Journal of Computational Mathematics, Vol.23, No.3, 2005, 225–232.

PRECONDITIONED SPECTRAL PROJECTED GRADIENT METHOD ON CONVEX SETS *1)

Lenys Bello
(*Dpto. de Matemática, Facultad de Ciencias (FACYT), Universidad de Carabobo, Valencia, Venezuela*)

Marcos Raydan
(*Dpto. de Computación, Universidad Central de Venezuela, Ap. 47002, Caracas 1041-A, Venezuela*)

Abstract

The spectral gradient method has proved to be effective for solving large-scale unconstrained optimization problems. It has been recently extended and combined with the projected gradient method for solving optimization problems on convex sets. This combination includes the use of nonmonotone line search techniques to preserve the fast local convergence. In this work we further extend the spectral choice of steplength to accept preconditioned directions when a good preconditioner is available. We present an algorithm that combines the spectral projected gradient method with preconditioning strategies to increase the local speed of convergence while keeping the global properties. We discuss implementation details for solving large-scale problems.

Mathematics subject classification: 49M, 90C, 65K.
Key words: Spectral gradient method, Projected gradient method, Preconditioning techniques, Nonmonotone line search.

1. Introduction

We consider the optimization problem

$$\text{minimize } \{f(x):\ x \in \Omega\},$$

where Ω is a nonempty closed and convex set in $\Re^n$, n is large, f is continuously differentiable, $g(x) = \nabla f(x)$ is available, and $G(x) \approx \nabla^2 f(x)$ is also available and will be considered as a preconditioner. Our main objective is to develop a preconditioned and projected extension of the spectral gradient method to solve this problem.

Spectral gradient methods are nonmonotone schemes that have recently received considerable attention in the numerical analysis and optimization literature. They were introduced by Barzilai and Borwein [1], the convergence for quadratics was established by Raydan [17], and more recently, a proof of the R-linear rate of convergence for convex quadratics was given by Dai and Liao [10]. A complete review is presented by Fletcher [11], and the asymptotic behavior is studied by Dai and Fletcher [9].

The spectral gradient methods have been applied succesfully to find local minimizers of large scale problems ([4, 5, 7, 8, 12, 18, 20]), to solve box-constrained quadratic optimization (Bielschowsky et al. [2]), and also to minimize general smooth functions on convex sets [6]. Preconditioned spectral gradient versions have also been developed ([16, 13, 15]). However, the combination of preconditioning techniques to accelerate the process and projected techniques on convex sets, for robustness and regularity, has not been studied.

* Received December 10, 2003; final revised July 17, 2004.
1) This author was partially supported by UCV-PROJECT 97-003769.

The paper is divided into sections as follows. In Section 2 we describe briefly the preconditioned spectral gradient method, and we recall some of the important aspects of the spectral projected gradient method on convex sets. In Section 3 we present the new algorithm that combines both ideas taking into account the difficult cases. In Section 4 we present preliminary and encouraging numerical results, and some final remarks.

2. Previous Extensions

We describe briefly the most important properties of the preconditioned spectral gradient method and the spectral projected gradient method on convex sets.

2.1 Preconditioned spectral gradient method

The iterates of the Preconditioned Spectral Gradient (PSG) method presented by Glunt, Hayden, and Raydan [13] are defined by

$$x_{k+1} = x_k - \alpha_k^{-1} z_k \ ,$$

where $z_k = G_k^{-1} g_k$, G_k is a nonsingular approximation to the Hessian of f at x_k and the scalar α_k is given by

$$\alpha_k = (-\alpha_{k-1}) \frac{z_{k-1}^t y_{k-1}}{z_{k-1}^t g_{k-1}} \ ,$$

where x_0 and α_0 are given initial data (see also [16]).

The PSG method requires no line search during the process but does not guarantee monotonic descent in the objective function. As a consequence, Raydan [18] proposed a globalization scheme for the spectral gradient algorithm that fits nicely with the nonmonotone behavior of this family of methods. Roughly speaking, the algorithm forces the step to satisfy this weak condition:

$$f(x_{k+1}) \leq \max_{0 \leq j \leq M} f(x_{k-j}) + \gamma g_k^t (x_{k+1} - x_k) \ ,$$

where M is a nonnegative integer and γ is a small positive number. When $M > 0$ this condition allows the objective function to increase at some iterations and still guarantees global convergence. This globalization strategy is based on the nonmonotone line search technique of Grippo, Lampariello and Lucidi [14].

A direct combination of the PSG method and the nonmonotone globalization strategy described above produces an algorithm fully described in Luengo et. al. [15].

2.2 Spectral projected gradient method

There have been many different variations of the projected gradient method that can be viewed as the constrained extensions of the optimal gradient method for unconstrained minimization. They all have the common property of maintaining feasibility of the iterates by frequently projecting trial steps on the feasible convex set. In particular, Birgin et al. [6, 3] combine the projected gradient method with recently developed ingredients in optimization, as follows. The algorithm starts with $x_0 \in \Re^n$ and is based on the spectral projected gradient direction $d_k = P(x_k - \alpha_k g(x_k)) - x_k$, where α_k is the spectral choice of steplength $\frac{\langle s_{k-1}, s_{k-1} \rangle}{\langle s_{k-1}, y_{k-1} \rangle}$, and for $z \in \Re^n$, $P(z)$ is the projection on Ω. In the case of rejection of the first trial point, $x_k + d_k$, the next ones are computed along the same direction, i.e., $x_+ = x_k + \lambda d_k$, using a nonmonotone line search to force the following condition

$$f(x_+) \leq \max_{0 \leq j \leq \ \min \ \{k, M-1\}} f(x_{k-j}) + \gamma \lambda \langle d_k, g(x_k) \rangle,$$

where $M \geq 1$ is a given integer. As a consequence, the projection operation must be performed only once per iteration. More details can be found in [6] and [3].