Efficient Generalized Conjugate Gradient Algorithms, Part 1: Theory

Y. LIU¹ AND C. STOREY²

Communicated by L. C. W. Dixon

Abstract. The effect of inexact line search on conjugacy is studied in unconstrained optimization. A generalized conjugate gradient method based on this effect is proposed and shown to have global convergence for a twice continuously differentiable function with a bounded level set.

Key Words. Unconstrained optimization, hybrid and restart conjugate gradient methods, inexact line search.

1. Introduction

Conjugate gradient methods are some of the most useful algorithms for unconstrained optimization of large problems by virtue of their storage saving properties. The general routine is given below.

Suppose that $f$ is a twice continuously differentiable function on its domain containing a bounded level set $L$,

$$L = \{x \in \mathbb{R}^n : f(x) \leq f(x_1)\}, \quad n \geq 2,$$

where $x_1$ is an initial point. Then, the minimum point of $f$ is to be found by a sequence of line searches on directions $s_k$, $k = 1, 2, \ldots$,

$$x_{k+1} = x_k + a_k s_k,$$

where

$$a_k = \arg \min_a f(x_k + as_k),$$

$$s_{k+1} = -g_{k+1} + b_k s_k,$$

¹ Visiting Scholar, Department of Mathematical Sciences, Loughborough University of Technology, Loughborough, Leicestershire, England.
² Professor, Department of Mathematical Sciences, Loughborough University of Technology, Loughborough, Leicestershire, England.
In (4), $g_k = f'(x_k)$ is the gradient of $f$ at $x_k$, $s_1 = -g_1$ is the steepest descent direction, and $b_k$ is determined by the conjugacy condition
\[ s_k^T H_{k+1} s_{k+1} = 0, \quad k = 1, 2, \ldots, \] (5)
where $H_{k+1} = f''(x_{k+1})$ is the Hessian of $f$ at $x_{k+1}$ and $s_k^T$ is the transpose of $s_k$. Thus, we could have $b_k$ of the form suggested by Daniel (Ref. 1),
\[ b_k = g_{k+1}^T H_{k+1} s_k / s_k^T H_{k+1} s_k. \] (6)

But in practice, in order to avoid computation of second derivatives and storage of matrices, condition (5) is changed to its difference form,
\[ s_{k+1}^T (g_{k+1} - g_k) = 0, \quad k = 1, 2, \ldots \] (7)
Accordingly, we have $b_k$ of the form suggested by Sorenson (Ref. 2),
\[ b_k = g_{k+1}^T (g_{k+1} - g_k) / s_k^T (g_{k+1} - g_k), \quad k = 1, 2, \ldots \] (8)
On the other hand, by (3) we have
\[ s_k^T g_{k+1} = 0, \quad k = 1, 2, \ldots; \] (9)
then, (8) becomes
\[ b_k = -g_{k+1}^T (g_{k+1} - g_k) / s_k^T g_k, \quad k = 1, 2, \ldots \] (10)
We note that $s_{k+1}$ in (4) is independent of the length of $s_k$ when $b_k$ takes the form of (6), (8), or (10). This property is useful in computations. On the other hand, if $b_k$ takes the Polak-Ribière form (Ref. 3)
\[ b_k = g_{k+1}^T (g_{k+1} - g_k) / g_k^T g_k, \quad k = 1, 2, \ldots \] (11)
which is obtained from (10) considering (4) and (9), this property is lost.
We know that, if $f$ is a quadratic function, then
\[ g_{k+1}^T g_k = 0, \quad k = 1, 2, \ldots, \]
and (11) takes the form suggested by Fletcher-Reeves (Ref. 4),
\[ b_k = g_{k+1}^T g_k / g_k^T g_k, \quad k = 1, 2, \ldots \] (12)
Fletcher-Reeves' method is the simplest of all conjugate gradient methods and its convergence is proved by Powell (Ref. 5), where $f$ is not restricted to being quadratic. Furthermore, Al-Baali (Ref. 6) extends this result to show the convergence of Fletcher-Reeves' method with inexact line search when $a_k$ satisfies the conditions
\[ |s_{k+1}^T s_k| < c_1 |s_k^T s_k|, \quad 0 < c_1 < 1/2, \] (13)
and
\[ f(x_{k+1}) \leq f(x_k) + c_2 a_k g_k^T s_k, \quad 0 < c_2 < 1/2. \] (14)