Journal of Computational Mathematics
Vol.36, No.6, 2018, 881–902.
http://www.global-sci.org/jcm
doi:10.4208/jcm.1706-m2016-0721

A GENERAL TWO-LEVEL SUBSPACE METHOD FOR NONLINEAR OPTIMIZATION*

Cheng Chen
University of Chinese Academy of Sciences, Beijing 100190, China
LSEC, ICMSEC, Academy of Mathematics and Systems Science, Chinese Academy of Sciences Beijing, China
Email: cchen@lsec.cc.ac.cn
Zaiwen Wen
Beijing International Center for Mathematical Research, Peking University, Beijing 100871, China
Email: wenzw@pku.edu.cn
Yaxiang Yuan
LSEC, ICMSEC, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China
Email: yyx@lsec.cc.ac.cn

Abstract

A new two-level subspace method is proposed for solving the general unconstrained minimization formulations discretized from infinite-dimensional optimization problems. At each iteration, the algorithm executes either a direct step on the current level or a coarse subspace correction step. In the coarse subspace correction step, we augment the traditional coarse grid space by a two-dimensional subspace spanned by the coordinate direction and the gradient direction at the current point. Global convergence is proved and convergence rate is studied under some mild conditions on the discretized functions. Preliminary numerical experiments on a few variational problems show that our two-level subspace method is promising.

Mathematics subject classification: 65N06, 65B99.
Key words: Nonlinear optimization, Convex and nonconvex problems, Subspace technique, Multigrid/multilevel method, Large-scale problems.

1. Introduction

Consider an infinite-dimensional minimization problem

$$\min_{u\in\mathcal{V}} \ \mathcal{F}(u), \tag{1.1}$$

where $\mathcal{F}$ is a mapping from $\mathcal{V}$ to $\mathbb{R}$ and $\mathcal{V}$ is the infinite-dimensional space where u lives in. Infinite-dimensional optimization problems are a major source of large-scale finite dimensional optimization problems, such as partial differential equations (PDEs) and optimal control problems governed by PDEs. Since it is very hard or almost impossible to obtain explicit solutions for these problems, they are usually solved numerically either by a "discretize-then-optimize" strategy or an "optimize-then-discretize" strategy. For these kind of problems, usually a very

* Received August 17, 2016 / Revised version received March 14, 2017 / Accepted June 28, 2017 / Published online August 7, 2018 /

fine discretization is needed to obtain a satisfactory discretization error, but the computational cost is much expensive. In this paper, we follow the second strategy and propose a new numerical scheme to solve them.

Quite a few numerical optimization methods for large-scale problems have been developed using a fundamental technique named subspace optimization directly or indirectly. It attracts more and more attention in recent years [1–3]. The conjugate gradient method arose originally in [4] to solve linear systems and were introduced in nonlinear minimization in [5]. It defines a new search direction by a given linear combination of the negative gradient direction and the previous search direction. Yuan and Stoer [6] viewed the conjugate gradient method from the subspace point of view, namely, to find a best trial direction, even an approximate minimum, in the 2-dimensional subspace spanned by the two conjugate directions. Another popular method in nonlinear programming is the limited-memory quasi-Newton method proposed by Shanno [7] and Nocedal [8]. It generates the quasi-Newton matrix by using some historical information. The block coordinate descent (BCD) and the alternating direction method of multipliers (ADMM) are de facto subspace techniques. More general subspace methods and latest developments are referred to [3, 9–11].

Although existing optimization methods can be applied to solve problem (1.1), they make little use of its underlying hierarchical structure. In contrast, multigrid/multilevel method is a more natural concept. It was originally proposed for solving linear elliptic partial differential equations with simple boundary value and proved to work well [12–15]. It takes advantage of different levels discretization of infinite-dimensional problems to execute the coarse grid corrections recursively with a combination of smoothing steps on fine grid. It not only reduces the computational cost but also accelerates the convergence rate. It is well-known that good performance of iterative methods may depend on a good initial guess. The mesh refinement, or full multigrid method [14, 15], uses the nested iteration idea to solve fine grid problems with an initial point interpolated from the solution of the next coarser grid. Multigrid methods were also extended to solve nonlinear PDE problems. One approach is called Newton-MG method [14–16], in which a linear expansion at the current iterate is used in outer iterations and multigrid methods were used for Jacobian systems in inner iterations. Another extension is full approximation scheme (FAS) [15–17], in which the multigrid methodology is directly applied to the original system of nonlinear equations and its corresponding system of nonlinear residual equations. It obtains a full approximation rather than an error correction term in coarse grid problems. A combination of Newton-MG and FAS was proposed by Yavneh and Dardyk [18]. The other extension is projection multilevel method [19–21], which regards a series of discretization spaces as projections from the infinite-dimensional space, and represents them with nodal or finite element. Taking projections onto various subspaces, it solves the problems by correcting the current iterate.

Multigrid method has also be applied to infinite-dimensional optimization problems, especially optimal control problems governed by PDEs [22–25]. It is used for solving the KKT systems derived from optimality conditions and inner loops of optimization scheme derived from original problems. An approach was proposed by Nash [26] and developed in [27–30] for solving the unconstrained convex infinite-dimensional optimization problems, in which a linear term is added in the discretized nonlinear problems at each level other than the finest one to enforce first order coherence in the neighborhood of current iterate between the neighboring levels of grid. This is a new reinterpretation of multigrid from an optimization point of view and uses it as outer iterative scheme [25]. Based on this scheme, Wen and Goldfarb [31] proposed a