 0.$$

According to the definition of the spectral radius

$$\rho(G'(z_*)) = max\{|\mu_1|, \ldots, |\mu_{n+r}|\} = 0. \tag{11}$$

Now to finish the proof of the lemma, it is necessary to use the theorem of Ostrowsky, which proves: if $\rho(G'(z_*)) < 1$ then the $G(z)$ which is defined by formula (9) is a contraction. The process (8) is nothing else that a method of successive approximation for the solution of the equations $G(z) = z$, that is finding of immovable point of the operator $G(z)$.

The equality (11) quarantees that the sequence $\{z_*\}$ converges to the point z_*, more rapidly than any geometrical progression. Indeed, for $\forall$ given $\varepsilon/2 > 0$ exist a norma $\| \bullet \|$ in E^n (see, for example, [19], p. 47) such that

$$\|G'(z_*)\| \leq \rho(G'(z_*)) + \frac{\varepsilon}{2} = \frac{\varepsilon}{2},$$

according to (11). For arbitrary $\varepsilon/2 > 0$ according to differentiation of $G(z)$ it is possible to indicate a number $N = N(\varepsilon)$ such that

$$\|G(z_k) - G(z_*) - G'(z_*)(z_k - z_*)\| \leq \frac{\varepsilon}{2}\|(z_k - z_*)\|,$$

when $k \geq N$. Taking this in consideration, for $k \geq N$

$$\|z_{k+1} - z_*\| \leq \|G(z_k) - G(z_*) - G'(z_*)(z_k - z_*)\| +$$
$$+ \|G'(z_*)\| \ \|z_k - z_*\| \leq \varepsilon\|z_k - z_*\|.$$

Lemma is proved.

Remark. If $h_i(x), 1 \leq i \leq r$ are the linear functions then $x_k \to x_*$ superlinearly. In fact, in this case from (8) taking in consideration (10) will have

$$z_{k+1} = z_k - [F'(z_k)]^{-1} F(z_k),$$

that is the matrix $\nabla h(x_k)^T D_k \nabla h(x_k)$ does not participate in the construction of the vector z_{k+1}. In the common case it is impossible to get a superlinear convergence for the sequence $\{x_k\}$.

The construction of the vector x_{k+1}, which is the solution of the problem $(7),(4)$ does not depend on the selection of the matrix D_k. This matrix influences sufficiently only on the properties of the sequence $\{\lambda_k\}$. It is possible to snow that in the sollouing way. Let the point z_k be constucted and let $\bar{x}_{k+1}$, $\bar{\lambda}_{k+1}$ be the solution of the problem $(3),(4)$. Taking in consideration (10) it's easy to see, that the solution of the problem $(7),(4)$ is connected with $\bar{x}_{k+1}, \bar{\lambda}_{k+1}$ in the following way: $x_{k+1} = \bar{x}_{k+1}, \lambda_{k+1} = \bar{\lambda}_{k+1} - D_k h(x_k)$. From this follows, that the selection of the matrix D_k directly shoues up when constructing vector λ_{k+1}. This selection finally will influence on the properties of the sequence $\{x_k\}$ because the value of λ_k is required while forming the problems $(7),(4)$. The character of convergence is set in the following lemma.

Lemma 2 *For the sequence* $\{x_k\}$*, constructed by method* $(7),(4)$ *the following appraisal is true*

$$\|x_{k+1} - x_*\| \leq \mu_k \|x_{k-1} - x_*\|, \tag{12}$$

where $\mu_k \to 0$*, while* $k \to \infty$*.*

Proof. Let

$$P_k = I - \nabla h(x_k)^T [\nabla h(x_k) \nabla h(x_k)^T]^{-1} \nabla h(x_k).$$

In the paper [6] the following statement is set if the matrix $M(z_k)$ in the method $(7),(4)$ is such that

$$\lim_{k\to\infty}\frac{\|P_k(M(z_k)-\nabla^2_{xx}L(z_*))P_k(x_{k+1}-x_k)\|}{\|x_{k+1}-x_k\|}=0 \qquad (13)$$

then for the sequence $\{x_k\}$ the appraisal (12) is true. To prove the lemma it is enough to show, that for the matrix $M(z_k)$, selected in the form (5) the condition (13) is true.

In fact for any vector $y \in E^n$

$$\nabla h(x_k)P_k y = 0.$$

Hence,

$$(M(z_k) - \nabla^2_{xx}L(z_k))P_k y = 0.$$

Taking this in consideration, after some transfofmations, we obtain

$$\|P_k(M(z_k)-\nabla^2_{xx}L(z_*))P_k(x_{k+1}-x_k)\| =$$
$$= \|P_k(\nabla^2_{xx}L(z_k)-\nabla^2_{xx}L(z_*))P_k(x_{k+1}-x_k)\| \le \gamma_k\|x_{k+1}-x_k\|,$$

where

$$\gamma_k = \|\nabla^2_{xx}L(z_k)-\nabla^2_{xx}L(z_*)\| \to 0,$$

while $k\to\infty$, because $z_k \to z_*$, according to lemma I.

3 The Description of the Quasi-Newton Algorithms

The properties of the method (7), (4) will be kept, if the matrix $M(z_k)$ will be replaced by some other matrix A_k, which is sufficiently close to the matrix $M(z_k)$. Let's pass now to the construction of the matrices A_k (which approximate $M(z_k)$, using the methods studied in [18].

It's necessary to note that the problem of construction of these matrices is complicated, because obviosly there does not exist function, the matrix of the second derivatives of which (calculated in the point z_k) would be exactly equal to $M(z_k)$ because the second addendum in the matrix $M(z_k)$ contains only the first derivatives of the functions $h_i(x), 1 \leq i \leq r$.

However one can take the function

$$\Psi(z) = L(z) + \frac{1}{2}\|\sqrt{D}h(x)\|^2.$$

In fact

$$\nabla_x \Psi(z) = \nabla_x L(z) + \nabla h(x)^T D h(x),$$

$$\nabla^2_{xx} \Psi(z) = \nabla^2_{xx} L(z) + \nabla h(x)^T D \nabla h(x) + h''(x)^T D h(x).$$

Near the solution x_* it is possible to neglect the matrix $h''(x)^T Dh(x)$. Hense, in the formulae (3.1), (4.1), (5.1), (6.1), (6.3) from the paper [18], for construction of the matrices A_k, it is possible to select as e_{k-i}, a vector

$$e_{k-i} = \nabla_x \Psi(x_{k-i} + r_{k-i}, \lambda_k) - \nabla_x \Psi(x_{k-i}, \lambda_k).$$

Consider another move simple method of construction of the vectors $e_{k-i}, 1 \leq i \leq n$, than the method above. Let e_{k-i} be calculated according to the formulae

$$e_{k-i} = \nabla_x L(x_{k-i} + r_{k-i}, \lambda_k) - \nabla_x L(x_{k-i}, \lambda_k) +$$

$$+ \nabla h(x_k)^T D_{k-i}[h(x_{k-i} + r_{k-i}) - h(x_{k-i})], 1 \leq i \leq n \qquad (14)$$

From the point of view of calculations it is more simple to take as D_{k-i} a product of the unit matrix I with a positive number σ_{k-i}:

$$D_{k-i} = \sigma_{k-i} I.$$

In this case

$$(e_{k-i}, r_{k-i}) = (a_{k-i}, r_{k-i}) + \sigma_{k-i}(b_{k-i}, \nabla h(x_k) r_{k-i}), \qquad (15)$$

where

$$a_{k-i} = \nabla_x L(x_{k-i} + r_{k-i}, \lambda_k) - \nabla_x L(x_{k-i}, \lambda_k),$$

$$b_{k-i} = h(x_{k-i} + r_{k-i}) - h(x_{k-i}).$$

Two cases are possible on each i.

1) $\nabla h(x_k) r_{k-i} = 0$. According to the suppositions made above (see section 1) it means, that the vector r_{k-i} belongs to subspace, on which the matrix $\nabla^2_{xx} L(z_*)$ is positively defined.

   Hence, $(e_{k-i}, r_{k-i}) \geq m \|r_{k-i}\|^2$ on any selection of σ_{k-i}. Here, m — a constant that characterize (see section 1) a positive definite of the matrix $\nabla^2_{xx} L(z_*)$.

2) $\nabla h(x_k) r_{k-i} \neq 0$. Using the Lagrange's formula for operators, it is possible to set, that the ratio $(a_{k-i}, r_{k-i})/(b_{k-i}, \nabla h(x_k) r_{k-i})$ is bounded. Let

$$\sigma_{k-i} = \delta - (a_{k-i}, r_{k-i})/(b_{k-i}, \nabla h(x_k) r_{k-i}), \qquad (16)$$

   where δ is an arbitrary positive number, chosen such that $\sigma_{k-i} > 0$.

Then from (15) we obtain

$$(e_{k-i}, r_{k-i}) = \delta (b_{k-i}, \nabla h(x_k) r_{k-i}) \qquad (17)$$

While the gradients of the functions $h_i(x), 1 \leq i \leq r$ are linearly independent in the neighbourhood of the points x_* the following appraisal is true

$$(e_{k-i}, r_{k-i}) \geq \delta m_1 \|r_{k-i}\|^2, 0 < m_1 < \infty.$$

Basing on that, it is possible now to construct the matrixes A_k, in the following way [18,21,22]. Suppose, that $k = \xi n + s, \xi = 0, 1, 2, \ldots, 0 \leq s \leq n-1$. We will denote $w_{\xi n+s} = \eta_{\xi n+s} v_{s+1}$, where $v_1, v_2, \ldots, v_n$ is a system of unit vectors. Let us calculate the matrix

$$A_k = \sum_{i=0}^{n-1} \frac{e_{k-i} e_{k-i}^T}{r_{k-i}^T e_{k-i}}. \tag{18}$$

Here the vectors e_{k-i} are chosen according to the formula (14), where

$$D_{k-i} = \sigma_{k-i} I,$$

where σ_{k-i} is either positive number or it is chosen from the condition (16).

The construction of the vectors r_{k-i} is performed according to the formulae

$$r_{\xi n} = w_{\xi n}, \quad r_k = w_k - \sum_{j=o}^{s-1} \frac{w_k^T e_{\xi n+j}}{e_{\xi n+j}^T r_{\xi n+j}} r_{\xi n+j}.$$

The factors η_k are chosen so that to satisfy the following condition $0 < c \le |\eta_{k+1}|/|\eta_k| \le 1$ and $lim_{k \to 0} \eta_k = 0$. The vectors r_k, e_k are constructed recursively while using Gram-Schmidt modified algorithm [18,21,22].

Let's note, that if $(a_{k-i}, r_{k-i}) \le 0$ then it is possible to choose as δ in (16) any positive number but the scalar product (e_{k-i}, r_{k-i}) could be calculated directly by formula (17).

Analogically to that in [16,17] the verity of the following appraisals is set

$$(r_{k-i}, e_{k-j}) = \circ(\|r_{k-i}\| \|e_{k-j}\|), \qquad i \ne j. \tag{19}$$

Taking in consideration these appraials we'll have

$$A_k r_{k-i} = e_{k-i} + \mu_{k-i}, 1 \le i \le n,$$

where $\|\mu_{k-i}\| \to 0$, while $k \to \infty$.

Repeating now the reasoning from the lemma 2.3.1.from [2], we'll set that

$$\lim_{k \to \infty} \|A_k - M(z_k)\| = 0 \tag{20}$$

If in adition the condition of Lipshitz for matrix $\nabla^2_{xx}L(x,\lambda)$ concerning x holds then the following appraisal is true:

$$\|A_k - M(z_k)\| \leq c|\eta_k|, 0 < c < \infty, \tag{21}$$

where η_k — a number used in the definition of the vectors r_{k-i}, (see [18]).

If the condition (19) holds true, then

$$\lim_{k\to\infty} \|(A_k - \nabla^2_{xx}L(z_k))q\| = 0 \tag{22}$$

on any vector q such that $\nabla h(x_k)q = 0$.

Hence, the matrices A_k approximate $\nabla^2_{xx}L(z)$ on the tangent subspace to the multitude $\{x|h_i(x) = 0, \quad 1 \leq i \leq r\}$ at the point x_k, that is on that subspace, where $\nabla^2_{xx}L(z)$ is positive definite. Taking in consideration (20), (22), analogically to that as it was done in the proof of lemma 2, it is possible to show, that the considered Quasi-Newton methods will converge with the velocity, evaluated by inecuality (12).

Let us mention the following. It is impossible to use the methods [18] for approximation of the matrix $\nabla^2_{xx}L(z_k)$ because it is posible that at some i will be $(r_i, e_i) = 0$. The exeption is the method of indecisive directions [2]. While using this method, it is better to construct the matrices A_k, which approximate $\nabla^2_{xx}L(z_k)$ directly. To the constructed matrix A_k it is possible to add now a matrix $\nabla h(x_k)^T D_k \nabla h(x_k)$, trying to get a positive definite by choosing the matrix D_k.

It's necessary to mention that it is profitable to ensure the positive definite of requiered matrices, because in this case the auxiliary problems are rather effectively solved.

4 Another method of the constructing of the approximating matrices

Let's consider briefly one more method (that essentially differs from the considered above) of the construction of the matrices A_k, that ensures the subfilment of the condition (22).

According to the suppositions done above, the space

$$R(x_k) \equiv R_k = \{q|\nabla h(x_k)q = 0\}$$

has the dimentions $n-r$. Let $b_1, \ldots, b_{n-r}$ the basis of this space and let A_k is the matrix, defined by formula (18) in which $w_{k-i} = \eta_{k-i}b_i, 1 \leq i \leq n-r$, bat the values $e_{k-i}, 1 \leq i \leq n-r$ are replaced by the expressions a_{k-i} (see (15)) and addition is done by all i from 1 to $n-r$. Taking this constructions in consideration and reasoning in the same way as in [16–19] the verity of the appraisal (19) used, but then supposing that this appraisal holds true the verity of the condition (22) is proved. The matrices A_k satisfy the following condition $(A_k y, y) \geq 0$ for any $y \in E^n$.

Concerning the construction of the basis of subspace R_k it could be obtained in the following way (see for example [23]). With the help of elementary operations and rearrangements the matrix $\nabla h(x_k)$ is reduced to the matrix S_k strairs aspect. Then the system of equations $\nabla h(x_k)q = 0$ is equivalent to the system $S_k q = 0$, which (taking in consideration that rang $(\nabla h(x_k)) = r$) has $n-r$ fred variables. Giving in turn to each of free variables the value 1, but to the rest of free variables the values 0 and solving the system $S_k q = 0$ relative to the remained variables, we'll obtain search basis $b_1, \ldots, b_{n-r}$ of the space R_k.

References

[1] Fiacco A.V., McCormick G.P. Nonlinear programming: Sequential unconstrained minimization techniques. John Wiley & Sons, New York, 1968.

[2] Pshenichnyi B.N., Danilin Yu.M. Numerical methods in extremal problems. Mir Publishers, Moscow, 1975.

[3] Robinson S.M. A quadratically convergent algorithm for general programming problems. Mathematical Programming, vol.3, 1972, p.145–156.

[4] Garcia Palomares U.M., Mangasarian O.L. Superlinearly convergent quasi — Newton algorithms for nonlinearly constrained optimization problems. Computer Sciences Technical Report 195, University of Wisconsin, Madison (Wisc.), 1974.

[5] Han S.P. Superlinearly convergent variable metric algorithms for general nonlinear programming problems. Mathematical Programming, vol.11, No.3, 1976, p.263–282.

[6] Powell M.J.D. Algorithms for nonlinear constraints that use Lagrangian functions. Mathematical Programming, vol.14, No.2, 1978, p.20–32.

[7] Bertsekas D.P. Constrained optimization and Lagrange multipliers methods. Academic Press, New York, 1982.

[8] Bonnans J.F. An introduction to Newton type algorithms for nonlinearly constrained optimization problems. Internasional Series of Numerical Mathematics, vol.87, Basel, 1989, p.1–17.

[9] Bogs P.T., Tolle J.W., Wang P. On the local convergence of quasi — Newton methods for constrained optimization. SIAM J. Cont. Optimiz., vol.20, 1982, p.161–171.

[10] Byrd R.H. On the convergence of constrained optimization methods with accurate Hessian information on a subspace. SIAM J. Numer. Anal., vol.27, No.1, 1990, p.141–153.

[11] Coleman T.F., Conn A.R. On the local convergence of a quasi — Newton method for the nonliear programming problem. SIAM J. Numer. Anal., vol.21, No.4, 1984, p.755–769.

[12] Wright S.J. Convergence of projected Hessian approximations in quasi — Newton methods for the nonlinear programming problem. IMA J. Numer. Anal., vol.6, No.4, 1986, p.463–474.

[13] Tapia R. On secant updates for use in general constrained optimization. Mathematics of Computation, vol.51, No.183, 1988, p.181–202.

[14] Fontecilla R. Local convergence of secant methods for nonlinear constrained optimization. SIAM J. Numer. Anal., vol.25, No.3, 1988, p.692–712.

[15] Gurwitz C.B., Overton M.L. Sequential quadratic program- ming methods based on approximating a projected Hessian matrix. SIAM J. Sci. Statist. Comput., vol.10, No.4, 1989, p.631–653.

[16] Danilin Yu.M. Metodi sopreajionnih napravlenii, ne trebuiuscie resenia adnomernih zadaci minimizatii. Dokl. Akad. Nauk SSSR, vol.218, No.3, 1974, p.513–516 (Russian)

[17] Bulanyi A.P., Danilin Yu.M. Kvaziniutonovskie algoritmi minimizatii, osnovannie na postroenii sistem sopreajionnih vectorov. J. V. M. i M. F., vol.18, No.4, 1978, p.877-885 (Russian)

[18] Danilin Yu.M., Moraru V.A. Kvaziniutonovskie metodi uslovnoi optimizatii. Kibernetika, No.5, 1979, p.80-86 (Russian)

[19] Ortega J.M., Rheinboldt W.C. Iterative solutions of nonlinear equations in several variables. Academic Press, New-York, 1970.

[20] Numerical methods for constrained optimization. Edited by P.E.Gill and W.Murrey. Academic Press, New York, 1974.

[21] Moraru V.A. Kvaziniutonovskie metodi dlea resenia extremalinih zadaci s ogroniceniami tipa ravenstv. Kibernetika, No.5, 1984, p.124-125 (Russian)

[22] Moraru V. Quasi-Newton methods for constrained optimization. Proceedings of the 9^{th} Romanian Symposium on Computer Science. Iasi, November 12-13, 1993, p.405-408.

[23] Strang G. Linear Algebra and its applications. Academic Press, New-York, 1976.

Vasile Moraru,
Dept. of Computing Mathematics and Programming,
Tehnical University of Moldova
Bd. Stefan cel Mare, 168, Kishinev,
277012, Moldova

Received 15 December, 1995