DOI: 10.13656/j.cnki.gxkx.2004.02.004
广西科学 Guangxi Sciences 2004, 11 (2): 91~ 99, 105

An Approximate Gauss-Newton Based BFGS Method for Solving Symmetric Nonlinear Equations*
解非线性对称方程组问题的近似高斯-牛顿基础 BFGS方法

Wei Zengxin Yuan Gonglin Lian Zhigang
韦增欣 袁功林 连志钢

(College of Math.& Info. Sci., Guangxi Univ., 100 Daxuelu, Nanning, Guangxi, 530004, China)
(广西大学数学与信息科学学院 南宁市大学路 100号 530004)

Abstract An approximate Gauss-Newton based BFGS method for solving symmetric nonlinear equations is presented. The proposed method always generates descent directions whatever linear search is used. The global and superlinear convergence of the proposed method under suitable conditions is proved. Numerical results show that the proposed method is successful.

Key words BFGS method, Gauss-Newton method, symmetric equation, global convergence, superlinear convergence

摘要 给出一个解非线性对称方程组问题的近似高斯-牛顿基础 BFGS方法.该方法无论使用何种线性搜索,此方法产生的方向总是下降的.证明在适当的条件下,该方法的全局收敛性和超线性收敛性,给出数值检验结果.

关键词 BFGS方法 高斯-牛顿方式 对称方程 全局收敛 超线性收敛
中图法分类号 O241.7

1 Introduction

It's well known that the Quasi-Newton methods play a special role in solving unconstrained optimization problems in References [1~ 6]. Some modified BFGS methods with global and superlinear convergence for nonconvex minimization problems have been proposed in References [7~ 11]. We propose an approximate Gauss-Newton based BFGS method which can generate descent directions for the norm function for solving symmetric nonlinear equations in this paper.

Consider a general Quasi-Newton method for solving the following system of nonlinear equations:

$$F(x) = 0, x \in R^n, \qquad (1.1)$$

where $F: R^n \to R^n$ is continuously differentiable. Equation (1.1) can generate a sequence of iterates $\{x_k\}$, $x_{k+1} = x_k + d_k$, where d_k is a solution of the following system of linear equations

$$B_k d_k + F(x_k) = 0. \qquad (1.2)$$

If in Equation (1.2), matrix B_k is replaced by the Jacobian $F'(x_k)$ of the function F at x_k, then the method would be reduced to the well-know Newton method.

$$F'(x_k) d_k + F(x_k) = 0. \qquad (1.3)$$

An interesting feature of Equation (1.3) is its local superlinear convergence property without computation of Jacobians. To enlarge the convergence domain of a Quasi-Newton method, line search technique or trust region strategy can be exploited. In this paper, we use a backtracking line search technique to globalize a Quasi-Newton method.

A scalar $\lambda_k \geqslant 0$ which satisfies the following equation is a line search step at iteration k of an iterative method

$$\|F(x_k + \lambda_k d_k)\| \leqslant \|F(x_k)\|, \qquad (1.4)$$

$x_{k+1} = x_k + \lambda_k d_k$ is the next iterate. Where λ_k is called the steplength. We define the norm function as $\theta(x) = \frac{1}{2}\|F(x)\|^2 x \in R^n$. So the nonlinear equation problem (1.1) is equivalent to the following global optimization problem

$$\min \theta(x), x \in R^n, \qquad (1.5)$$

2003-09-30收稿.
* Supported by National Natural Science Foundation of China (10161002) and Natural Science Foundation of Guangxi (9811020).

and Equation (1.4) is equivalent to the following equation

$$\theta(x_k + \lambda_k d_k) \leqslant \theta(x_k). \quad (1.6)$$

A sequence x_k satisfying Equations (1.4) and (1.6) are generated by an iterative method that is a norm descent method. The inequality (1.6) holds for all $\lambda_k \geqslant 0$ sufficiently small if d_k is a descent direction of θ at x_k. Accordingly, the related iterative method is a norm descent method. In particular, Newton's method is a norm descent method in the sense that $\|g_{k+1}\| \leqslant \|g_k\|$ holds on every iterations. However, if the vector d_k is determined from B_k, then d_k may not be a descent direction of θ at x_k even if B_k is symmetric and positive definite. Therefore, in order to globalize a Quasi-Newton method, by the means of modifying a Quasi-Newton direction, Li and Fukushima presented an approximately norm descent line search technique and established global and superlinear convergence of a Gauss-Newton based BFGS method for solving symmetric nonlinear equations[12].

In this paper, we update B_k by combining a modified BFGS formula such that B_{k+1} inherits positive definiteness of B_k whatever line search is used. We adjust the steplength and the search direction simultaneously so that the generated iterate sequence satisfies Inequality (1.6). Under suitable conditions, we establish a global convergence theorem which shows that there exists an accumulation point that is a stationary point of Problem (1.5) even if $F'(x)$ is singular everywhere. We also get the superlinear convergence of the modified method.

In the next section, we present a way to generate an approximately descent Quasi-Newton direction for θ and then propose a new BFGS method for solving Equation (1.1), and describe some properties of the modified method. In section 3, we propose a modified algorithm. In sections 4 and 5, we establish the global and superlinear convergence of the modified method under some reasonable conditions respectively. In section 6, we report some numerical results.

2 Descent direction by a new Quasi-Newton method

First, we give the following Assumptions.

Assumption 1 (A) The level set

$$\mathrm{K} = \{x \in R^n \mid \theta(x) \leqslant \theta(x_0)\}$$

is bounded.

(B) F is continuously differentiable on an open convex set K_1 containing K, and $F'(x)$ is symmetric for every $x \in \mathrm{K}_1$.

(C) The Jaconbian of F bounded, and there exists a positive constant M_1 such that

$$\|F'(x)\| \leqslant M_1, \forall x \in \mathrm{K}_1. \quad (2.1)$$

(D) The sequence $\{x_k\}$ is contained in a bounded set K_1 in which F is continuously differentiable, and there exists a constant $m, M \geqslant 0$ such that

$$m\|x - y\| \leqslant \|F(x) - F(y)\| \leqslant M\|x - y\|. \quad (2.2)$$

According to (C) and (D) of Assumption 1, for all $x \in \mathrm{K}_1$ and $d \in R^n$, we have

$$m\|x - x^*\| \leqslant \|F(x)\| = \|F(x) - F(x^*)\| \leqslant M\|x - x^*\|, \quad (2.3)$$

$$m\|d\| \leqslant \|F'(x)d\| \leqslant M\|d\|, \quad (2.4)$$

$$\frac{1}{M}\|d\| \leqslant \|F'(x)d\| \leqslant \frac{1}{m}\|d\|. \quad (2.5)$$

In particular, for all $x \in \mathrm{K}_1$, we have

$$m\|x - x^*\| \leqslant \|F(x)\| = \|F(x) - F(x^*)\| \leqslant M\|x - x^*\|, \quad (2.6)$$

where x^* stands for the unique solution of Equation (1.1) in K_1. Then we have

$$\|F(x)\| \leqslant M, x \in \mathrm{K}_1. \quad (2.7)$$

Next, we describe a way to generate a descent Quasi-Newton direction for θ and then propose a new descent BFGS method for solving Equation (1.1). Recall that in Newton's methods, the Newton direction is a solution of the Newton Equation (1.3). If $F'(x_k)$ is singular, the Equation (1.3) may have no solution. To solve Equation (1.3), we may need to solve the least squares problem

$$\min \frac{1}{2}\|F'(x_k)d + F(x_k)\|^2,$$

in order to obtain a direction d_k, which results in the so-called Gauss-Newton equation

$$F'(x_k)^2 d + F'(x_k)F(x_k) = 0. \quad (2.8)$$

Where, we have used Assumption 1 (A). It is noticed that, if $F'(x_k)$ is nonsingular, Equation (2.8) is equivalent to Equation (1.3). In Reference [12], a Gauss-Newton based method was proposed where the Quasi-Newton direction is the solution of the following system of linear equation.

$$B_k d_k + \bar{q}_k = 0, \quad (2.9)$$

where $\overline{q_k}$ is an approximation of vector $F'(x_k)F(x_k)$ and B_k is an approximation of matrix $F'(x_k)^2$. Specifically, let λ_{k-1} be the steplength used in the previous iteration.

Therefore, we can define vector $\overline{q_k}$

$$\overline{q_k} = (F(x_k + \lambda_{k-1}F(x_k)) - F(x_k))/\lambda_{k-1} \approx F'(x_k)F(x_k),$$

and matrix B_k is updated by the BFGS formula in Reference [12]:

$$B_{k+1} = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k y_k^T}{y_k^T s_k}, \qquad (2.10)$$

where $y_k = F(x_k + \delta_k) - F(x_k)$, $s_k = x_{k+1} - s_k$ and $\delta_k = F_{k+1} - F_k$. It is clear that if $\|s_k\|$ is small, then

$$B_{k+1}s_k = y_k \approx F'(x_k)^2 s_k.$$

Now we replace y_k by y_k^*. In Reference [12]

$$y_k = F(x_k + \delta_k) - F(x_k) \approx F'(x_k)\delta_k \approx F'(x_k)F'(x_k)s_k.$$

We define $y_k^* = \frac{\delta_k^{*T} s_k}{\|s_k\|^2}\delta_k^*$, where $\delta_k^* = \delta_k + A_k s_k$,

$A_k = \gamma_k \frac{|g_{k+1}^T(g_{k+1} - g_k)|}{\|s_k\|}$, γ_k is a positive constant.

In the following we will discuss $y_k^* = \frac{\delta_k^{*T} s_k}{\|s_k\|^2}\delta_k^*$.

We know

$$\|y_k^* - F'(x_{k+1})^2 s_k\| = \left\|\frac{\delta_k^{*T} s_k}{\|s_k\|^2}\delta_k^* - F'(x_{k+1})^2 s_k\right\| \leqslant M'\|s_k\| + M''\|s_k\|,$$

where $M' = \frac{\|\delta_k\|^2}{\|s_k\|^2} + \|F'(x_{k+1})^2\|$ and $M'' = 2|MA_k| + |A_k^2|$, if $\|s_k\|$ is sufficiently small, and by using Assumption 1 (C) and (D), and Formulae (2.4) and (2.7), we obtain $MM \geqslant \frac{\|\delta_k\|^2}{\|s_k\|^2} \geqslant mm$, $|A_k| \leqslant \gamma_k MM$, by using $\|F'(x_{k+1})\| \leqslant M_1$. Then, we get $\|y_k^* - F'(x_{k+1})^2 s_k\| \to 0$.

Therefore, we can get

$$y_k^* \approx F'(x_{k+1})^2 s_k. \qquad (2.11)$$

Replacing y_k by y_k^*, we can get the approximate Guass-Newton-Based BFGS update formula

$$B_{k+1} = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k^* y_k^{*T}}{y_k^{*T} s_k}, \qquad (2.12)$$

where $s_k = x_{k+1} - s_k$, $y_k^* = \frac{\delta_k^{*T} s_k}{\|s_k\|^2}\delta_k^*$, where $\delta_k^* = \delta_k + A_k s_k$, $A_k = \gamma_k \frac{|g_{k+1}^T(g_{k+1} - g_k)|}{\|s_k\|}$, γ_k is a positive constant.

If B_{k+1} is updated by the formula (2.12), we can get

$$s_k^T B_{k+1} s_k = y_k^{*T} s_k = \frac{\delta_k^{*T} s_k}{\|s_k\|^2}\delta_k^{*T} s_k = \frac{(\delta_k^{*T} s_k)^2}{\|s_k\|^2} > 0, \qquad (2.13)$$

therefore, B_{k+1} is positive symmetric matrix whatever line search is used. Since the solution d_k of Formula (2.9) may not be a descent direction of θ at x_k when x_k is far away from a solution of Equation (1.1), it is not possible to get a steplength $\lambda_k \geqslant 0$ satisfying Formula (1.6). Instead, we use the following line search, the $\lambda_k > 0$ satisfies the equation.

$$\theta(x_k + \lambda_k d_k) - \theta(x_k) \leqslant -\sigma_1\|\lambda_k d_k\|^2 - \sigma_2\|\lambda_k F(x_k)\|^2 + \varepsilon_k\|F(x_k)\|^2, \qquad (2.14)$$

where σ_1 and σ_2 are positive constants, and ε_k satisfies

$$\sum_{k=0}^{\infty} \varepsilon_k < \infty. \qquad (2.15)$$

Because ε_k is small, we can get $\{x_k\}$ is approximately norm descent. Observe that

$$\lim_{\lambda_{k-1} \to 0} \overline{q_k} = F'(x_k)F(x_k) \triangleq q_k.$$

Accordingly, the solution of Eqaution (2.9) with $\overline{q_k}$ instead of $\overline{q_k}$ is $\tilde{d}_k = -B_k^{-1}F'(x_k)F(x_k)$. Using Assumption 1 (B), then we can get $\tilde{q}_k$ is a descent direction of θ at x_k. This observation prompts us to regard λ_{k-1} as a parameter. When this parameter is adjusted to be small enough, the solution of Eqaution (2.9) is a descent direction of θ at x_k. The following process gives details of its realization. Let

$$q_k(\lambda) = (F(x_k + \lambda F(x_k)) - F(x_k))/\lambda. \qquad (2.16)$$

Consider the system of linear equation with parameter λ

$$B_k d_k + q_k(\lambda) = 0. \qquad (2.17)$$

Let $d(\lambda)$ be the solution of Equation (2.17). The following lemma shows that when $\lambda > 0$ is sufficiently small, every solution of Equation (2.17) is a descent direction of θ at x_k.

Lemma 2.1 Let σ_1 and σ_2 be positive constants and B_k be a symmetric and positive definite matrix. If x_k is not a stationary point of Formula (1.5), then there exists a constant $\bar{\lambda} > 0$ depending on k such that when $\lambda \in (0, \bar{\lambda})$, the unique solution $d(\lambda)$ of Equation (2.17) satisfies

$$\nabla\theta(x_k)^T d(\lambda) < 0. \qquad (2.18)$$

Moreover, the inequality

$$\theta(x_k + \lambda d(\lambda)) - \theta(x_k) \leqslant -\sigma_1\|\lambda d(\lambda)\|^2 -$$

$\sigma_2\|\alpha F(x_k)\|^2$ (2.19)

holds for all $\alpha > 0$ sufficiently small.

Proof According to Equation (2.16), we have

$$\lim_{\alpha\to 0} q_k(\alpha) = F'(x_k)F(x_k).$$

Therefore, we obtain from Equation (2.17) that

$$\lim_{\alpha\to 0^+} \nabla\theta(x_k)^T d(\alpha) = -\lim_{\alpha\to 0^+} F(x_k)^T F'(x_k)B_k^{-1}q_k(\alpha) = -F(x_k)^T F'(x_k)B_k^{-1}F'(x_k)F(x_k).$$

Since $F'(x_k)$ is symmetric and $F'(x_k)F(x_k)\neq 0$, as x_k is not a stationary point of Formula (1.5), the last equality and positive definiteness of B_k imply Formula (2.18). Inequality (2.19) has been proved. Notice that

$$\lim_{\alpha\to 0^+} \frac{(\theta(x_k+\alpha d(\alpha)) - \theta(x_k))}{\alpha} = \lim_{\alpha\to 0^+} \nabla\theta(x_k)^T d(\alpha) = -F(x_k)^T F'(x_k)B_k^{-1}F'(x_k)F(x_k) < 0.$$

Since the right hand side of Inequality (2.19) is $o(\alpha)$. Then the Inequality (2.19) holds for all $\alpha > 0$ sufficiently small. The proof is complete.

From the above lemma, we can find a descent quasi-Newton direction by adjusting parameter α.

3 The statement of algorithms

In this section, we give an algorithm. Firstly, we give two procedures:

Procedure 1 Let constant $\rho\in(0,1)$ be given. Let i_k be the smallest nonnegative integer such that Inequality (2.19) holds with $\alpha = \rho^i, i = 0, 1, \cdots$. Let $d_k = d(\rho^{i_k})$ and $q_k = q_k(\rho^{i_k})$.

Procedure 1 ensures that the value of θ at $x_k + \rho^{i_k} d_k$ is less than that of θ at x_k, though d_k may not necessarily be a descent direction of θ at x_k. It is reasonable to let the scalar ρ^{i_k} be the steplength. However, this steplength may be very small if i_k is large. To enlarge steplength, we exploit the following forward procedure.

Procedure 2 Let i_k and d_k be determined by procedure 1. If $i_k = 0$, let $\alpha_k = 1$. Otherwise, let j_k be the largest positive integer $j\in\{0, 1, 2, \cdots, i_k - 1\}$ satisfying

$$\theta(x_k + \rho^{i_k-j}d_k) - \theta(x_k) \leqslant -\sigma_1\|\rho^{i_k-j}d_k\|^2 - \sigma_2\|\rho^{i_k-j}F(x_k)\|^2, \quad (3.1)$$

let $\alpha_k = \rho^{i_k-j_k}$.

Based on the above process, we propose a norm descent Gauss-Newton based BFGS method as follows.

Algorithm 1

Step 0 Choose an initial symmetric positive definite matrix $B_0\in R^{n\times n}$. Let $x_0\in R^n$. Let $k:=0$;

Step 1 Stop if $g_k = 0$. Otherwise, determine d_k and λ_k by Procedures 1 and 2. Let the next iterative be $x_{k+1} = x_k + \alpha_k d_k$.

Step 2 Put $s_k = x_{k+1} - x_k = \alpha_k d_k$, $W_k = F(x_{k+1}) - F(x_k)$, and $y_k^* = \frac{W_k^T s_k}{\|s_k\|^2}W_k$. Update B_k to get B_{k+1} by Formula (2.12).

Step 3 Let $k:=k+1$. Go to step 1.

Now let's see the properties of Algorithm 1.

Property In Algorithm 1, If $i_k = 0$, let $\alpha_k = 1$. Otherwise, let j_k be the largest positive integer $j\in\{0, 1, 2, \cdots, i_{k-1}\}$ satisfying

$$\theta(x_k + \rho^{i_k-j}d_k)) - \theta(x_k) \leqslant -\sigma_1\|\rho^{i_k-j}d_k\|^2 - \sigma_2\|\rho^{i_k-j}F(x_k)\|^2, \quad (3.2)$$

where $\alpha_k = \rho^{i_k-j_k}$.

Note that Formula (3.2) is satisfied with $j = 0$ and B_k is symmetric positive definite matrix. Therefore, Algorithm 1 is well-defined.

d_k and α_k are generated by the algorithm. From the algorithm, it is easy to see that

$$\theta(x_k + \alpha_k d_k) - \theta(x_k) \leqslant -\sigma_1\|\alpha_k d_k\|^2 - \sigma_2\|\alpha_k F(x_k)\|^2, \quad (3.3)$$

which corresponds to Formula (2.14) with $\beta = 0$. It is also easy to see that if $\alpha_k \neq 1$, then $\alpha'_k = \alpha_k/\rho$ satisfies

$$\theta(x_k + \alpha'_k d_k) - \theta(x_k) > -\sigma_1\|\alpha'_k d_k\|^2 - \sigma_2\|\alpha'_k F(x_k)\|^2. \quad (3.4)$$

Notice that the algorithm generates a direction d_k which satisfies

$$B_k d_k + q_k = 0, \quad (3.5)$$

where $q_k = q_k(\rho^{i_k})$. Vector q_k differs from $q_k(\alpha_k)$ if $j_k \neq 0$.

4 Global convergence analysis

In this section, we will establish global and superlinear convergence for Algorithm 1. In a similar way we can get the global and superlinear convergence of Algorithm 2. Let $\{x_k\}$ and $\{B_k\}$ stand for the sequences of iterates and matrices generated by Algorithm 1 respectively. From Algorithm 1, we can obtain the following lemma straightforward.

According to Assumption 1(D), Formulae (2.6) and (2.7), it is not difficulty to deduce that

$$\|y_k^*\| \leqslant \frac{\|\gamma_k\|^2}{\|s_k\|} \leqslant (M + \mu_k MM)^2\|s_k\| \qquad (4.1)$$

and

$$A_k \leqslant \mu_k M\|F(x_{k+1}) - F(x^*)\| \leqslant \mu_k MM\|x_{k+1} - x^*\|, \qquad (4.2)$$

if k is large enough, we have $x_{k+1} \to x^*$, and by choosing suitable μ_k, we can get $A_k \to 0$.

Define

$$G_k = \int_0^1 F'(x_k + t\lambda_{k-1}F_k)\,dt. \qquad (4.3)$$

Then we have $\lambda_{k-1}^{-1}[F(x_k + \lambda_{k-1}F_k) - F_k] = G_kF_k$.

Hence Formula (2.17) can be rewritten as

$$B_kd_k + G_kF_k = 0. \qquad (4.4)$$

Lemma 4.1 Let Assumption 1 be satisfied. The sequence $\{\theta(x_k)\}$ is strictly decreasing. In addition, the following statements hold.

(1) If $s_k \to 0$, then there are constants $m_1 > 0$ and $M_2 > 0$ such that for all k sufficiently large

$$M_2\|s_k\|^2 \geqslant y_k^{*T}s_k \geqslant m_1\|s_k\|^2. \qquad (4.5)$$

(2) Suppose that $\lambda_k = 1$ holds only for a finite number of k. Then we have

$$\sum_{k=0}^{\infty}\|\lambda_kF_k\|^2 < \infty \qquad (4.6)$$

and

$$\sum_{k=0}^{\infty}\|\lambda_kd_k\|^2 = \sum_{k=0}^{\infty}\|s_k\|^2 < \infty. \qquad (4.7)$$

Moreover, Formula (4.5) holds for all k sufficiently large.

Proof (1) By Assumption 1 and the mean-value theorem, we have

$$\gamma_k^{*T}s_k = s_k^T(F_{k+1} - F_k) + A_ks_k = s_k^T\int_0^1 F'(x_k + ts_k)\,dt\,s_k + A_ks_k = s_k^T[\int_0^1 F'(x_k + ts_k)\,dt]^2s_k + s_k^T\int_0^1 F'(x_k + t\gamma_k)\,dt\,\gamma_k - s_k^T[\int_0^1 F'(x_k + ts_k)\,dt]^2s_k + s_k^T\int_0^1 F'(x_k + ts_k)\,dt\,s_k - s_k^T\int_0^1 F'(x_k + t\gamma_k)\,dt\,\gamma_k + A_ks_k = \|\int_0^1 F'(x_k + ts_k)\,dt\,s_k\|^2 + s_k^T\int_0^1 F'(x_k + t\gamma_k) - F'(x_k + ts_k))\,dt\int_0^1 F'(x_k + ts_k)\,dt\,s_k + s_k^T[\int_0^1 F'(x_k + ts_k)s_k - F'(x_k + t\gamma_k)\gamma_k)\,dt] + A_ks_k \geqslant \|F_{k+1} - F_k\| - \|s_k^T[\int_0^1 F'(x_k + t\gamma_k) - F'(x_k + t\gamma_k))\,dt]\|\|\int_0^1 F'(x_k + ts_k)\,dt\,s_k\| - \|s_k^T\int_0^1 F'(x_k + ts_k)s_k - F'(x_k + t\gamma_k)\gamma_k\,dt]\| + A_ks_k \geqslant m\|s_k\|^2 - \|s_k\|^2M\int_0^1 F'(x_k + t\gamma_k) - F'(x_k + ts_k)\|\,dt - \|s_k\|\int_0^1\|F'(x_k + ts_k)s_k - F'(x_k + t\gamma_k)\gamma_k\|\,dt + A_ks_k = [m - M\int_0^1\|F'(x_k + t\gamma_k) - F'(x_k + ts_k)\|\,dt]\|s_k\|^2 - [\int_0^1\|F'(x_k + ts_k)s_k - F'(x_k + t\gamma_k)\gamma_k\|\,dt]\|s_k\| + A_ks_k, \qquad (4.8)$$

where the last inequality follows Formulae (2.2) and (2.4). If $s_k \to 0$, then $\gamma_k = F_{k+1} - F_k \to 0$. By the continuity of $F'(x_k)$, we have $F'(x_k + ts_k) \to F'(x_k)$, $F'(x_k + t\gamma_k) \to F'(x_k)$ and using $A_k \to 0$. So we can obtain $\gamma_k^Ts_k \geqslant m'\|s_k\|^2$, $y_k^{*T}s_k = \frac{\gamma_k^Ts_k}{\|s_k\|^2}\gamma_k^Ts_k$, therefore, we get the left hand side of Formula (4.5). By using Formulae (2.2) and (4.1), we have

$$y_k^{*T}s_k \leqslant \|\gamma_k\|^2 \leqslant (M + \mu_kMM)^2\|s_k\|^2.$$

Then we get the right hand side of Formula (4.5). Therefore, we get Formula (4.5).

(2) If $\lambda_k = 1$ holds for only finitely many k's, then Formula (3.1) is used to determine a steplength λ_k for all k sufficiently large. By Formula (2.14), we have

$$\sigma_1\|\lambda_kF_k\|^2 + \sigma_2\|s_k\|^2 \leqslant \|F_k\|^2 - \|F_{k+1}\|^2 + \varepsilon_k\|g_k\|^2.$$

Since $\{\|F_k\|\}$ is bounded and $\{\varepsilon_k\}$ satisfies Formula (2.15), we get Formulae (4.6) and (4.7) by summing these inequalities. In particular, $\|s_k\| \to 0$, which also implies that Formula (4.5) holds for all k sufficiently large. The proof is complete.

We are going to establish a global convergence theorem of Algorithm 1 to show that under Assumption 1, there exists an accumlation point of $\{x_k\}$ which is a stationary point of Formula (1.5), namely,

$$\liminf_{k\to\infty}\|\nabla\theta(x_k)\| = 0. \qquad (4.9)$$

It is obtained from Lemma 4.1 that if $\lim_{k\to\infty}\sup\lambda_k \neq 0$, then $\lim_{k\to\infty}\|F(x_k)\| = 0$, hence, Formula (4.9) holds. So, we need only to show Formula (4.9) for the case $\lim_{k\to\infty}\lambda_k = 0$.

We do it by assuming

$$\liminf_{k\to\infty}\|\nabla\theta(x_k)\| > 0 \qquad (4.10)$$

to deduce a contradiction.

Notice that Formula (4.10) particularly implies that there is a constant $\varepsilon > 0$ such that $\|F(x_k)\| \geqslant \varepsilon$ for all k.

Therefore, from Formulae (4.1), (4.5) and Theorem 2.1 in Reference [2] we get the following lemma

Lemma 4.2 If Formula (4.10) holds, then there are positive constants U_i, $i = 1, 2, 3$, such that for any positive integer k inequalities

$$\|B_i s_i\| \leqslant U_1 \|s_i\|, U_2 \|s_i\|^2 \leqslant s_k^T B_i s_i \leqslant U_3 \|s_i\|^2 \quad (4.11)$$

hold for at least $[k/2]$ many $i \leqslant k$.

Inequalities (4.11) and (3.5) imply that there are at least $[k/2]$ many $i \geqslant k$ satisfying

$$\|q_i\| = \|B_i d_i\| \leqslant U_1 \|d_i\|, \|d_i\| \leqslant U_2^{-1} \|q_i\|. \quad (4.12)$$

Next, we prove the global convergence of Algorithm 1.

Theorem 4.1 Let Assumption 1 be satisfied, and x_k be generated by Algorithm 1. Then Formula (4.9) holds.

Proof We need only to show Formula (4.9) for the case $\lim_{k\to\infty} T_k = 0$. In this case, Inequality (3.4) holds for all sufficiently large k. Suppose contrarily that Formula (4.9) is not held or equivalent, Formula (4.10) holds. Denote by K the set of indices i such that Formula (4.11) holds. Then K is infinite. Since $\{x_k\} \subset K$ is bounded. Let $K_1 \subset K$ and subsequences $\{x_k\}_{k\in K_1}$ and $\{d_k\}_{k\in K_1}$ converge to x^* and d^*, respectively. Then we have

$$\lim_{k\in K_1} q_k = 5\theta(x^*). \quad (4.13)$$

Dividing both sides of Inequality (3.4) by T_k and then taking limits as $k \to \infty$ with $k \in K_1$, we get

$$5\theta(x^*)^T d^* \geqslant 0. \quad (4.14)$$

On the other hand, taking inner product with d_k in Formula (3.5), we get

$$0 = d_k^T B_k d_k + q_k^T d_k \geqslant U_2 \|d_k\|^2 + q_k^T d_k.$$

Taking limits in both sides as $k \to \infty$ with $k \in K_1$ yields

$$5\theta(x^*) d^* \leqslant - U_2 \|d^*\|^2.$$

This together with Formula (4.14) implies $d^* = 0$. It then follows Formula (4.12) that $\lim_{k\in K_1} q_k = 0$, which together with Formula (4.13) yields a contradiction with Formula (4.10). The contradiction proves Formula (4.9).

Remarks In the modified BFGS method in this paper, the iterative matrix B_k is always positive definite, and the similar updating technique is also adopted as used in Reference [11]. Consequently, we establish Theorem (4.1) which shows that the iterative sequence has an accumulation point which is a stationary point of problem $\min \theta(x) = \frac{1}{2}\|F(x)\|^2$. It may not be a solution of the nonlinear equation (1.1) if the Jacobian is singular at that point. The following theorem shows a strong convergence property of Algorithm 1.

Theorem 4.2 Let Assumption 1 hold. Suppose that the sequence $\{x_k\}$ generated by Algorithm 1 has a subsequence converging to a stationary x^* in which $F'(x^*)$ is nonsingular. Then x^* is a solution of Equation (1.1). Moreover, the whole sequence $\{x_k\}$ converges to x_k^*.

Proof Since x^* satisfies $5\theta(x^*) = F'(x^*)F(x^*)$, we have $F(x^*) = 0$ if $F(x^*)$ is nonsingular. Since $\{\theta(x_k)\}$ is convergent, every accumulation point of $\{x_k\}$ is a solution of Equation (1.1). By the nonsingularity of $F'(x^*)$ again, x^* is an isolated limit point of $\{x_k\}$. From Formulae (4.6) and (4.7), we have $x_{k+1} - x_k \to 0$ as $k \to \infty$. Therefore, the whole sequence $\{x_k\}$ converges to x^*. Next, we give a choice of $\_k$ to satisfy Formula (4.2).

Let Y and m_3, m_4 $(m_3 < m_4)$ be positive constants and $U \in [0, 1]$, we define the Q:

$$Q = \{k \mid m_3 \|W_k\| \leqslant Y g_{k+1}^T (g_{k+1} - g_k) \mid \leqslant m_4 \|W_k\|\}$$

and choose $\_k$ as follows:

$$\_k = \begin{cases} Y, & \text{if } k \in Q, \\ (U m_3 + (1 - U) m_4) \dfrac{m_3 \|W_k\|}{| g_{k+1}^T (g_{k+1} - g_k)|}, & \text{if } k \notin Q. \end{cases}$$

It is very clear that Formula (4.2) holds.

In the following we will prove the superlinear convergence of Algorithm 1.

5 Superlinear convergence analysis

To obtain superlinear convergence of Algorithm 1, we need the following Assumption 2.

Assumption 2 $F'(x)$ is Hölder continuous at x^*; i.e., there are positive constants M_3 and e such that for every x in a neighborhood of x^*,

$$\|F'(x) - F'(x^*)\| \leqslant M_3 \|x - x^*\|^e. \quad (5.1)$$

The following lemma shows that, like the ordinary BFGS method, the Dennis–Moré condition in References [4, 5] ensures superlinear convergence of

Algorithm 1. Recall that B_k is updated so as to approximate $F_k'^2 = F_k'F_k'^T$ in Algorithm 1.

Lemma 5.1 Let Assumption 1 hold. If

$$\lim_{k\to 0}\frac{\|(B_k - F'(x^*)^2)p_k\|}{\|p_k\|} = 0, \qquad (5.2)$$

then $\lambda_k \equiv 1$ for all k sufficiently large. Moreover, $\{x_k\}$ converges superlinearly.

Proof Omitted. For the proof see Reference [12][Lemma 3.5].

This lemma shows that to establish superlinear convergence of Algorithm 1, it verifys that $\{x_k\}$ satisfies the Dennis-Moré condition (5.2).

Lemma 5.2 Let Assumption 1 hold. Then, for any fixed $\nu > 0$, we have

$$\sum_{k=0}^{\infty}\|x_k - x^*\|^{\nu} < \infty. \qquad (5.3)$$

Moreover, we have

$$\sum_{k=0}^{\infty}\varepsilon_k(\nu) < \infty, \qquad (5.4)$$

where $\varepsilon_k(\nu) = \max\{\|x_k - x^*\|^{\nu}, \|x_{k+1} - x^*\|^{\nu}\}$.

Proof Omitted. For the proof see Reference [12][Lemma 3.6].

Lemma 5.3 Let Assumptions 1 and 2 hold. Then, there exist positive constants M_4 and m_2 such that for all k sufficiently large

$$\|y_k^* - F'^2(x^*)s_k\| \leqslant M_4\varepsilon_k\|s_k\| \text{ and } \|y_k^*\| \geqslant m_2\|s_k\|, \qquad (5.5)$$

where $\varepsilon_k = \max\{\|x_k - x^*\|^{\nu}, \|x_{k+1} - x^*\|^{\nu}\}$.

Proof Since $x_k \to x^*$, Formula (5.1) holds for all k large enough.

$$\gamma_k^{*T}s_k\gamma_k^* = (\gamma_k + A_ks_k)^Ts_k(\gamma_k + A_ks_k) = (\gamma_k^Ts_k + A_ks_k^Ts_k)(\gamma_k + A_ks_k) = \gamma_k^Ts_k\gamma_k + A_k\gamma_k^Ts_ks_k + A_ks_k^Ts_k\gamma_k + A_k^2s_k^Ts_ks_k. \qquad (5.6)$$

By using Formula (4.2), we obtain

$$\|A_k\| \leqslant C\|x_{k+1} - x^*\| \leqslant C\varepsilon_k. \qquad (5.7)$$

Where $C = \_{}_kMM$.

By using the mean value theorem we have for all k sufficiently large

$$\begin{aligned}&\|\gamma_k^Ts_k\gamma_k - s_k^Ts_kF'^2(x^*)s_k\| = \\ &\|s_k^TF'(\xi_1)s_kF'(\xi_1)s_k - s_k^Ts_kF'(x^*)F'(x^*)s_k\| \leqslant \\ &\|s_k^TF'(\xi_1)s_kF'(\xi_1)s_k - s_k^TF'(\xi_1)s_kF'(x^*)s_k\| + \\ &\|s_k^TF'(\xi_1)s_kF'(x^*)s_k - s_k^Ts_kF'(x^*)F'(x^*)s_k\| \leqslant \\ &M\|s_k\|^3\|F'(\xi_1) - F'(x^*)\| + \\ &\|F'(x^*)s_ks_k^TF'(\xi_1)s_k - s_k^Ts_kF'(x^*)(s_ks_k^T)^{-1}(s_ks_k^T)\cdot \\ &F'(x^*)s_k\| \leqslant MM_3\|s_k\|^3\varepsilon_k + \|F'(x^*)\|\cdot \\ &\|s_ks_k^TF'(\xi_1)s_k - (s_ks_k^T)^{-1}s_k(s_k^Ts_k)s_k^TF'(x^*)s_k\| = \\ &MM_3\|s_k\|^3\varepsilon_k + \|F'(x^*)\|\|s_ks_k^TF'(\xi_1)s_k - \\ &(s_ks_k^T)^{-1}(s_k^Ts_k)s_ks_k^TF'(x^*)s_k\| \leqslant MM_3\varepsilon_k\|s_k\|^3 + \\ &M\|s_k\|\|s_k^T\|\|F'(\xi_1) - F'(x^*)\|\|s_k\| \leqslant \\ &2MM_3\varepsilon_k\|s_k\|^3. \end{aligned} \qquad (5.8)$$

Where $\xi_1 \in (x_k, x_{k+1})$. Using Formulae (5.6), (5.7) and (5.8), we can get

$$\begin{aligned}&\|y_k^* - F'(x^*)^2s_k\| = \\ &\left\|\frac{\gamma_k^{*T}s_k\gamma_k^* - s_k^Ts_kF'(x^*)^2s_k}{s_k^Ts_k}\right\| \leqslant \frac{\|\gamma_k^Ts_k\gamma_k - s_k^Ts_kF_*^2s_k\|}{\|s_k\|^2} + \\ &\frac{\|A_k\gamma_k^Ts_ks_k\|}{\|s_k\|^2} + \frac{\|A_ks_k^Ts_k\gamma_k\|}{\|s_k\|^2} + \frac{\|A_k^2s_k^Ts_ks_k\|}{\|s_k\|^2} \leqslant \\ &2MM_3\varepsilon_k\|s_k\| + CM\varepsilon_k\|s_k\| + CM\varepsilon_k\|s_k\| + C^2\varepsilon_k\|s_k\| = \\ &M_4\varepsilon_k\|s_k\|,\end{aligned} \qquad (5.9)$$

where $M_4 = 2MM_3 + 2CM + C^2$. Therefore, the first inequality of Formula (5.5) holds. Moreover, by Formulae (2.2) and (4.8) we have

$$\|y_k^*\| \geqslant m'\|\gamma_k^*\| \geqslant m'm\|s_k\|.$$

Therefore, the second inequality of Formula (5.5) holds. The proof is complete.

Remarks Denote $P = F'(x^*)^{-1}$. For an $n\times n$ matrix A, define a matrix norm $\|A\|_P = \|PAP\|_F$, where $\|\cdot\|_F$ denotes the Frobenius norm of a matrix. Here let H_k and H_{k+1} stand for the inverse matrices of B_k and B_{k+1}, respectively. The following lemma shows that the BFGS Formula (2.12) exhibits a property similar to that of the ordinary BFGS formula.

Lemma 5.4 Under Assumption 1 (A) and (B), there exist positive constants M_5, M_6, M_7, and $\kappa \in (0,1)$ such that for all k sufficiently large

$$\|B_{k+1} - F'(x^*)^2\|_P \leqslant \|B_k - F'(x^*)^2\|_P + M_5\varepsilon_k, \qquad (5.10)$$

$$\|H_{k+1} - F'(x^*)^{-2}\|_P^{-1} \leqslant \left(1 - \frac{1}{2}\kappa K_k^2 + M_6\varepsilon_k\right)\|H_k - F'(x^*)^{-2}\|_P^{-1} + M_7\varepsilon_k, \qquad (5.11)$$

where $\varepsilon_k(\nu) = \max\{\|x_k - x^*\|^{\nu}, \|x_{k+1} - x^*\|^{\nu}\}$ and K_k is given by

$$K_k = \frac{\|p^{-1}[H_k - F'^{-2}(x^*)]y_k^*\|}{\|H_k - F'^{-1}(x^*)\|_P^{-1}\|py_k^*\|}. \qquad (5.12)$$

In particular, $\{\|B_k\|\}$ and $\{\|H_k\|\}$ are bounded.

Proof Omitted. For the proof see Reference [12][Lemma 3.8].

The following theorem shows that the new descent directions of the presented Qusi-Newton method also have the property of superlinear convergence.

Theorem 5.1 Let Assumption 1 and Assumption 2 hold. Then

$$\lim_{k\to\infty}\frac{\|(B_k - F'(x^*)^2 s_k\|}{\|s_k\|} = 0. \qquad (5.13)$$

Moreover, $\{x_k\}$ generated by Algorithm 1 converges superlinearly and $\lambda_k \equiv 1$ for all k sufficiently large.

Proof Formula (5.11) can be rewrititen as

$$\frac{1}{2}k\__k^2\|H_k - F'(x^*)^{-2}\|_{P^{-1}} \leqslant \|H_k - F'(x^*)^{-2}\|_{P^{-1}} - \|H_{k+1} - F'(x^*)^{-2}\|_{P^{-1}} + M_6\|H_k - F'(x^*)^{-2}\|_{P^{-1}}\, i_k + M_7\, i_k.$$

We know that $\{\|H_k - F'(x^*)^{-2}\|\}$ is bounded and $i_k = i_k(V)$ satisfies Formula (5.4). Therefore, summing the above inequalities, we obtain

$$\frac{1}{2}k\sum_{k=0}^{\infty}\__k^2\|H_k - F'(x^*)^{-2}\|_{P^{-1}} < \infty. \qquad (5.14)$$

According to Formula (5.12), we have

$$\lim_{k\to\infty}\__k^2\|H_k - F'(x^*)^{-2}\|_{P^{-1}}\| = \lim_{k\to\infty}\frac{\|P^{-1}(H_k - F'(x^*)^{-2})y_k^*\|^2}{\|H_k - F'(x^*)^{-2}\|_{P^{-1}}\|Py_k^*\|^2} = 0.$$

Since $\|H_k - F'(x^*)^{-2}\|_{P^{-1}}$ is bounded, it follows

$$\lim_{k\to\infty}\frac{\|P^{-1}(H_k - F'(x^*)^{-2})y_k^*\|}{\|Py_k^*\|} = 0. \qquad (5.15)$$

According to Formulae (2.5) and (4.1), we have

$$\|Py_k^*\| = \|F'(x^*)^{-1}y_k^*\| \leqslant \frac{1}{m}\|y_k^*\| \leqslant \frac{M^2}{m}\|s_k\|.$$

By Formula (2.4), we get

$$\|P^{-1}(H_k - F'(x^*)^{-2})y_k^*\| = \|F'(x^*)(H_k - F'(x^*)^{-2})y_k^*\| \geqslant m\|(H_k - F'(x^*)^{-2})_k^*\|.$$

Therefore, Formula (5.15) implies

$$\lim_{k\to\infty}\frac{\|(H_k - F'(x^*)^{-2})y_k^*\|}{\|s_k\|} = 0. \qquad (5.16)$$

On the other hand, we have

$$\|(H_k - F'(x^*)^{-2})y_k^*\| = \|H_k(F'(x^*)^2 - B_k)F'(x^*)^{-2}y_k^*\| \geqslant \|H_k(F'(x^*)^2 - B_k)s_k\| - \|H_k(F'(x^*)^2 - B_k)(s_k - F'(x^*)^{-2}y_k^*)\| = \|H_k(F'(x^*)^2 - B_k)s_k\| - \|H_k(F'(x^*)^2 - B_k)F'(x^*)^{-2}(y_k^* - F'(x^*)^2 s_k)\| \geqslant \|H_k(F'(x^*)^2 - B_k)s_k\| - M_4\, i_k\|H_k(F'(x^*)^2 - B_k)F'(x^*)^{-2}\|\|s_k\| = \|H_k(F'(x^*)^2 - B_k)s_k\| - o(\|s_k\|),$$

where the last inequality follows from Formula (5.5). We know that $\{\|B_k\|\}$ and $\{\|H_k\|\}$ are bounded, and $\{H_k\}$ is uniformly nonsingular. Therefore, there is a constant $m_3 > 0$ such that $\|H_k(F'(x^*)^2 - B_k)s_k\| \geqslant m_3\|(F'(x^*)^2 - B_k)s_k\|$ for all k. So we have

$$\|(H_k - F'(x^*)^{-2})y_k^*\| \geqslant \|F'(x^*)^2 - B_k)s_k\| - o(\|s_k\|)$$

and hence Formula (5.16) yields Formula (5.13). In view of Lemma 5.1, the proof is complete.

6 Numerical results

In this section, we report results of some preliminary numerical experiments by the proposed method. We solve the following problem which is similar to the problem 1 in various sizes in Reference [12].

Problem 1 The discretized two-poit boundary value problem[6],

$$g(x) \triangleq Ax + \frac{1}{(n+1)^2}F(x) = 0,$$

when A is the $n\times n$ tridiagonal matrix given by

$$A = \begin{bmatrix} 8 & -1 & & & \\ -1 & 8 & -1 & & \\ & -1 & 8 & -1 & \\ & & \ddots & \ddots & \ddots \\ & & & \ddots & \ddots & -1 \\ & & & & -1 & 8 \end{bmatrix}$$

and $F(x) = (F_1(x), F_2(x), \cdots, F_n(x))^T$ with $F_i(x) = \cos x_i - 1, i = 1, 2, \cdots, n$. In Algorithm 1, we choose $\_{}_k$ as follows:

$$\_{}_k = \begin{cases} 1, & \text{if } k \in \Omega, \\ \dfrac{m_3\|g_{k+1} - g_k\|}{|g_{k+1}^T(g_{k+1} - g_k)|}, & \text{if } k \notin \Omega \end{cases}$$

where $m_3 = 10^{-13}$, $m_4 = 10^{-5}$, and

$$\Omega = \{k \mid m_3\|g_{k+1} - g_k\| \leqslant |g_{k+1}^T(g_{k+1} - g_k)| \leqslant m_4\|g_{k+1} - g_k\|\}.$$

In the experiments, the parameters of Algorithm 1 were set to be $d = 0.1$, $T_1 = 0.001$, $T_1 = T_2 = 10^{-5}$, and in the nonlinear equation problem the initial matrix B_0 was set to be $B_0 = A$. The iteration stopped when the condition $\|F(x)\| \leqslant 10^{-6}$ was satisfied. The columns of the tables have the following meanings.

Problem: the name of the test problem in MATLAB;

Dim: the dimension of the problem;

NI: the number of iterations;

NF: the number of function;

The numerical results indicate that the proposed method performs well for Problem 1. Moreover, the initial points have not much influence upon the number of iterations. It is observed that the method is useful when the dimension of the problem becomes larger. The test results for nonlinear equation problems as follows:

(1) Small-scale problems refer to Table 1;

(2) Large-scale problems refer to Table 2.

Table 1 Test results for Algorithm 1

$x0$	(10,⋯,10)	(100,⋯,100)	(1000,⋯,1000)	(−10,⋯,−10)	(−100,⋯,−100)	(−1000,⋯,−1000)
Dim	NI/NF	NI/NF	NI/NF	NI/NF	NI/NF	NI/NF
$n=9$	21/137	23/145	24/147	21/135	23/145	24/147
$n=50$	69/579	71/589	70/583	70/587	71/591	72/593
$n=99$	125/1093	126/1097	128/1109	126/1103	129/1115	127/1099
$n=100$	126/1109	127/1117	128/1113	128/1111	127/1105	128/1119
$n=200$	215/1943	222/1997	242/2151	226/2035	222/1999	234/2101
$x0$	(10,0,10,0,⋯)	(100,0,100,0,⋯)	(1000,0,1000,0,⋯)	(−10,0,−10,0,⋯)	(−100,0,−100,0,⋯)	(−1000,0,−1000,0,⋯)
Dim	NI/NF	NI/NF	NI/NF	NI/NF	NI/NF	NI/NF
$n=9$	21/139	21/139	24/153	20/135	23/149	22/143
$n=50$	68/577	71/587	71/587	69/579	69/579	71/587
$n=99$	126/1107	125/1099	127/1101	127/1109	127/1113	128/1121
$n=100$	126/1087	127/1097	125/1095	126/1089	126/1097	128/1101
$n=200$	220/1989	224/2025	233/2095	209/1905	225/2033	237/2121

Table 2 Test results for Algorithm 1

$x0$	(10,⋯,10)	(100,⋯,100)	(1000,⋯,1000)	(−10,⋯,−10)	(−100,⋯,−100)	(−1000,⋯,−1000)
Dim	NI/NF	NI/NF	NI/NF	NI/NF	NI/NF	NI/NF
$n=300$	220/1995	241/2191	266/2423	215/1951	246/2245	261/2377
$n=350$	234/2127	249/2267	276/2517	240/2185	255/2323	286/2611
$n=400$	226/2063	245/2241	277/2527	220/2011	250/2283	280/2561
$n=450$	242/2197	274/2487	304/2761	247/2243	268/2443	309/2807
$n=500$	245/2233	269/2449	297/2717	240/2191	268/2443	293/2681
$x0$	(10,0,10,0,⋯)	(100,0,100,0,⋯)	(1000,0,1000,0,⋯)	(−10,0,−10,0,⋯)	(−100,0,−100,0,⋯)	(−1000,0,−1000,0,⋯)
Dim	NI/NF	NI/NF	NI/NF	NI/NF	NI/NF	NI/NF
$n=300$	214/1961	243/2227	274/2505	214/1961	247/2261	273/2501
$n=350$	239/2185	264/2413	289/2643	241/2201	262/2395	289/2643
$n=400$	228/2089	255/2335	289/2645	228/2089	255/2335	290/2651
$n=450$	236/2159	269/2459	305/2783	236/2159	269/2459	297/2709
$n=500$	242/2209	268/2447	290/2647	242/2209	268/2447	304/2775

From the numerical results, we can propose that Algorithm 1 is useful for symmetric nonlinear equation problems, specially, for the large-scale problems. In summary, the presented numerical results reveal that the given method has potential advantages when applying to symmetric nonlinear equations whose function is not difficult to compute while the dimension is large.

References

1 Broyden C G, Dennis J E, Moré J J. On the local and supelinear convergence of quasi-Newton methods. J Inst Math Appl, 1973, (12): 223~246.

2 Byrd R, Nocedal J. A tool for the analysis of quasi-Newton methods with application to unconstrained minimization. SIAM Journal on Numerical Analysis, 1989, (26): 727~739.

3 Byrd R, Nocedal J, Yuan Y. Global convergence of a class of quasi-Newton methods on convex problems. SIAM Journal on Numerical Analysis, 1987, (24): 1171~1189.

4 Dennis J E, Moré J J. A characteization of superlinear convergence and its application to quasi-Newton methods. Math Comp, 1974, (28): 549~560.

5 Dennis J E Jr, Moré J J. Quasi-Newton methods, motivation and theory. SIAM Rev, 1977, (19): 46~89.

6 Yuan Y, Sun W. Theory and Methods of Optimization. Beijing: Science Press of China, 1999.

7 Davidon W C. Variable metric methods for minimization. Argonne National Labs Report Argonne IL, 1959.

8 Li D, Fukushima M. A modified BFGS method and its global convergence in nonconvex minimization. Journal of Computational and Applied Mathematics, 2001, (129): 15~35.

(下转第 105页 Continue on page 105)

模型(12)的求解问题归结为求方程组(13)满足条件

$k_t \geq 0, (t = 1, 2, \cdots, 12)$,

的解.解方程组(13)有

$X = (20, 18, 20, 15, 30, 10)^T$,

$K = (k_1, k_2, \cdots, k_{12})^T = (\frac{10}{7}, \frac{30}{7}, \frac{5}{7}, \frac{15}{7}, 1, 1, 0, 0, \cdots, 0)^T$.

求解得到该运输公司调度运输工具 A_1 和 A_2 的数量分别是 20辆和 18辆;运输工具 A_1 装载货物 D_1 和 D_2 的数量分别是 20(箱 /辆)和 15(箱 /辆);运输工具 A_2 装载货物 D_1 和 D_2 的数量分别是 30(箱 /辆)和 10(箱 /辆);最小费用为 1420元.说明优化后的模型不但获利最大,而且运输工具的承载能力得到充分利用.

参考文献

1 袁亚湘,孙文瑜.最优化理论与方法.北京:科学出版社,2001.

2 陈宝林.最优化理论与算法.北京:清华大学出版社,2000.

3 Taha H A. Operations Research-An Introduction.吴立煦译.上海:上海人民出版社,1985.

4 Kaufmann A. Integer and mixed programming. In: Theory and Applications. London Academic Press INC LTD, 1977.

5 Fletcher R. Practical Methods of Optimization. In Constrained Optimization. New York: Johon Wiley and Sons, 1983.

6 Powell M J D. Nonlinear Optimization. London: Academic Press, 1982.

7 姜启源.数学模型.北京:高等教育出版社.1993.

8 傅远德.线性规划和整数规划.成都:成都科技大学出版社,1989.

(责任编辑:黎贞崇)

(上接第 99页 Continue from page 99)

9 Powell M J D. A new algorithm for unconstrained optimation. In: Rosen J B, Mangasarian O L, Ritter K. eds. Nonlinear Programming, New York: Academic Press, 1970.

10 Wei Z, Qi L, Chen X. An SQP-typemethod and its application in stochastic programming. Journal of Optimization Theory and Applications, 2003, (116): 205 ~ 228.

11 Wei Z, Yu G, Yuan G, et al. . The Superlinear Convergence of a Modified BFGS-type Method for Unconstrained Optimization. To be appeared in Computational and Applications, 2002.

12 Li D, Fukushima M. A globally and superlinear convergent Gauss-Newton-based BFGS method for symmetric nonlinear equations. SIAAM J Number Anal, 1999, (37): 152~ 172.

13 Moré J J, Garbow B S, Hillstrome K E. Testing unconstrained optimization software. ACM Trans Math Software, 1981, (7): 17~ 41.

14 Li D, Fukushima M. On the global convergence of the BFGS method for nonconvex unconstrained optimiation problems. SIAM J Optim, 2001, (11): 1054~ 1064.

15 Li D, Fukushima M. Descent Directions of Quasi-Newton Methods for Symmetric Nonlinear Equations, 2002.

(责任编辑:蒋汉明 黎贞崇)