CSIAM Trans. Appl. Math.
doi: 10.4208/csiam-am.2020-0008

Vol. **1**, No. 1, pp. 142-154
March 2020

An Implicit Evaluation Method of Vector 2-Norms Arising from Sphere Constrained Quadratic Optimizations

T. Sogabe[1,*], A. Suzuki[2] and S.-L. Zhang[1]

[1] *Department of Applied Physics, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8603, Japan.*
[2] *Department of Computational Science and Engineering, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8603, Japan.*

Received 13 February 2020; Accepted 20 March 2020

Abstract. An implicit evaluation method of vector 2-norms is presented for function evaluations arising from sphere constrained quadratic optimizations. The efficiency of the method in terms of computational costs mainly comes from the well-known shifted conjugate gradient method, and the robustness of the method comes from the fact that it never suffers from cancellations when the coefficient matrix is symmetric positive definite. Numerical experiments indicates that the method is promising for reducing computational costs of Ye's hybrid method for solving sphere constrained quadratic optimizations.

AMS subject classifications: 65F10, 65K05, 90C20
Key words: Shifted linear systems, symmetric positive definite, the shifted conjugate gradient method, sphere constrained quadratic optimizations, Ye's hybrid method.

1 Introduction

Given a symmetric positive definite matrix $A\in\mathbb{R}^{N\times N}$ and a vector $b\in\mathbb{R}^N$, finding a root of the function

$$f(\sigma)=\|(A+\sigma I)^{-1}b\|_2-1 \tag{1.1}$$

over the set of all positive real numbers $\mathbb{R}^+$ is of prime importance in solving sphere constrained quadratic optimizations (SQO hereafter)

$$\begin{aligned}&\text{minimize} \quad q(x)=\frac{1}{2}(x,Ax)-(b,x)\\ &\text{subject to} \quad x\in S=\{x\in\mathbb{R}^N:\|x\|_2\le 1\}.\end{aligned}$$

*Corresponding author. *Email addresses:* sogabe@na.nuap.nagoya-u.ac.jp (T. Sogabe), suzuki@na.cse.nagoya-u.ac.jp (A. Suzuki), zhang@na.nuap.nagoya-u.ac.jp (S.-L. Zhang)

Here, the symbol $(\cdot,\cdot)$ denotes the standard dot product, i.e., $(x,y)=x^{\mathrm{T}}y$. The SQO problems have a rich variety of applications such as trust region algorithms for nonlinear programming (see [1] and references therein), regularization methods for ill-posed problems [8], and graph partitioning problems [6].

Of many root-finding algorithms, e.g. the bisection method, the false position method, and Newton's method, one of the most well-known root-finding algorithms for (1.1) is Ye's hybrid method [10] that is combining Newton's method and a special binary search. The most time consuming part of Ye's hybrid method is evaluating the function (1.1) for some given values, i.e.,

$$f(\sigma_\ell)=\|(A+\sigma_\ell I)^{-1}b\|_2-1, \quad \ell=1,2,\cdots,m. \tag{1.2}$$

For computing the 2-norms in (1.2), one may choose sparse direct solvers (see, e.g., [3]), since the sparse pattern of $A+\sigma_\ell I$ does not change for all $\sigma_\ell\in\mathbb{R}^+$ and thus one symbolic factorization may be enough. This is very cost-efficient in this respect. The approach, however, requires m times numerical factorizations that will be highly expensive. The Householder triangularization of $A+\sigma_\ell I$ may be more useful when the matrices are small, since the triangularizations can be done simultaneously, i.e., $Q^{\mathrm{T}}(A+\sigma_\ell I)Q=T+\sigma_\ell I$ for $\ell=1,\cdots,m$, where T is a tridiagonal matrix, and the Cholesky factorizations of tridiagonal matrices $T+\sigma_\ell I$ for $\ell=1,\cdots,m$ can be done in $\mathcal{O}(mN)$. When matrices are large, the memory requirement of $\mathcal{O}(N^2)$ may be a bottleneck.

The 2-norms in (1.2) can be viewed as the solution 2-norms $\|x^{(\ell)}\|_2$ of the following, so-called, shifted linear systems:

$$(A+\sigma_\ell I)x^{(\ell)}=b, \quad \ell=1,2,\cdots,m. \tag{1.3}$$

Krylov subspace methods for solving shifted linear systems have received much attention since only one Krylov subspace $K_n(A,b):=\mathrm{span}\{b,Ab,\cdots,A^{n-1}b\}$ is required because of the shift-invariance property $K_n(A,b)=K_n(A+\sigma_\ell I)$ (see, an excellent survey, [9]). Since the coefficient matrices are symmetric positive definite, it is natural to use the shifted Conjugate Gradient (shifted CG) method for solving (1.3) whose algorithm in a complete form is described in [7]; the symmetric version of the shifted Bi-CG [5], and an accurate variant of the shifted CG method was proposed for the case $(A^{\mathrm{T}}A+\sigma_\ell I)x^{(\ell)}=A^{\mathrm{T}}b$ [4]. This motivates us to use the shifted CG method for efficiently evaluating the function (1.1).

The purpose of the paper is to give an efficient evaluation method for computing the solution 2-norms $\|x^{(\ell)}\|_2$ of (1.3) based on the shifted CG method. The paper is organized as follows: In the next section, a brief explanation of the shifted CG method is given. In Section 3, explicit and implicit evaluation methods for evaluating function values (1.2) are described. The main contribution of the paper is the implicit evaluation method. In Section 4, the explicit/implicit evaluation methods are modified so that they can be applicable to Ye's hybrid method. In Section 5, the results of some numerical experiments are described. Finally some concluding remarks are made in Section 6.