# A feasible DY conjugate gradient method for linear equality constraints

**Can LI**

College of Mathematics, Honghe University, Mengzi 661199, China

E-mail: Lican841014@163.com

**Abstract.** In this paper, we propose a feasible conjugate gradient method for solving linear equality constrained optimization problem. The method is an extension of the Dai-Yuan conjugate gradient method proposed by Dai and Yuan to linear equality constrained optimization problem. It can be applied to solve large linear equality constrained problem due to lower storage requirement. An attractive property of the method is that the generated direction is always feasible and descent direction. Under mild conditions, the global convergence of the proposed method with exact line search is established. Numerical experiments are also given which show the efficiency of the method.

## 1. Introduction
Consider the unconstrained optimization problem

$$\min f(x), \; x \in R^n,\tag{1}$$

where $f: R^n \to R$ is a continuously differentiable function. The conjugate gradient method for solving (1) is given by

$$x_{k+1} = x_k + \alpha_k d_k,\tag{2}$$

where $\alpha_k$ is the step length, and $d_k$ is search direction defined by

$$d_k = \begin{cases} -\nabla f(x_k) & k = 0, \\ -\nabla f(x_k) + \beta_k d_{k-1} & k \geq 1, \end{cases}\tag{3}$$

where $\beta_k$ is a parameter. The conjugate gradient method is available for large-scale unconstrained optimization because its storage is relatively small. Numerical results [1] showed that, if $f$ is easy to be computed and if its dimension $n$ is vary large, the conjugate gradient method is still the best choice for solving unconstrained optimization problem (1).

If the objective function $f(x)$ is a strictly convex quadratic function, and if $\alpha_k$ is computed by the exact line search, namely,

$$f(x_k + \alpha_k d_k) = \min_{\alpha > 0} f(x_k + \alpha d_k).\tag{4}$$

then the method (2) and (3) is called the linear conjugate gradient method. The linear conjugate gradient method was originally propose by Hestenes and Stiefel [2] for solving the linear system of equations $Ax = b$ and several formulas of $\beta_k$ were considered, which are equivalent for strictly convex quadratic objective function.

On the other hand, (2) and (3) is called the nonlinear conjugate gradient method for general

unconstrained optimization problem. The nonlinear conjugate gradient method was first proposed by Fletcher and Reeves [3]. Some well-known conjugate gradient methods include the Hestenses-Stiefel method , Fletcher-Reeves method, Polak-Ribiere-Polyak method, Conjugate-Descent method, Liu-Storey method, and Dai-Yuan method . In this paper, we are interested in the DY method[4], in which $\beta_k$ is defined by

$$\beta_k^{DY} = \frac{\| \nabla f(x_k) \|^2}{d_{k-1}^{\mathrm{T}} y_{k-1}} \tag{5}$$

where $y_{k-1} = \nabla f(x_k) - \nabla f(x_{k-1})$, and $\| \cdot \|$ stands for the Euclidean norm. It was shown in [4] that such a method can guarantee the descent property of each direction provided the steplength satisfies the Wolfe conditions. In this case, the global convergence of the method was also proved in [4] under some mild assumptions on the objective function.

Recently, in view of the advantage of the conjugate gradient method for solving unconstrained optimization problem, Li and Li[5] extend it to the following linear equality constrained optimization problem

$$\min f(x)$$
$$\text{s.t. } Ax = b, \tag{6}$$

where $f : R^n \to R$ is continuously differentiable function, $A \in R^{m \times n}$ with $m < n$ is of full rank, and $b \in R^m$. The direction generated by the method is always feasible and descent direction. Under suitable conditions, the feasible Fletcher-Reeves method with exact line search is globally convergent.

In this paper, we further study the feasible conjugate gradient method for linear equality constrained optimization problem. We focus our attention to the DY conjugate gradient method. We will extend the DY method for solving unconstrained optimization problem to the linear equality constrained optimization problem.

## 2. Algorithm

In this section, we extend the DY method to linear equality constrained optimization problem (6). First we simply recall the feasible Fletcher-Reeves method in [5] for solving equality constrained optimization problem. For the problem (6), we separate the coefficient matrix A into $A = (B, N)$, where $B \in R^{m \times m}$ is nonsingular. For any $x \in D$, let $x = \begin{pmatrix} x^B \\ x^N \end{pmatrix}$, we get

$$x^B = B^{-1}b - B^{-1}Nx^N. \tag{7}$$

The problem (6) is equivalent to the following unconstrained optimization problem

$$\min F(x^N) \stackrel{\Delta}{=} f(x^B, x^N) = f(B^{-1}b - B^{-1}Nx^N, x^N).$$

In order to get feasible direction, the following lemma is given by Li and Li in [5].

**Lemma 1.** Let $x_k$ be the current iterate, $\bar{d}_k = \begin{pmatrix} \bar{d}_k^B \\ \bar{d}_k^N \end{pmatrix}$ is answer of the following linear programming

$$\min \quad \nabla f(x_k)^{\mathrm{T}} d$$
$$s.t \quad Ad = 0, \quad \| d^N \| \le 1.$$

Then the answer $\bar{d}_k = \begin{pmatrix} \bar{d}_k^B \\ \bar{d}_k^N \end{pmatrix}$ is given by

$$\bar{d}_k^N = -\frac{\nabla F(x_k^N)}{\| \nabla F(x_k^N) \|}, \quad \bar{d}_k^B = -B^{-1}N\bar{d}_k^N. \tag{8}$$

Let $z_k = \nabla f(x_k)^{\mathrm{T}} \overline{d}_k = \nabla F(x_k^N)^{\mathrm{T}} \overline{d}_k^N = -\|\nabla F(x_k^N)\|$ , and $-g_k = -z_k \overline{d}_k$ . Similar to the DY conjugate method in [4] for solving unconstrained optimization problem, we propose the following DY- type conjugate gradient method for solving the linear equality constrained optimization problem (8) which we call FDY method, that is, the direction $d_k$ is defined by

$$d_k = \begin{cases} -g_k & k = 0, \\ -g_k + \beta_k^{FDY} d_{k-1} & k \geq 1, \end{cases} \tag{9}$$

where $\beta_k^{FDY} = \dfrac{z_k^2}{d_{k-1}^{\mathrm{T}} y_{k-1}}$ , $y_{k-1} = \nabla f(x_k) - \nabla f(x_{k-1})$ . It is easy to see from (8) and (9) that for any

$k \geq 0$ , $Ad_k = 0$ .This implies $d_k$ that provides a feasible direction of $f$ at $x_k$ .

Based on the above discussion, we propose a feasible DY method which we call FDY method as follows.

**Algorithm 1(FDY method).**

Step1: Given constant $\varepsilon > 0$ , choose an initial point $x_0 \in D$ , let $k := 0$ .

Step2: If $|z_k| \leq \varepsilon$ , stop, get solution $x_k$ . Otherwise go to step 3.

Step3: Compute $d_k$ by (9).

Step4: Determine step size $\alpha_k$ by some line search.

Step5: Let $x_{k+1} = x_k + \alpha_k d_k$ and $k := k+1$ .Go to step2.


**3. Global convergence**

In this section, we prove the global convergence of Algorithm 1 with exact line search and with Wolfe line search respectively. To establish the global convergence theorem of Algorithm 1, we assume that the objective function satisfies the following condition that the level set $\Omega = \{x \in R^n \mid f(x) \leq f(x_0)\}$ is bounded.Next, we prove the global convergence of Algorithm 1 with exact line search (4).

**Theorem 2.** Let the sequence $\{x_k\}$ be generated by Algorithm 1 with exact line search，then

$$\liminf_{k \to \infty} \|\nabla F(x_k^N)\| = 0. \tag{10}$$

**Proof.** For the sake of contradiction，we suppose that the conclusion is not true. Then there exists a constant $\varepsilon > 0$ such that

$$\|\nabla F(x_k^N)\| \geq \varepsilon, \ \forall k. \tag{11}$$

From the definition of $d_k$ , we have

$$\nabla f(x_k)^{\mathrm{T}} d_k = -\nabla f(x_k)^{\mathrm{T}} g_k + \beta_k^{FDY} \nabla f(x_k)^{\mathrm{T}} d_{k-1}.$$

Since the steplength $\alpha_k$ is determined by exact line search, it holds that

$$\nabla f(x_k)^{\mathrm{T}} d_k = -\nabla f(x_k)^{\mathrm{T}} g_k.$$

Recalling the definition of $g_k$ , we obtain from (17) that

$$\nabla f(x_k)^{\mathrm{T}} d_k = -\|\nabla F(x_k^N)\|^2 < 0. \tag{12}$$

Since $\{f(x_k)\}$ is decreasing, it is clear that the sequence $\{x_k\}$ generated by Algorithm 1 is contained in $\Omega$ . This implies that $\{x_k\}$ is bound. That is, there is an infinite index set $K_1$ such that

$$f^* = \lim_{k(\in K_1) \to \infty} f(x_k) = f(\lim_{k(\in K_1) \to \infty} x_k) = f(x^*).$$

At the same time, the sequence $\{x_{k+1}| \, k \in K_1\}$ is bound. Similarly, there exists an infinite index set $K_2 \subset K_1$ such that

$$f^* = \lim_{k(\in K_2)\to\infty} f(x_{k+1}) = f(\lim_{k(\in K_2)\to\infty} x_{k+1}) = f(\overline{x}^*).$$

That is

$$f(x^*) = f(\overline{x}^*) = f^*. \tag{13}$$

Notice that the step length $\alpha_k$ is determined by exact line search, that is

$$f(x_{k+1}) = f(x_k + \alpha_k d_k) \le f(x_k + \alpha d_k), \quad \forall \alpha > 0.$$

Let $k \to \infty$, we have that for all $k \in K_2 \subset K_1$

$$f(\overline{x}^*) \le f(x^* + \alpha d^*), \quad \forall \alpha > 0. \tag{14}$$

By Taylor's expansion, there is

$$f(x^* + \alpha d^*) = f(x^*) + \alpha \nabla f(x^*)^{\mathrm{T}} d^* + o(\alpha). \tag{15}$$

It is follows from (15) and (12) that when $\alpha > 0$ is sufficient small $f(x^* + \alpha d^*) < f(x^*)$. With (14), we have $f(\overline{x}^*) \le f(x^* + \alpha d^*) < f(x^*)$. This contradicts (13). This contradiction shows that (10) is true.

## 4. Numerical experiment

In this section, we report some numerical experiments about Algorithm 1 with exact line search, and compare it with the feasible FR method proposed in [5]. We stop the iteration if the iteration number exceeds 5000 or the function evaluation number exceeds 7000 on the following inequality is satisfied $|z_k| \le 10^{-5}$.

In this paper, all codes are written in Matlab and run on PC with 2.66 GHZ CPU processor and 1GB RAM memory and Windows XP operation system. The results are listed in Table1. In Table1, Name denotes the name of the test problem, Dim denotes the dimension of the test problems, NF and NI denote the number of function evaluations and the iteration number respectively, Time denotes CPU time in second.

In table 1, we test the Algorithm 1 by eleven test problems. We can see from table 1 that the method terminates at the solution of all test problems. The results of numerical experiments indicate that the proposed method in this paper is efficiency method for some test problems.

Table 1   Numerical Results

| Name | Dim | FFR method | | | FDY method | | |
|---|---|---|---|---|---|---|---|
| | | NI | NF | Time | NI | NF | Time |
| MAD1 | 2 | 2 | 3 | 0.0160 | 1 | 2 | 0.0321 |
| MAD2 | 2 | 1 | 2 | 0.0150 | 4 | 5 | 0.0160 |
| MAD4 | 2 | 46 | 47 | 0.1410 | 13 | 14 | 0.0160 |
| Hs028 | 3 | 14 | 15 | 0.0160 | 58 | 59 | 0.0150 |
| Hs048 | 5 | 3 | 4 | 0.0160 | 3 | 4 | 0.0160 |
| Hs049 | 5 | 41 | 42 | 0.0150 | 37 | 38 | 0.0160 |
| Hs050 | 5 | 10 | 11 | 0.0150 | 11 | 12 | 0.0310 |
| Hs051 | 5 | 2 | 3 | 0.0160 | 2 | 3 | 0.0160 |
| Hs052 | 5 | 2 | 3 | 0.0160 | 2 | 3 | 0.0150 |
| Wong2 | 10 | 130 | 131 | 0.2030 | 106 | 107 | 0.0940 |
| Wong3 | 20 | 817 | 818 | 1.6400 | 1266 | 1267 | 1.6400 |

**Acknowledgment**

**References**
[1] Liu D C, Nocedal J. On the limited memory method for large scale optimization, Math. Prog., 1989,45:503-528.
[2] Hestenes M R, Stiefel E L. Method of conjugate gradient for solving linear systems. J.Res.Nat.Bur.Stand., 1952, 49:409-432.
[3] Fletcher R, Reeves, C. Function minimization by conjugate gradient method. J.Comput., 1964, 7:149-154.
[4] Dai Y H, Yuan Y. A nonlinear conjugate gradient method with a strong global convergence property. SIAM J.Optim., 1999, 10:177-182.
[5] Li C, Li D H. An extension of conjugate gradient method to linear equality constrained optimization problem, Applied Mathematics and Computation. 2013, 219: 10909-10914.