Available online at www.sciencedirect.com
ScienceDirect
Mathematics and Computers in Simulation 140 (2017) 1–23
MATHEMATICS AND COMPUTERS IN SIMULATION
www.elsevier.com/locate/matcom

Original articles

A Hessian-free Newton–Raphson method for the configuration of physics systems featured by numerically asymmetric force field[☆]

Yu Liang[a,*], Zhenjun Shi[b], Peter W. Chung[c]

[a] *Department of Computer Science and Engineering, University of Tennessee at Chattanooga, Chattanooga, TN 37403, USA*
[b] *Department of Mathematics and Computer Sciences, Central State University, Wilberforce, OH 45384, USA*
[c] *Department of Mechanical Engineering, University of Maryland, College Park, MD 20742, USA*

Received 7 September 2012; received in revised form 7 July 2015; accepted 1 November 2016
Available online 10 February 2017

Abstract

Numerically asymmetric force field is examined in particle-oriented problems such as Quasicontinuum modeling and simulation. To configure the ground state of a large-scale physical system featured by numerically asymmetric force field, we propose a Hessian-free Newton–Raphson method where the Newton equation is solved using central-difference based BiCGstab algorithm (denoted as HFNR-BiCGstab-diff for simplicity). A detailed analytical and experimental investigation on the convergence performance of the HFNR-BiCGstab-diff algorithm is given in this paper. A pure HFNR-BiCGstab-diff algorithm may suffer from unreliable start-up, particularly in the case that the initial guess is far from the minimizer. As a remedy, a hybrid method that couples HFNR-BiCGstab-diff with preconditioned nonlinear conjugate gradient algorithm (PNCG) is developed to achieve optimal computational performance. The algorithms addressed in this paper have been implemented using POSIX-Thread C++. Their performance has been evaluated using three-dimensional Quasicontinuum simulation problems, which are featured by asymmetric force field and large dimensional sizes up to 122,808 degree-of-freedom, as benchmarks. The numerical experiment on IBM SP2 demonstrates that, compared to alternative unconstrained optimization methods such as preconditioned nonlinear conjugate gradient algorithm, the hybrid algorithm saves 20%–60% running time.

Keywords: Unconstrained optimization; Hessian-free Newton–Raphson method; BiCGstab; Central-difference; Nonlinear conjugate gradient

1. Introduction

As its major contributions, this paper examines the numerically asymmetric force field derived from particle-oriented modeling and simulation, and then develops efficient solvers corresponding to the numerical asymmetric force field.

☆ This work is jointly sponsored by the National Science Foundation with contract number 1647175, National Science Foundation with contract Number 1240734, National Science Foundation with contract Number 1111542, U.S. Army Research Laboratory with contract Number DAAD19-01-2-0014.
* Corresponding author.
E-mail address: yu-liang@utc.edu (Y. Liang).

http://dx.doi.org/10.1016/j.matcom.2016.11.011
0378-4754/

An unconstrained optimization [5,10,11,13,35]

$$\min_{\mathbf{u}\in\Re^N} \Pi(\mathbf{u}), \tag{1}$$

where $\Pi(\mathbf{u})$ is the potential function corresponding to a system configuration $\mathbf{u}$, is generally used to configure the ground state of physical systems. This paper aims to use an inexact Newton–Raphson algorithm [4,2] to solve the unconstrained optimization problem arising from physical systems that suffer from asymmetric force field.

As one of the most popular and efficient unconstrained optimization methods, the Newton–Raphson algorithm [5,10,12,34,35,42,43,6,17,19] for Eq. (1) can be formulated as

$$\mathbf{u}^{(k+1)} = \mathbf{u}^{(k)} + \alpha_{k+1}\mathbf{d}^{(k)}, \quad \text{(Line search)} \tag{2}$$

and

$$\mathbf{G}^{(k)}\mathbf{d}^{(k)} = -\mathbf{g}^{(k)}, \quad \text{(Newton equations)} \tag{3}$$

where k indicates the nonlinear iterative number, gradient $\mathbf{g}^{(k)}$ and Hessian $\mathbf{G}^{(k)}$ are the first- and second-order derivatives of $\Pi(\mathbf{u})$ at point $\mathbf{u}^{(k)}$, respectively, and α_{k+1} is obtained by backtracking line search algorithm [32,35].

For large-scale problems, an explicit formulation of Hessian $\mathbf{G}^{(k)}$ is extremely costly. Therefore, the Newton–Raphson method suffers from relatively high storage and computing complexity in solving Newton equations to obtain the search directions [55]. As remedies, two following techniques are employed in this paper:

- truncated strategies (or inexact method) [5,9,12,17,20,33,34,42,43,54,56], where only an approximate solution to Newton equations is obtained, and
- Hessian-free or Hessian-reduced techniques [10,13,17,20,33,34,39], where an explicit formulation about Hessian is circumvented.

In order to reduce the computational cost of Hessian-related operations involved in Newton–Raphson algorithm, iterative algorithms are emphatically investigated to solve the Newton equations (Eq. (3)). As the only Hessian-related operations in the implementation of iterative algorithms, the matrix–vector product was approximated by central difference

$$\mathbf{G}^{(k)}\mathbf{v} = \frac{\mathbf{g}(\mathbf{u}^{(k)} + \tau\mathbf{v}) - \mathbf{g}(\mathbf{u}^{(k)} - \tau\mathbf{v})}{2\tau} + \mathcal{O}(\tau^2\|\mathbf{v}\|^3), \tag{4}$$

where τ is called the finite-difference interval and it will be determined according to the curvature of Π. Eq. (4) is derived from multivariate Taylor series.

Mathematically, Hessian $\mathbf{G}^{(k)}$ is supposed to be symmetric [52]. Namely, $\frac{\partial^2\prod}{\partial\mathbf{u}_i\partial\mathbf{u}_j} \equiv \frac{\partial^2\prod}{\partial\mathbf{u}_j\partial\mathbf{u}_i}$ where i and j indicate different components of a physics system. This hypothesis is also consistent with Newton's Third Law, which states that "for every action, there is an equal and opposite reaction". However, due to the application of central-difference strategy and cluster summary rule for particle-oriented modeling, the numerical symmetry of Hessian is not guaranteed. This issue is observed in our numerical experiments [20], where those symmetric linear solvers such as Conjugate Gradient [40] converge extremely slowly in solving the Newton equations. As one of the most efficient nonsymmetric linear solvers, the BiCGstab algorithm [40] based on the central difference technique (denoted as BiCGstab-diff for simplicity) is used to solve the Newton equations. The Newton–Raphson algorithm based on BiCGstab-diff (HFNR-BiCGstab-diff) is addressed in detail in terms of feasibility, accuracy, and convergence rate.

From the point of view of iterative convergence rate, a Newton–Raphson algorithm is superior to some other algorithms such as steep-descent [37] and nonlinear conjugate gradient methods [1,36,6,7] due to its quadratic convergence rate under some mild conditions. However, from the point of view of convergence time, Newton–Raphson algorithm is not always the winner, particularly during the early stage of calculation. As has been demonstrated in the numerical experiments, only in case that an initial guess ($\mathbf{u}^{(0)}$) is close enough to the minimizer ($\mathbf{u}^*$), the advantage of Newton–Raphson algorithm [5,10,12,34,35,42,43] can be fully embodied. As a result, this work investigates a hybrid strategy which combines the preconditioned nonlinear conjugate gradient (PNCG) [1,6,7,15,36,37] with the HFNR-BiCGstab-diff.

The nomenclature related to this paper is given in Table 1.