Journal of Computational and Applied Mathematics 298 (2016) 24–39

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

Small-sample statistical condition estimation of large-scale generalized eigenvalue problems

Peter Chang-Yi Weng *, Frederick Kin Hing Phoa

Institute of Statistical Science, Academia Sinica, Taiwan

ARTICLE INFO

Article history:
Received 15 December 2014

MSC:
15A12
15A18
15A22
65F15
65F35
65F50

Keywords:
Statistical condition estimation
Generalized eigenvalue problem
Deflating subspace
Large-scale problem
Sylvester equation
Newton's method

ABSTRACT

We consider the evaluation of the sensitivity or condition number of (generalized) eigenvalue problems for a large and sparse real matrix (or matrix pair) in $\mathbb{R}^{n\times n}$, through some (coupled) Sylvester equation using Newton's method. The technique of the statistical condition estimation has been adapted to the sensitivity of symmetric matrices as well as general matrices with special structures under some assumptions on various types of perturbations.

1. Introduction

For $A, B \in \mathbb{R}^{n\times n}$, consider the generalized eigenvalue problem (GEP):

$$Ax = \lambda Bx, \quad x \neq 0,$$

with $B = I_n$ for the standard eigenvalue problem (SEP). From [1], we can extend to large-scale SEP and compute the sensitivity of the SEP associated with a large and sparse real matrix A (to be specified). Consider the block-Schur decomposition on A

$$P^\top AP = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}, \tag{1}$$

with $A_{ij} \equiv P_i^\top AP_j$ $(i, j = 1, 2)$ and $P \equiv [P_1, P_2] \in \mathbb{R}^{n\times n}$ being orthogonal (or $P^{-1} = P^\top$). We assume that P is in Householder factors [2, p. 224], so that vector multiplications by P can be computed in $O(n)$ flops. Here, $P_1 \in \mathbb{R}^{n\times m}$ $(m \ll n)$ is an accurate estimate to the basis of some invariant subspace associated with A_{11} (see [3]) and the subspectra of the submatrices A_{11} and A_{22} are nonintersecting, giving

$$\sigma(A_{11}) \cap \sigma(A_{22}) = \phi, \tag{2}$$

* Corresponding author.
E-mail address: pweng@stat.sinica.edu.tw (P.C.-Y. Weng).
URL: http://www.stat.sinica.edu.tw/pweng/ (P.C.-Y. Weng).

http://dx.doi.org/10.1016/j.cam.2015.11.022

thus the invariant subspace approximated by span(P_1) is isolated and well-defined. From the definition of an invariant subspace of the matrix A, providing some correction $R \in \mathbb{R}^{m\times(n-m)}$ and the fact that

$$\left\{P\begin{bmatrix} I_m & R \\ 0 & I_{n-m}\end{bmatrix}\right\}^{-1} = \begin{bmatrix} I_m & -R \\ 0 & I_{n-m}\end{bmatrix}P^\top,$$

we have

$$[I_m, \ -R]\,P^\top AP\left[R^\top, \ I_{n-m}\right]^\top = 0,$$

leading to the Sylvester equation

$$A_{12} = RA_{22} - A_{11}R, \tag{3}$$

giving

$$T^{-1}P^\top APT = \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22}\end{bmatrix}, \quad \text{where } T \equiv \begin{bmatrix} I & R \\ 0 & I\end{bmatrix}. \tag{4}$$

Then define

$$T^{-1}P^\top EPT \equiv \begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22}\end{bmatrix}, \tag{5}$$

assume that E is a perturbed matrix such as $\hat{A} = A+E$, devised from a distribution $\mathcal{E} = \{E : \|E\| \le \epsilon\|A\|, \text{ for a scalar } \epsilon > 0\}$.

The perturbed matrix $\hat{A}$ can also be devised on the block-Schur decomposition, combining with (1)

$$P^\top\hat{A}P \equiv \begin{bmatrix} \hat{A}_{11} & \hat{A}_{12} \\ \hat{A}_{21} & \hat{A}_{22}\end{bmatrix} = \begin{bmatrix} A_{11}+E_{11} & A_{12}+E_{12} \\ E_{21} & A_{22}+E_{22}\end{bmatrix}, \tag{6}$$

where $E_{ij} \equiv P_i^\top EP_j$ $(i, j = 1, 2)$.

The Sylvester equation in (3) has appeared frequently in papers associated with SEPs and their perturbation or error analysis and it plays vital roles in many applications such as matrix eigendecompositions [2], control theory [4], model reduction [5], image processing [6], numerical solution of matrix differential Riccati equations [7] and many more. The large-scale Sylvester equation has been solved via the Krylov subspace based algorithms and Alternating-Directional-Implicit (ADI) iterations. The related work can be found in [8–11] and its solution of (3) from refinement is to get the correction R, please consult [12,13,3].

Theorem 1.1 will be modified from [3, Theorem 4.1], which is for the condition numbers of the average eigenvalue of A_{11} and the invariant subspace spanned by the columns of P_1 (the sensitivities of the GEP will be discussed in Theorem 3.1 later).

Theorem 1.1. *Let $A \in \mathbb{R}^{n\times n}$ and $M = [M_1\ M_2]$ be unitary such that*

$$M^\top AM = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22}\end{bmatrix},$$

where the right-hand side is partitioned conformably with M. Define the operator $\mathcal{T}_A$ by $\mathcal{T}_A(Q) = QA_{11} - A_{22}Q$. If the nonsymmetric algebraic Riccati equation (NARE)

$$A_{21} + A_{22}S - SA_{11} - SA_{12}S = 0 \tag{7}$$

has a solution S, then the columns of

$$\hat{M}_1 = (M_1 + M_2S)(I + S^\top S)^{-\frac{1}{2}}$$

are orthogonal and span an invariant subspace of A, and the matrix

$$\tilde{A}_{11} = \hat{M}_1^\top A\hat{M}_1$$

is similar to the matrix $A_{11} + A_{12}S$.

The conditions for the solvability of (7) ((28) for GEP in Section 3) should be satisfied, see [12,14,13,15,1,3] and the perturbation analysis of (7) was also presented in [3].

The condition number of a problem calculates the sensitivity of the solution to small perturbations in the input. We call the problem "well conditioned" if its condition number is small; otherwise the problem is "ill conditioned". Some examples of the famous condition number problems are referred to [16–19]. For the condition numbers of the average eigenvalue of A_{11} and the invariant subspace, the technique we adapt is the statistical condition estimation (SCE). SCE has been usually used in the framework of Monte Carlo trials [20,21] where the SEP has to be solved [22–26] and extended to applications