



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.

© 2015 Elsevier B.V. All rights reserved.

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

$$P^T A P = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}, \quad (1)$$

with $A_{ij} \equiv P_i^T A P_j$ ($i, j = 1, 2$) and $P \equiv [P_1, P_2] \in \mathbb{R}^{n \times n}$ being orthogonal (or $P^{-1} = P^T$). 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

$$\sigma(A_{11}) \cap \sigma(A_{22}) = \emptyset, \quad (2)$$

^{*} 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).

thus the invariant subspace approximated by $\text{span}(P_1)$ 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

$$[I_m, -R] P^\top A P \begin{bmatrix} R^\top & I_{n-m} \end{bmatrix}^\top = 0,$$

leading to the Sylvester equation

$$A_{12} = RA_{22} - A_{11}R, \quad (3)$$

giving

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

Then define

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

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

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

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

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

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

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

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)

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

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

such as linear systems [27], least squares problems [28], and eigenvalues [29–31]. However, this paper is possibly the first on condition estimation of large-scale GEP that is of practical use. Firstly, we apply Newton's method to solve the (coupled) Sylvester equation (3) (or (33)), then we should choose one type of perturbations such as norm-bounded perturbations or componentwise relative perturbations. The issue of choosing such assumption on the type of perturbations is still a problem. The survey of the componentwise relative perturbations can be referred to [32] and also found in [33] for the extension. The following is to adapt the SCE method to estimate the error of the average eigenvalue of A_{11} or the invariant subspace P_1 and finally compute the sensitivities of them associated with the Wallis factors.

For the rest of the paper, we shall make the following assumptions:

- A1. A (and B) are large and structured, with a small number of nonzero elements in each row or column. Also, linear systems associated with $(A - \gamma I)$ (or $(A - \gamma B)$) can be solved efficiently, in the sense of [33, Section 13.3].
- A2. The dimension of the invariant subspace (or deflating subspaces) is small relative to the size of the eigenvalue problem; i.e., $m \ll n$.

Note that the Assumption A1 is quite general, as indicated by [33].

We shall consider the condition number of the large-scale SEP in detail in Section 2, which also contains the formulation for the errors in the average eigenvalues of A_{11} and the invariant subspace and the efficient method for the solvability of Sylvester equations in terms of refinement. Modifications required for the computations in real arithmetic associated with complex eigenvalues are also considered. The sensitivity of the large-scale GEP is treated in Section 3 analogously. Section 4 introduces the method of SCE and applies this estimation method on the large-scale SEP and GEP under a variety of assumptions on the type of perturbations. The algorithms for the computation of sensitivities of average eigenvalue of A_{11} and (A_{11}, B_{11}) , invariant subspace and deflating subspaces on the large-scale SEP and GEP are also provided. Several large-scale examples have been conducted in Section 5 to illustrate the efficiency of our algorithms. Finally, we draw our concluding remarks in Section 6.

2. Perturbation of the standard eigenvalue problem

From Theorem 1.1, we change A and M into \hat{A} and P , respectively, then the invariant subspace spanned by the columns of P_1 is perturbed to the span of the columns of $P_1 + P_2X$,

$$\mathcal{R}\{\hat{P}_1\} = \mathcal{R}\{P_1 + P_2X\},$$

where X satisfies the following equation using (6) and (7)

$$\mathcal{T}_{\hat{A}}(X) = X\hat{A}_{11} - \hat{A}_{22}X = \hat{A}_{21} - X\hat{A}_{12}X = E_{21} - X\hat{A}_{12}X. \quad (8)$$

Since E is small compared with A associated with (2), we will get

$$\|\mathcal{T}_{\hat{A}}^{-1}\|_F \approx \|\mathcal{T}_A^{-1}\|_F.$$

Under these assumptions, $\|X\|_F$ is on the order of $\|E\|_F$ and $\|E\|$ is small, combining with (8)

$$X = \mathcal{T}_{\hat{A}}^{-1}(E_{21} - X\hat{A}_{12}X) \approx \mathcal{T}_A^{-1}(E_{21} - X\hat{A}_{12}X) \approx \mathcal{T}_A^{-1}(E_{21}). \quad (9)$$

For (9), the first-order approximation of the error in the invariant subspace P_1 induced by the perturbation E is listed below

$$\Delta_{P_1} = \mathcal{T}_A^{-1}(P_2^\top EP_1). \quad (10)$$

From Theorem 1.1 associated with (4) and (5), there is an orthogonal matrix \hat{M}_1 such that the spectrum of

$$\hat{M}_1^\top T^{-1} P^\top \hat{A} P T \hat{M}_1 \equiv \hat{A}_{11},$$

is equal to the spectrum of $A_{11} + D_{11} + D_{12}S$, where S is a solution of the NARE. We can obtain that

$$\sigma(\hat{A}_{11}) = \sigma(A_{11} + D_{11} + D_{12}S).$$

Since S is small, we can get

$$\sigma(\hat{A}_{11}) \approx \sigma(A_{11} + D_{11}) = \sigma(A_{11} + (P_1^\top - RP_2^\top)EP_1). \quad (11)$$

When A_{11} has m eigenvalues, the average eigenvalue of A_{11}

$$\mu(A_{11}) = \frac{\text{trace}(A_{11})}{m}.$$

From (11), the error in the average eigenvalue of A_{11} perturbed by E can be measured

$$\Delta_\mu(A_{11}) = \frac{\text{trace}((P_1^\top - RP_2^\top)EP_1)}{m}. \quad (12)$$

2.1. Solution to Sylvester equations

In order to satisfy (4), we have to solve the Sylvester equation

$$RA_{22} - A_{11}R - A_{12} = 0. \quad (13)$$

Using Newton's method on (13) with $R_0 = 0$ and we may solve the Sylvester equations in the form: (for $k \geq 0$)

$$R_{k+1}A_{22} - A_{11}R_{k+1} = A_{12}. \quad (14)$$

For the convergence of Newton's method in (13), please refer to [12,14], which are predicated by [34,35]. Under favourable conditions in [12,34,35], the iterates R_k will converge to the correction R quadratically. In order to compute efficiently, we transform (14) to

$$A_{22}^\top W_1 - W_1 A_{11}^\top = A_{12}^\top, \quad (15)$$

where $W_1 = R_{k+1}^\top$ and it has a slight difference from [15]. Consider the Schur decomposition [2, p. 192] on the small A_{11}^\top such that

$$A_{11}^\top = \hat{Q}Z\hat{Q}^\top \in \mathbb{R}^{m \times m}$$

where \hat{Q} is unitary, Z is upper triangular with z_j ($j = 1, \dots, m$) being the j th diagonal element and $\eta_j \in \mathbb{R}^{j-1}$ ($j = 2, \dots, m$) contains elements above z_j in Z . We can transform (15) to

$$A_{22}^\top W_1 - W_1 \hat{Q}Z\hat{Q}^\top = A_{12}^\top$$

where $X = W_1 \hat{Q}$ and $\hat{A}_{12} = A_{12}^\top \hat{Q}$, then

$$A_{22}^\top X - XZ = \hat{A}_{12}. \quad (16)$$

For $j = 1, \dots, m$, we solve the Sylvester equation (16), providing

$$(A_{22}^\top - z_j I)x_j = (\hat{A}_{12})_j + X_{-j}\eta_j, \quad (17)$$

where $(\cdot)_j$ denotes the j th column, $x_j = (X)_j$ and $X_{-j} = [x_1, \dots, x_{j-1}]$ ($j = 2, \dots, m$) and X_{-1} vanishes. Note that $A_{22}^\top - z_j I$ is nonsingular since the spectra of A_{11} and A_{22} are nonintersecting. As a matter of fact, in many applications, the spectra are far apart and $A_{22}^\top - z_j I$ should be well-conditioned. Note that near-singular linear systems are modified to well-conditioned ones [36,37], but we will solve a smaller well-conditioned linear system via a slightly bigger ill-conditioned one, because of the inherent structure in A^\top . The error analysis has been discussed in [15, Section 2.3.1]. In order to solve the linear systems in (17) efficiently, we will apply the efficient inversion of $A^\top - \gamma I$. For the unified treatment of SEP and GEP, we assume

$$\begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} \equiv P_l^\top (A - \gamma B) P_r, \quad (18)$$

where $P = P_l = P_r$ and $B = I_n$ for the standard eigenvalue problem (SEP). It is vital that P_l and P_r are stored in Householder factors, so that the multiplication by P_l and P_r can be carried out in $O(n)$ computational complexity. In order to compute the inversion of $A^\top - \gamma I$ easily, we modify (18) to get

$$P_r^\top (A^\top - \gamma B^\top) P_l = \begin{bmatrix} C_{11}^\top & C_{21}^\top \\ C_{12}^\top & C_{22}^\top \end{bmatrix}.$$

Let $\hat{J}_{ij} \equiv P_{l_i}^\top (A^\top - \gamma B^\top)^{-1} P_{r_j}$ ($i, j = 1, 2$). we have $\hat{J}_{11} = (C_{11}^\top - C_{21}^\top C_{22}^{-\top} C_{12}^\top)^{-1}$

$$\begin{aligned} P_{l_i}^\top (A^\top - \gamma B^\top)^{-1} P_{r_j} &\equiv \begin{bmatrix} \hat{J}_{11} & \hat{J}_{12} \\ \hat{J}_{21} & \hat{J}_{22} \end{bmatrix} = \begin{bmatrix} C_{11}^\top & C_{21}^\top \\ C_{12}^\top & C_{22}^\top \end{bmatrix}^{-1}, \\ &= \left[\begin{array}{c|c} \hat{J}_{11} & -\hat{J}_{11} C_{21}^\top C_{22}^{-\top} \\ \hline -C_{22}^{-\top} C_{12}^\top \hat{J}_{11} & C_{22}^{-\top} + C_{22}^{-\top} C_{12}^\top \hat{J}_{11} C_{21}^\top C_{22}^{-\top} \end{array} \right]. \end{aligned}$$

We obtain

$$C_{22}^{-\top} v = \hat{J}_{22} v - C_{22}^{-\top} C_{12}^\top \hat{J}_{11} C_{21}^\top C_{22}^{-\top} v. \quad (19)$$

Replacing $C_{22}^{-\top} C_{12}^\top \hat{J}_{11}$ and $C_{21}^\top C_{22}^{-\top}$ by \hat{J}_{21} and $\hat{J}_{11}^{-1} \hat{J}_{12}$, respectively in (19)

$$C_{22}^{-\top} v = \hat{J}_{22} v - \hat{J}_{21} \hat{J}_{11}^{-1} \hat{J}_{12} v. \quad (20)$$

The products $\hat{J}_{12}v$ and $\hat{J}_{22}v$ can be computed through

$$\begin{bmatrix} \hat{J}_{11} & \hat{J}_{12} \\ \hat{J}_{21} & \hat{J}_{22} \end{bmatrix} \begin{bmatrix} 0 \\ v \end{bmatrix} = P_{l_i}^\top (A^\top - \gamma B^\top)^{-1} P_{r_j} \begin{bmatrix} 0 \\ v \end{bmatrix},$$

which is equivalent to the solution

$$(A^\top - \gamma B^\top)y = P_{r_j} \begin{bmatrix} 0 \\ v \end{bmatrix},$$

with $\begin{bmatrix} \hat{J}_{12}v \\ \hat{J}_{22}v \end{bmatrix} = P_{l_i}^\top y$. From (20), we also require \hat{J}_{11} and \hat{J}_{21} and they can be computed from the definition of \hat{J}_{ij} . All the calculations involved in $\hat{J}_{22}v, \hat{J}_{12}v, \hat{J}_{21}$ and \hat{J}_{11} can then be achieved in $O(n)$ flops under our assumptions. For (17), we apply (20) to compute the inversion of $(A_{22}^\top - z_j I)$ efficiently using the structure of A^\top , then retrieve $R = \hat{Q}X^\top$.

Using the real Schur form of A_{11} on (13), we also consider the complex conjugate pair of eigenvalues $z_j, z_{j+1} = \bar{z}_j$ and the corresponding eigenvectors $x_j, x_{j+1} = \bar{x}_j$; We have the real $X_j \in \mathbb{R}^{(n-m) \times 2}$, spanning the same column space as $[x_j, x_{j+1}]$. Then, we have the real 2×2 block Z_j instead of z_j, z_{j+1} , with $\eta_j, \eta_{j+1} \in \mathbb{R}^{j-1}$ above the block Z_j . With $(\cdot)_{j_1 j_2}$ denoting the j_1 to j_2 columns, (16) is changed into

$$A_{22}^\top X_j - X_j Z_j = (\hat{A}_{12})_{j,j+1} + X_{-j}[\eta_j, \eta_{j+1}]. \quad (21)$$

Applying the Kronecker product, (21) becomes

$$M_{kj}v(X_j) = r_j,$$

with $M_{kj} \equiv I_2 \otimes A_{22}^\top - Z_j^\top \otimes I_{n-m}$, $r_j \equiv v[(\hat{A}_{12})_{j,j+1} + X_{-j}[\eta_j, \eta_{j+1}]]$ and $v[\cdot]$ stacking columns. The method of the inversion of M_{kj} is to adapt the similar structures as A_{22}^\top repeated in the two diagonal subblocks, which can be achieved efficiently.

Let $Z_j = \begin{bmatrix} z_j & \beta_{j1} \\ \beta_{j2} & z_{j+1} \end{bmatrix}$, giving

$$M_{kj} = \left[\begin{array}{c|c} A_{22}^\top - z_j I_{n-m} & -\beta_{j2} I_{n-m} \\ \hline -\beta_{j1} I_{n-m} & A_{22}^\top - z_{j+1} I_{n-m} \end{array} \right]. \quad (22)$$

Using the structure of A^\top on (22), we get

$$P_l^\top N_j P_r = \begin{bmatrix} * & * \\ * & M_{kj} \end{bmatrix},$$

with

$$N_j = I_2 \otimes A^\top - Z_j^\top \otimes I_n, \quad P_l = P_r = \begin{bmatrix} P & 0 \\ 0 & P \end{bmatrix} \begin{bmatrix} I_m & 0 & 0 & 0 \\ 0 & 0 & I_{n-m} & 0 \\ 0 & I_m & 0 & 0 \\ 0 & 0 & 0 & I_{n-m} \end{bmatrix}.$$

The inversion of M_{kj} can be computed efficiently by inverting N_j as N_j is structured, P is a projection and M_{kj} is the lower-right corner block in $P_l^\top N_j P_r$.

Consider the sensitivity of the invariant subspace, we will solve the following Sylvester equation from (9)

$$XA_{11} - A_{22}X = E_{21}.$$

Analogously, we adapt Newton's method, combining with the efficient inversion of $A - \gamma I$ for the structure of A , then we will also retrieve a solution X .

3. Perturbation of the generalized eigenvalue problem

Given a large and structured real matrix pair (A, B) , A and $B \in \mathbb{R}^{n \times n}$, we seek to obtain eigenvalues λ_i and the corresponding eigenvectors $x_i \neq 0$ such that $Ax_i = \lambda_i Bx_i$ in the GEP. Consider the eigenvalue decomposition on $(A - \lambda B)$, we get

$$U^\top (A - \lambda B)V = H - \lambda K, \quad (23)$$

with U and $V \in \mathbb{R}^{n \times n}$ being orthogonal and H is upper quasi-triangular and K is upper triangular. A quasi-triangular matrix is block upper triangular with 1×1 or 2×2 blocks on the diagonal and 2×2 blocks on the diagonal of $H - \lambda K$ are pairs of complex conjugate eigenvalues. Giving two conformal partitions $U = [U_1, U_2]$ and $V = [V_1, V_2]$ on (23):

$$\begin{bmatrix} U_1^\top \\ U_2^\top \end{bmatrix} (A - \lambda B) \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} H_{11} & H_{12} \\ 0 & H_{22} \end{bmatrix} - \lambda \begin{bmatrix} K_{11} & K_{12} \\ 0 & K_{22} \end{bmatrix}.$$

Now, $\mathcal{L} = \text{span}(U_1)$ and $\mathcal{R} = \text{span}(V_1)$ form a pair of deflating subspaces associated with (H_{11}, K_{11}) , called left and right deflating subspaces of the spectrum of $(A - \lambda B)$.

The generalized block-Schur decomposition on the matrix pair (A, B) is

$$P^\top (A, B) Q = \left(\begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}, \begin{bmatrix} B_{11} & B_{12} \\ 0 & B_{22} \end{bmatrix} \right), \quad (24)$$

with $A_{ij} = P_i^\top A Q_j$ and $B_{ij} = P_i^\top B Q_j$ ($i, j = 1, 2$), $P \equiv [P_1, P_2]$ and $Q \equiv [Q_1, Q_2] \in \mathbb{R}^{n \times n}$ being real orthogonal. In order to compute efficiently, we store P and Q in Householder factors and $P_1, Q_1 \in \mathbb{R}^{n \times m}$ ($m \ll n$) contain the bases of the approximate left and right deflating subspaces associated with (A_{11}, B_{11}) . The subspectra of the submatrix pairs (A_{11}, B_{11}) and (A_{22}, B_{22}) are nonintersecting such as

$$\sigma(A_{11}, B_{11}) \cap \sigma(A_{22}, B_{22}) = \emptyset. \quad (25)$$

Define

$$G \equiv \begin{bmatrix} I & L \\ 0 & I \end{bmatrix} \quad \text{and} \quad T \equiv \begin{bmatrix} I & R \\ 0 & I \end{bmatrix}.$$

The diagonalization of the generalized block-Schur decomposition can be transformed into

$$G^{-1} \left(\begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}, \begin{bmatrix} B_{11} & B_{12} \\ 0 & B_{22} \end{bmatrix} \right) T = \left(\begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix}, \begin{bmatrix} B_{11} & 0 \\ 0 & B_{22} \end{bmatrix} \right). \quad (26)$$

In order to satisfy (26), we will get the following generalized Sylvester equation

$$L A_{22} - A_{11} R = A_{12}, \quad L B_{22} - B_{11} R = B_{12}. \quad (27)$$

Add the perturbation (E, F) on the matrix pair (A, B) such as $\hat{A} = A + E$ and $\hat{B} = B + F$, giving

$$G^{-1} P^\top E Q T \equiv \begin{bmatrix} I_{11} & I_{12} \\ I_{21} & I_{22} \end{bmatrix} \quad \text{and} \quad G^{-1} P^\top F Q T \equiv \begin{bmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{bmatrix}.$$

Consider the generalized block-Schur decomposition on the perturbed matrix pair (\hat{A}, \hat{B})

$$\begin{aligned} P^\top (\hat{A}, \hat{B}) Q &\equiv \left(\begin{bmatrix} \hat{A}_{11} & \hat{A}_{12} \\ \hat{A}_{21} & \hat{A}_{22} \end{bmatrix}, \begin{bmatrix} \hat{B}_{11} & \hat{B}_{12} \\ \hat{B}_{21} & \hat{B}_{22} \end{bmatrix} \right), \\ &= \left(\begin{bmatrix} A_{11} + E_{11} & A_{12} + E_{12} \\ E_{21} & A_{22} + E_{22} \end{bmatrix}, \begin{bmatrix} B_{11} + F_{11} & B_{12} + F_{12} \\ F_{21} & B_{22} + F_{22} \end{bmatrix} \right), \end{aligned}$$

where $E_{ij} = P_i^\top E Q_j$ and $F_{ij} = P_i^\top F Q_j$ ($i, j = 1, 2$).

Theorem 3.1. Let $A, B \in \mathbb{R}^{n \times n}$, $M = [M_1 \ M_2]$ and $N = [N_1 \ N_2]$ be unitary such that

$$M^\top (A, B) N = \left(\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \right),$$

where the right-hand side matrices are partitioned conformably with M and N . Define the operator $\tilde{\mathcal{T}}$ on the matrix pair (A, B) by $\tilde{\mathcal{T}}_{(A,B)}(P, Q) = (P A_{11} - A_{22} Q, P B_{11} - B_{22} Q)$. If the coupled NARE

$$\begin{aligned} A_{21} + A_{22} S_2 - S_1 A_{11} - S_1 A_{12} S_2 &= 0, \\ B_{21} + B_{22} S_2 - S_1 B_{11} - S_1 B_{12} S_2 &= 0, \end{aligned} \quad (28)$$

has a solution pair (S_1, S_2) , then the columns of

$$\begin{aligned} \tilde{M}_1 &= (M_1 + M_2 S_2)(I + S_1^\top S_2)^{-\frac{1}{2}}, \\ \tilde{N}_1 &= (N_1 + N_2 S_2)(I + S_1^\top S_2)^{-\frac{1}{2}}, \end{aligned}$$

are orthogonal respectively and span left and right deflating subspaces of (A, B) , and the matrix pair

$$(\tilde{A}_{11}, \tilde{B}_{11}) = \tilde{M}_1^\top (A, B) \tilde{N}_1,$$

is similar to the matrix pair $(A_{11} + A_{12} S_2, B_{11} + B_{12} S_2)$.

We change the matrix pair (A, B) and (M, N) into (\hat{A}, \hat{B}) and (P, Q) , respectively then the left and right deflating subspaces spanned by the columns of (P_1, Q_1) are perturbed to the span of the columns of $(P_1 + P_2X, Q_1 + Q_2Y)$,

$$\mathcal{R}\{(\hat{P}_1, \hat{Q}_1)\} = \mathcal{R}\{(P_1 + P_2X, Q_1 + Q_2Y)\},$$

where (X, Y) satisfies the following equation from [Theorem 3.1](#)

$$\begin{aligned}\tilde{\mathcal{T}}_{(\hat{A}, \hat{B})}^{-1}(X, Y) &= (X\hat{A}_{11} - \hat{A}_{22}Y, X\hat{B}_{11} - \hat{B}_{22}Y), \\ &= (\hat{A}_{21} - X\hat{A}_{12}Y, \hat{B}_{21} - X\hat{B}_{12}Y), \\ &= (E_{21} - X\hat{A}_{12}Y, F_{21} - X\hat{B}_{12}Y).\end{aligned}\quad (29)$$

Since (E, F) is small compared with (A, B) associated with [\(2\)](#) and [\(25\)](#), we will get

$$\|\tilde{\mathcal{T}}_{(\hat{A}, \hat{B})}^{-1}\|_F \approx \|\tilde{\mathcal{T}}_{(A, B)}^{-1}\|_F.$$

Under these assumptions, $\|(X, Y)\|_F$ is on the order of $\|(E, F)\|_F$ and $\|E\|, \|F\|$ are small, combining with [\(29\)](#)

$$\begin{aligned}(X, Y) &= \tilde{\mathcal{T}}_{(\hat{A}, \hat{B})}^{-1}(E_{21} - X\hat{A}_{12}Y, F_{21} - X\hat{B}_{12}Y), \\ &\approx \tilde{\mathcal{T}}_{(A, B)}^{-1}(E_{21} - X\hat{A}_{12}Y, F_{21} - X\hat{B}_{12}Y), \\ &= \tilde{\mathcal{T}}_{(A, B)}^{-1}(E_{21}, F_{21}).\end{aligned}\quad (30)$$

For [\(30\)](#), the first-order approximation of the error in the left and right deflating subspaces (P_1, Q_1) induced by the perturbations (E, F) is listed below

$$\Delta_{(P_1, Q_1)} = \tilde{\mathcal{T}}_{(A, B)}^{-1}(P_2^\top EQ_1, P_2^\top FQ_1).$$

Applying [Theorem 3.1](#), there exist two matrices \tilde{M}_1 and \tilde{N}_1 with orthogonal columns such that the spectrum of

$$\tilde{M}_1^\top G^{-1}P^\top(\hat{A}, \hat{B})QT\tilde{N}_1 \equiv (\hat{A}_{11}, \hat{B}_{11}),$$

is equal to the spectrum of $(A_{11} + I_{11} + I_{12}S_2, B_{11} + J_{11} + J_{12}S_2)$, where S_2 solves the coupled NARE. We can get

$$\sigma(\hat{A}_{11}, \hat{B}_{11}) = \sigma(A_{11} + I_{11} + I_{12}S_2, B_{11} + J_{11} + J_{12}S_2). \quad (31)$$

As S_2 is small enough to disregard terms $I_{12}S_2$ and $J_{12}S_2$, [\(31\)](#) is equivalent to

$$\sigma(\hat{A}_{11}, \hat{B}_{11}) \approx \sigma(A_{11} + (P_1^\top - LP_2^\top)EQ_1, B_{11} + (P_1^\top - LP_2^\top)FQ_1).$$

The error in eigenvalues of the matrix pair (A_{11}, B_{11}) can be measured

$$\Delta_\mu(A_{11}, B_{11}) = \mu((P_1^\top - LP_2^\top)EQ_1, (P_1^\top - LP_2^\top)FQ_1). \quad (32)$$

3.1. Generalized Sylvester equations

For [\(27\)](#), it will yield the coupled Sylvester equation

$$LA_{22} - A_{11}R = A_{12}, \quad LB_{22} - B_{11}R = B_{12}. \quad (33)$$

Similar to [\(13\)](#), we apply Newton's method directly on [\(33\)](#) to get the coupled Sylvester equation in the forms: (for $k \geq 0$)

$$L_{k+1}A_{22} - A_{11}R_{k+1} = A_{12}, \quad L_{k+1}B_{22} - B_{11}R_{k+1} = B_{12}. \quad (34)$$

Consider the generalized Schur decomposition [\[2, p. 253\]](#) on $A_{11}, B_{11} \in \mathbb{R}^{m \times m}$

$$A_{11} = \hat{Q}_1 Z_A \hat{Q}_2^\top, \quad B_{11} = \hat{Q}_1 Z_B \hat{Q}_2^\top,$$

with \hat{Q}_1 and \hat{Q}_2 being unitary, Z_A and Z_B being upper triangular with a_j and b_j being the diagonal elements, respectively and η_j, ζ_j containing elements above a_j and b_j in Z_A and Z_B , respectively. The coupled Sylvester equation [\(34\)](#) is then changed into

$$XA_{22} - Z_A Y = \tilde{A}_{12}, \quad XB_{22} - Z_B Y = \tilde{B}_{12}, \quad (35)$$

where $X = \hat{Q}_1^\top L_{k+1}$, $Y = \hat{Q}_2^\top R_{k+1}$, $\tilde{A}_{12} = \hat{Q}_1^\top A_{12}$ and $\tilde{B}_{12} = \hat{Q}_1^\top B_{12}$. For $j = 1, \dots, m$ with $y_j = (Y)_j$ and $Y_{-j} = [y_1, \dots, y_{j-1}]$, [\(35\)](#) can be transformed into

$$\begin{aligned}x_j A_{22} - a_j y_j &= (\tilde{A}_{12})_j + \eta_j Y_{-j}, \\ x_j B_{22} - b_j y_j &= (\tilde{B}_{12})_j + \zeta_j Y_{-j}.\end{aligned}\quad (36)$$

(36) is equivalent to

$$[x_j, y_j] \tilde{M}_{kj} = [z_{j1}, z_{j2}] \equiv [(\tilde{A}_{12})_j + \eta_j Y_{-j}, (\tilde{B}_{12})_j + \zeta_j Y_{-j}], \quad (37)$$

where $\tilde{M}_{kj} \equiv \left[\begin{array}{c|c} A_{22} & B_{22} \\ \hline -a_j I & -b_j I \end{array} \right]$. Assume without loss of generality that $a_j \neq 0$, we obtain

$$\tilde{M}_{kj} = \left[\begin{array}{c|c} A_{22} & \check{N}_j \\ \hline -a_j I & 0 \end{array} \right] \left[\begin{array}{c|c} I & a_j^{-1} b_j I \\ \hline 0 & I \end{array} \right], \quad \check{N}_j \equiv B_{22} - a_j^{-1} b_j A_{22}. \quad (38)$$

Since $\sigma(A_{11}, B_{11}) \cap \sigma(A_{22}, B_{22}) = \emptyset$, \check{N}_j in (38) is nonsingular. From (37) and (38), we can get the following solutions of the coupled Sylvester equation (36)

$$\begin{aligned} x_j &= (-a_j^{-1} b_j z_{j1} + z_{j2}) \check{N}_j^{-1} = (z_{j1} - a_j b_j^{-1} z_{j2}) (A_{22} - a_j b_j^{-1} B_{22})^{-1}, \\ y_j &= a_j^{-1} (x_j A_{22} - z_{j1}). \end{aligned} \quad (39)$$

Applying the structure of the matrix pair (A, B) with the efficient inversion of $(A - \lambda B)$, we can compute the inversion of \check{N}_j efficiently to obtain x_j then take x_j into (39) to get y_j . After getting (X, Y) , we can retrieve $(R, L) = (\hat{Q}_2 Y, \hat{Q}_1 X)$.

We also consider the real generalized Schur form of (A_{11}, B_{11}) , giving

$$\begin{aligned} X_j A_{22} - Z_{aj} Y_j &= \hat{z}_{j1} \equiv (\tilde{A}_{12})_{j,j+1} + [\eta_j, \eta_{j+1}] Y_{-j}, \\ X_j B_{22} - Z_{bj} Y_j &= \hat{z}_{j2} \equiv (\tilde{B}_{12})_{j,j+1} + [\zeta_j, \zeta_{j+1}] Y_{-j}, \end{aligned} \quad (40)$$

with $Z_{aj}, Z_{bj} \in \mathbb{R}^{2 \times 2}$ and Z_{bj} being upper triangular such that

$$Z_{aj} = \begin{bmatrix} a_{j11} & a_{j12} \\ a_{j21} & a_{j22} \end{bmatrix}, \quad Z_{bj} = \begin{bmatrix} b_{j11} & b_{j12} \\ b_{j21} & b_{j22} \end{bmatrix}.$$

Applying the Kronecker product, (40) is transformed into

$$\hat{M}_{kj} \begin{bmatrix} v(X_j) \\ v(Y_j) \end{bmatrix} = \begin{bmatrix} v(\hat{z}_{j1}) \\ v(\hat{z}_{j2}) \end{bmatrix}, \quad \hat{M}_{kj} \equiv \left[\begin{array}{c|c} A_{22}^\top \otimes I_2 & I_{n-m} \otimes -Z_{aj} \\ \hline B_{22}^\top \otimes I_2 & I_{n-m} \otimes -Z_{bj} \end{array} \right]. \quad (41)$$

Without loss of generality, let Z_{aj} be nonsingular and \hat{M}_{kj} be equivalent to

$$\hat{M}_{kj} \equiv \left[\begin{array}{c|c} I & 0 \\ \hline I_{n-m} \otimes Z_{bj} Z_{aj}^{-1} & I \end{array} \right] \left[\begin{array}{c|c} A_{22}^\top \otimes I_2 & I_{n-m} \otimes -Z_{aj} \\ \hline \hat{N}_j & 0 \end{array} \right], \quad (42)$$

with $\hat{N}_j \equiv B_{22}^\top \otimes I_2 - A_{22}^\top \otimes Z_{bj} Z_{aj}^{-1}$. As $\sigma(A_{11}, B_{11})$ does not intersect $\sigma(A_{22}, B_{22})$, \hat{N}_j is nonsingular and well-conditioned. From (41) and (42), the solution of (40) is listed below

$$\hat{N}_j v(X_j) = v(\hat{z}_{j2} - Z_{bj} Z_{aj}^{-1} \hat{z}_{j1}), \quad (43)$$

$$Z_{aj} Y_j = X_j A_{22} - \hat{z}_{j1}. \quad (44)$$

(43) can be transformed into

$$[A_{22}^\top \otimes I_2 - B_{22}^\top \otimes Z_{aj} Z_{bj}^{-1}] v(X_j) = v(\hat{z}_{j1} - Z_{aj} Z_{bj}^{-1} \hat{z}_{j2}). \quad (45)$$

(45) can be computed efficiently using the structure of the matrix pair (A^\top, B^\top) associated with (20), then we can get the solution X . Taking X in (44), we can obtain (X, Y) then retrieve $(R, L) = (\hat{Q}_2 Y, \hat{Q}_1 X)$.

For the condition number of the deflating subspaces, we use (30) to get the following coupled Sylvester equation:

$$\begin{aligned} X A_{11} - A_{22} Y &= E_{21}, \\ X B_{11} - B_{22} Y &= F_{21}. \end{aligned} \quad (46)$$

Similar to (33), we apply Newton's method on (46) firstly then use the structure of the matrix pair (A, B) from the efficient inversion of $(A - \lambda B)$ and finally obtain the solution (X, Y) .

4. Small-sample statistical condition estimation (SCE)

The method, which is to measure the effects on the solution of small random changes in the input data and scale the results properly then obtain condition estimation for each entry of a computed solution, called small-sample statistical

condition estimation (SCE). The SCE method was proposed by Gudmundsson, Kenney and Laub [1,38], which is applied to the estimation of condition numbers for linear systems [27], least-squares problems [28], (generalized) eigenvalue problems [1]. The advantage of SCE is flexibility, applying suitable perturbations according to different structures and the rigorous statistical theory for the probability of accuracy of the condition estimation is referred to [39,38]. First, we discuss an estimator for the sensitivity of scalar-valued functions of several variables and its statistical properties described, then extended to vector-valued functions.

Consider a function $f : \mathbb{R}^p \rightarrow \mathbb{R}$ differentiable at a point $x \in \mathbb{R}^p$, δz is the perturbation of x where $\|z\|_2 \leq 1$ and δ is a small positive number. Then the Taylor expansion of f has the form

$$f(x + \delta z) = f(x) + \delta f_x(x)z + \mathcal{O}(\delta^2),$$

where $f_x(x) = \frac{\partial f}{\partial x}(x)$ is the Fréchet derivative of f at x and the relative error in f using a first-order perturbation is given by

$$\frac{f(x + \delta z) - f(x)}{\delta} \approx f_x(x)z. \quad (47)$$

Then the standard measure of the local sensitivity of f at x is mainly determined by the norm of the Fréchet derivative like $\|f_x(x)\|$ [38]. Define the Wallis factor w_p [38], for an integer p

$$w_p = \begin{cases} 1 & \text{for } p = 1, \\ \frac{2}{\pi} & \text{for } p = 2, \\ \frac{1 \cdot 3 \cdot 5 \cdots (p-2)}{2 \cdot 4 \cdot 6 \cdots (p-1)} & \text{for odd } p > 2, \\ \frac{2 \cdot 2 \cdot 4 \cdot 6 \cdots (p-2)}{\pi \cdot 3 \cdot 5 \cdot 7 \cdots (p-1)} & \text{for even } p > 2. \end{cases} \quad (48)$$

This quantity can be approximated accurately, even for moderate value of p , by

$$w_p \approx \sqrt{\frac{2}{\pi(p - \frac{1}{2})}}. \quad (49)$$

When z is selected uniformly and randomly from the unit p -sphere S_{p-1} , then the condition estimator ξ of f at x can be written into the division of the absolute value of the $f_x(x)z$ and a scaling factor w_p :

$$\xi = |f_x(x)z| / w_p.$$

We can extend it to the subspace condition estimation. There exists randomly and uniformly orthonormal vectors $z_1, z_2, \dots, z_m \in S_{p-1}$, selected from the space of all p -dimensional subspaces of \mathbb{R}^n [40]. Then the norm of the projection of $f_x(x)$ onto the span of z_1, z_2, \dots, z_m is given

$$\sqrt{|f_x(x)z_1|^2 + |f_x(x)z_2|^2 + \cdots + |f_x(x)z_m|^2}.$$

And the subspace condition estimator is as listed

$$\xi(m) = \frac{w_m}{w_p} \sqrt{|f_x(x)z_1|^2 + |f_x(x)z_2|^2 + \cdots + |f_x(x)z_m|^2},$$

where w_p and w_m are Wallis factors with orders p and m , respectively.

If f is a vector-valued function and let $f : \mathbb{R}^p \rightarrow \mathbb{R}^q$ be a differentiable function at a point x . Define the Jacobian at x as

$$J(x) = \frac{\partial f}{\partial x}(x),$$

then the Taylor expansion can be given as follows:

$$f(x + \delta z) = f(x) + \delta J(x)z + \mathcal{O}(\delta^2).$$

Similar to (47), we can get the condition number of a vector-valued function f at the point $x \in \mathbb{R}^p$ such as $\|J(x)\|_2$ [41], then the following condition estimator of f at x is given

$$\varphi = \|J(x)z\| / w_p.$$

And the subspace condition estimator is also obtained

$$\varphi(m) = \frac{w_m}{w_p} \sqrt{\|J(x)z_1\|^2 + \|J(x)z_2\|^2 + \cdots + \|J(x)z_m\|^2}.$$

Next, we concentrate on an estimator of a $q \times p$ matrix L in the Frobenius norm, it is easily computed using a few matrix–vector multiplications and the following definition will be provided.

Definition 4.1. Let $L \in \mathbb{R}^{q \times n}$ and Z be a uniformly random matrix on the Stiefel manifold $V_{m,n}$ and $V_{m,n}$ is the Riemann space of such matrices with orthonormal columns in $\mathbb{R}^{n \times m}$. Define an estimator for $\|L\|_F$ by

$$\psi_L(m) = \sqrt{\frac{p}{m}} \|LZ\|_F.$$

Applying the rules of the Kronecker product on (12)

$$\begin{aligned} \text{trace}(A) &= (\text{vec}(I))^T \text{vec}(A), \\ \text{vec}(ABC) &= (C^T \otimes A) \text{vec}(B), \\ (A \otimes B)^T &= A^T \otimes B^T, \end{aligned} \quad (50)$$

we will obtain the sensitivity for the average eigenvalue of A_{11}

$$\begin{aligned} \phi &= \frac{(\text{vec}(I_k))^T \text{vec}((P_1^T - RP_2^T)EP_1)}{m}, \\ &= \frac{(\text{vec}(I_k))^T (P_1^T \otimes (P_1^T - RP_2^T)) \text{vec}(E)}{m}, \\ &= \frac{b^T \text{vec}(E)}{m}, \end{aligned}$$

where $b = (P_1 \otimes (P_1 - P_2 R^T)) \text{vec}(I_k)$.

For (9), we get the following Sylvester equation

$$XA_{11} - A_{22}X = P_2^T EP_1. \quad (51)$$

It can be written in the form

$$(A_{11}^T \otimes I - I \otimes A_{22}) \text{vec}(X) = \text{vec}(P_2^T EP_1).$$

The solution X can be provided

$$\text{vec}(X) = (A_{11}^T \otimes I - I \otimes A_{22})^{-1} \text{vec}(P_2^T EP_1).$$

Applying the vectorization on (10), giving

$$\text{vec}(\Delta_{P_1}) \approx \text{vec}(X) = (A_{11}^T \otimes I - I \otimes A_{22})^{-1} \text{vec}(P_2^T EP_1). \quad (52)$$

From (52), we can get the sensitivity for the invariant subspace spanned by the columns of P_1 using the property of the Kronecker product (50)

$$\psi = (A_{11}^T \otimes I - I \otimes A_{22})^{-1} (P_1^T \otimes P_2^T) \text{vec}(E) = C \text{vec}(E),$$

with $C \equiv (A_{11}^T \otimes I - I \otimes A_{22})^{-1} (P_1^T \otimes P_2^T)$. The following definition is for the formulations of sensitivities of the average eigenvalue of A_{11} and the invariant subspace, respectively.

Definition 4.2. Let the class of perturbations of A be given by

$$\mathcal{E} = \{E : E = \delta \text{unvec}(Sz), \delta > 0, z \in \mathbb{R}^p, \|z\|_2 \leq 1\},$$

for some fixed matrix S . Then the condition number for the average eigenvalue of the A_{11} is

$$\begin{aligned} \phi &= \left\| \frac{b^T \text{vec}(E)}{m} \right\|_2 = \left\| \frac{b^T \text{vec}(\delta \text{unvec}(Sz))}{m} \right\|_2, \\ &= \left\| \frac{b^T \delta Sz}{m} \right\|_2, \end{aligned} \quad (53)$$

and the condition number for the invariant subspace spanned by the columns of P_1 is

$$\begin{aligned} \psi &= \|C \text{vec}(E)\|_F = \|C \text{vec}(\delta \text{unvec}(Sz))\|_F, \\ &= \|C \delta Sz\|_F. \end{aligned} \quad (54)$$

(53) and (54) are simplified to

$$\begin{aligned}\tau(A, \mathcal{E}) &= \left\| \frac{b^\top S}{m} \right\|_2, \\ \rho(A, \mathcal{E}) &= \|CS\|_F.\end{aligned}$$

Definition 4.3. Let $Z \in \mathbb{R}^{p \times m}$ be an orthonormal basis for $Z \sim \Omega_p^m$, let z_i , $i = 1, 2, \dots, m$, denote the columns of Z , and let S be such that $\text{unvec}(\delta S z_i)$ is an element of \mathcal{E} for each $i = 1, 2, \dots, m$. Then condition estimators for the average eigenvalue of the A_{11} and the corresponding invariant subspace spanned by the columns of P_1 are

$$\phi_\tau = \frac{w_m}{w_p} \left\| \frac{b^\top SZ}{m} \right\|_2, \quad (55)$$

$$\psi_\rho = \sqrt{\frac{p}{m}} \|CSZ\|_F. \quad (56)$$

4.1. Perturbations of GEP

To unify the treatment for two kinds of perturbations of SEP and GEP, we skip B and F then the perturbations of GEP reduces to that of SEP.

4.1.1. Norm-bounded perturbations

One of the norm-bounded perturbations is defined

$$\|(E, F)\|_F \leq \epsilon \|(A, B)\|_F, \quad \text{for a scalar } \epsilon > 0,$$

and it can be characterized as

$$\mathcal{E}_{nb} = \{(E, F) : (E, F) = \epsilon \|(A, B)\|_F Z, Z \in \mathbb{R}^{2n^2}, \|Z\|_2 \leq 1\}.$$

From Definition 4.2, the matrix δS is provided by

$$\delta S = \epsilon \|(A, B)\|_F I_{2n^2}. \quad (57)$$

4.1.2. Componentwise relative perturbations

In the second case, when the entries of (A, B) are perturbed relative to their size

$$|e_{ij}| \leq \epsilon |a_{ij}|, \quad |f_{ij}| \leq \epsilon |b_{ij}|,$$

for some fixed scalar ϵ and $i, j = 1, 2, \dots, n$.

The class of componentwise relative perturbations can be characterized as follows: Let p denote the number of nonzero elements of (A, B) , $a = [a_1, a_2, \dots, a_p]^\top$ denote a vector consisting of nonzero elements of (A, B) , and $C \in \mathbb{R}^{2n^2 \times p}$ be a matrix of ones and zeros such that $C^\top C = I$ and $Ca^\top = \text{vec}(A, B)$, and let

$$D = \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & a_p \end{pmatrix}.$$

Then

$$\mathcal{E}_{cr} = \{(E, F) : (E, F) = \epsilon CDZ, Z \in \mathbb{R}^p, \|Z\|_2 \leq 1\},$$

and the matrix δS is given by

$$\delta S = \epsilon CD. \quad (58)$$

4.2. Algorithm

Here, we provide four algorithms, Algorithms 1–2 for condition numbers of large-scale SEP and others for sensitivities of large-scale GEP. Each of them will discuss the conditionings of the average eigenvalue of A_{11} or (A_{11}, B_{11}) and the corresponding invariant or deflating subspaces. Firstly, consider the condition estimates of SEP then move to that of GEP.

4.2.1. Condition numbers of SEP

We choose the large and sparse real matrix A and find the condition estimates of the average eigenvalue of A_{11} , ϕ_τ and the corresponding invariant subspace ψ_ρ . Solve the Sylvester equations through (13) and (51) to get solutions R and X respectively, then we apply norm-bounded perturbations (57) and componentwise relative perturbations (58) to find S and compute the Wallis factors w_p and w_m , using (55) and (56) to find sensitivities ϕ_τ and ψ_ρ . Algorithms 1 and 2 are summarized below and we would like to emphasize that care has to be exercised in Algorithms 1 and 2, where R and X in (13) and (51) are computed efficiently using (20).

Algorithm 1 (Condition Estimation of the Average Eigenvalue of A_{11})

Input:	A matrix $A \in \mathbb{R}^{n \times n}$, the set of assignment rules S ; the number of samples m , the number of independent variables p , $P = [P_1, P_2] \in \mathbb{R}^{n \times n}$ in Householder factors;
Output:	The condition number for the average eigenvalue of A_{11} , ϕ_τ ; Let the columns of Z be an orthonormal basis for $\mathcal{Z} \sim \Omega_p^m$; Compute the Wallis factors w_p and w_m using (48) and (49); Solve the Sylvester equation $RA_{22} - A_{11}R - A_{12} = 0$ from (3) to get R , using (17) and (21) associated with (20); For $i = 1, \dots, m$ Set z to be the column i of Z ; S can be computed by norm-bounded (57) or componentwise relative perturbations (58); Set $E = \text{unvec}(Sz)$, $E_{11} = P_1^\top EP_1$ and $E_{21} = P_2^\top EP_1$; Compute the error in the average eigenvalue of A_{11} , $\phi_i = \left \frac{\text{trace}((P_1^\top - RP_2^\top)EP_1)}{m} \right $; End for Compute the condition estimate in the average eigenvalue of A_{11} , ϕ_τ as defined in (55). End Do

Algorithm 2 (Condition Estimation of the Invariant Subspace)

Input:	A matrix $A \in \mathbb{R}^{n \times n}$, the set of assignment rules S , the number of samples m , the number of independent variables p , $P = [P_1, P_2] \in \mathbb{R}^{n \times n}$ in Householder factors;
Output:	The condition number of the invariant subspace spanned by the columns of P_1 , ψ_ρ ; Let the columns of Z be an orthonormal basis for $\mathcal{Z} \sim \Omega_p^m$; For $i = 1, \dots, m$ Set z to be the column i of Z ; The set of assignment rules S can be computed by (57) or (58); Set $E = \text{unvec}(Sz)$ and $E_{21} = P_2^\top EP_1$; Solve the Sylvester equation $XA_{11} - A_{22}X - E_{21} = 0$ from (51) to get X ; Compute the error in the invariant subspace, $\psi_i = \ X\ _F$; End for Compute the condition estimate of the invariant subspace, ψ_ρ as defined in (56). End Do

4.2.2. Condition numbers of GEP

Given a large and sparse real matrix pair (A, B) , we compute the condition numbers of the average eigenvalue of (A_{11}, B_{11}) , ϕ_τ and the corresponding deflating subspaces ψ_ρ . If A and B are banded or have favourable sparsity or structures, solving the coupled Sylvester equation (33) has been computed efficiently to get the solutions L and R under our assumptions. Then we adapt two kinds of perturbations like norm-bounded perturbations (57) and componentwise relative perturbations (58) to get S and calculate the Wallis factors w_p and w_m from (48) and (49), combining with (55) and (56) to obtain ϕ_τ and ψ_ρ , respectively. We summarize two algorithms for the GEP in Algorithms 3 and 4 below and need to emphasize that care has to be exercised in Algorithms 3 and 4, where R and L in (33) are computed efficiently using (20), (38) and (42).

Algorithm 3 (Condition Estimation of the Average Eigenvalue of (A_{11}, B_{11}))

Input:	A matrix pair (A, B) , where $A, B \in \mathbb{R}^{n \times n}$, the set of assignment rules S , the number of samples m , the number of independent variables p , $P = [P_1, P_2]$ and $Q = [Q_1, Q_2] \in \mathbb{R}^{n \times n}$ in Householder factors;
Output:	The condition estimate for the average eigenvalue of (A_{11}, B_{11}) , ϕ_τ ; Let the columns of Z be an orthonormal basis for $\mathcal{Z} \sim \Omega_p^m$; Compute the Wallis factors w_p and w_m using (48) and (49); Solve the coupled Sylvester equation (33) to get (L, R) , using (36) and (40) associated with (20), (38) and (42); For $i = 1, \dots, m$ Set z to be the column i of Z ; S can be computed by norm-bounded (57) or componentwise relative perturbations (58); Set $(E, F) = \text{unvec}(Sz)$, $E_{11} = P_1^\top EQ_1$, $E_{21} = P_2^\top EQ_1$, $F_{11} = P_1^\top FQ_1$ and $F_{21} = P_2^\top FQ_1$; Compute the error in the average eigenvalue of (A_{11}, B_{11}) , ϕ_i ; End for Compute the condition estimate in the average eigenvalue of (A_{11}, B_{11}) , ϕ_τ as defined in (55). End Do

Algorithm 4 (Condition Estimation of the Deflating Subspaces)

Input:	A matrix pair (A, B) , where $A, B \in \mathbb{R}^{n \times n}$, the set of assignment rules S , the number of samples m , the number of independent variables p , $P = [P_1, P_2]$ and $Q = [Q_1, Q_2] \in \mathbb{R}^{n \times n}$ in Householder factors;
Output:	The condition estimate of the left and right deflating subspaces, ψ_ρ ; Let the columns of Z be an orthonormal basis for $\mathcal{Z} \sim \Omega_p^m$; For $i = 1, \dots, m$ Set z to be the column i of Z ; The set of assignment rules S can be computed by (57) or (58); Set $(E, F) = \text{unvec}(Sz)$, $E_{21} = P_2^T E Q_1$ and $F_{21} = P_2^T F Q_1$; Solve the coupled Sylvester equation from (30) to get X and Y ; Compute the error in the left and right deflating subspaces, $\psi_i = \ (X, Y)\ _F$; End for Compute the condition estimate of the deflating subspaces, ψ_ρ as defined in (56). End Do

5. Numerical experiments

All the numerical results have been computed using MATLAB [42] Version R2013b on a desktop Acer with a 3.40 GHz Intel Core 2 Duo processor and 16 GB RAM, with machine accuracy $\epsilon_{ps} = 2.22 \times 10^{-16}$.

We have chosen a variety of examples. Examples 5.1–5.6 are artificial examples, where A (or B) is large-scale, sparse, possible with band structures. Our algorithm illustrates the quadratic convergence using Newton's method, computing the (coupled) Sylvester equation from (13) and (33) to get R and L , R respectively, then applies the flexible SCE within \mathcal{N}_A ($\mathcal{N}_{(A,B)}$) and \mathcal{C}_A ($\mathcal{C}_{(A,B)}$) to find sensitivities ϕ_τ and ψ_ρ . The numerical results are presented in Tables 1–8, where Examples 5.1–5.2 for the SEPs of $n = 1000, 2000, 4000, 8000$ and $16,000$; Examples 5.3–5.4 for the ϕ_τ of GEPs of $n = 1000, 2000, 4000, 8000$ and $12,000$ because of the computing resources; Examples 5.5–5.6 for the ψ_ρ of GEPs of $n = 1000, 2000, 4000$ and 8000 . In the former two examples 5.1–5.2 for SEPs, we provide the fixed numbers for the elements of A . Another two Examples 5.3–5.4 for GEPs, we choose the random elements using the normal distribution $\mathcal{N}(0, 1)$ on the three diagonal entries of the matrix pair (A, B) . In the other two examples 5.5–5.6 also for GEPs, we choose the matrix from Examples 5.1–5.2 respectively as one of matrix pair and make a modification of this matrix to get another one of that. Here m denotes the dimension of A_{11} (or B_{11}) and \mathcal{N}_A ($\mathcal{N}_{(A,B)}$), \mathcal{C}_A ($\mathcal{C}_{(A,B)}$) are represented as the norm-bounded perturbations and componentwise relative perturbations, respectively for the matrix A (or the matrix pair (A, B)).

Example 5.1. Example 5.1 for the SEP using SCE in Tables 1 and 2 is symmetric. The symmetric matrix A has zero components except on the three diagonals, which are $a_{i,i} = 2$ on the main diagonal entries and $a_{i,i+1} = a_{i+1,i} = -1$ on the first superdiagonal and subdiagonal entries, respectively for $i = 1, 2, \dots, n$ and is defined as below

$$A = \begin{pmatrix} 2 & -1 & \cdots & 0 \\ -1 & 2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & -1 \\ 0 & \cdots & -1 & 2 \end{pmatrix}.$$

Table 1 discusses sensitivities of the average eigenvalue of A_{11} within two different kinds of perturbations \mathcal{N}_A and \mathcal{C}_A . We set m into 4, which is the number of eigenvalues of A_{11} . For \mathcal{N}_A , the trend of the condition estimation has appeared fluctuation, but the sensitivities will be around 2 for \mathcal{C}_A from $n = 1000$ to $n = 16,000$ because of the norm of perturbation. Moreover, the results of \mathcal{C}_A are smaller than those of \mathcal{N}_A as we perturb on the nonzero entries of A and all elements of A for \mathcal{C}_A and \mathcal{N}_A , respectively and the norm of perturbations for \mathcal{C}_A is smaller than that for \mathcal{N}_A . In Table 2, the sensitivities of the invariant subspace P_1 under \mathcal{N}_A and \mathcal{C}_A are represented. The condition numbers under \mathcal{N}_A and \mathcal{C}_A are increasing with the increase of n from $n = 1000$ to $n = 8000$, and the increase rate of \mathcal{N}_A is faster than that of \mathcal{C}_A . In summary, the condition estimation ϕ_τ and ψ_ρ within \mathcal{C}_A is more stable than that within \mathcal{N}_A for the symmetric case with the fixed three diagonal elements.

Example 5.2. Example 5.2 in Tables 3 and 4 is the modification from [1, Example 1] and A has main diagonal entries $a_{i,i} = i$ and superdiagonal entries $a_{i,i+1} = 1$ and zeros elsewhere for $i = 1, 2, \dots, n$. The following matrix A is as listed

$$A = \begin{pmatrix} 1 & 1 & \cdots & 0 \\ 0 & 2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 0 & \cdots & 0 & n \end{pmatrix}.$$

In Tables 3 and 4, we compute the condition estimation for the nonsymmetric matrix A under \mathcal{N}_A and \mathcal{C}_A . Both results of ϕ_τ and ψ_ρ under \mathcal{N}_A and \mathcal{C}_A are increasing with the increase of n . The interesting thing is to find that the increase rate of

Table 1**Example 5.1** (Condition numbers of the average eigenvalue of A_{11}).

n	1000	2000	4000	8000	16 000
m	4	4	4	4	4
\mathcal{N}_A	3.9293e+01	1.9308e+02	9.3055e+01	2.0043e+03	1.3676e+03
\mathcal{C}_A	2.0498e+00	2.0656e+00	2.0412e+00	2.0014e+00	2.0504e+00

Table 2**Example 5.1** (Condition numbers of the invariant subspace P_1).

n	1000	2000	4000	8000	16 000
m	4	4	4	4	4
\mathcal{N}_A	2.1188e+06	3.8582e+07	9.1714e+07	1.1377e+09	1.0951e+09
\mathcal{C}_A	1.6115e+03	1.2189e+04	2.8830e+04	7.2271e+04	4.7723e+04

Table 3**Example 5.2** (Condition numbers of the average eigenvalue of A_{11}).

n	1000	2000	4000	8000	16 000
m	3	3	3	3	3
\mathcal{N}_A	1.4834e+04	4.3500e+04	1.2678e+05	4.2023e+05	1.2154e+06
\mathcal{C}_A	8.0461e+02	1.3320e+03	2.8360e+03	6.1709e+03	1.5243e+04

Table 4**Example 5.2** (Condition numbers of the invariant subspace P_1).

n	1000	2000	4000	8000	16 000
m	3	3	3	3	3
\mathcal{N}_A	5.1235e+04	9.2281e+04	2.2203e+05	8.6982e+05	2.1838e+06
\mathcal{C}_A	8.6063e+02	2.1019e+03	3.1810e+03	1.0581e+04	2.1207e+04

Table 5**Example 5.3** (Condition numbers of average eigenvalue of (A_{11}, B_{11})).

n	1000	2000	4000	8000	12 000
m	4	4	4	4	4
$\mathcal{N}_{(A,B)}$	3.1025e+03	1.2063e+04	2.1120e+04	3.0384e+04	8.7178e+04
$\mathcal{C}_{(A,B)}$	2.3789e+03	2.4490e+03	4.1024e+03	5.3238e+03	6.9356e+03

Table 6**Example 5.4** (Condition numbers of average eigenvalue of (A_{11}, B_{11})).

n	1000	2000	4000	8000	12 000
m	3	3	3	3	3
$\mathcal{N}_{(A,B)}$	1.1567e+03	1.5391e+03	8.2046e+03	3.6598e+04	1.4557e+04
$\mathcal{C}_{(A,B)}$	3.3098e+02	1.1637e+03	4.0172e+03	1.8216e+04	5.2850e+03

ϕ_τ within \mathcal{N}_A in Table 3 is around triple of the previous one, which is the same as that for the increase rate of the norm of perturbations with n , but the increase rate of ϕ_τ under \mathcal{C}_A starts 1.65 ($\frac{\mathcal{C}_A(n=2000)}{\mathcal{C}_A(n=1000)}$) then goes up to 2.47 ($\frac{\mathcal{C}_A(n=16,000)}{\mathcal{C}_A(n=8000)}$) step by step. To sum up, the increase rate of ϕ_τ under \mathcal{N}_A and \mathcal{C}_A for the nonsymmetric case in Example 5.2 is increasing gradually with n . Similar phenomenon has been appeared in Table 4 for the sensitivities of the invariant subspace P_1 within \mathcal{N}_A and \mathcal{C}_A .

Example 5.3. Example 5.3 for the GEP applying SCE in Table 5 is symmetric. The symmetric matrices A and B have zero components except on the three diagonals, whose nonzero components from $\mathcal{N}(0, 1)$ then they can be defined by the MATLAB command sprandsym [42].

In Table 5, we firstly discuss the sensitivities of the symmetric matrix pair (A, B) for ϕ_τ under $\mathcal{N}_{(A,B)}$ and $\mathcal{C}_{(A,B)}$. The results will be increasing gradually with the increase of n using $\mathcal{N}_{(A,B)}$ and $\mathcal{C}_{(A,B)}$.

Example 5.4. In Example 5.4, we discuss the condition estimation of the GEP for the nonsymmetric case of the matrix pair (A, B) using SCE and the numerical results are represented in Table 6. The nonsymmetric tridiagonal matrices A and B have nonzero components from $\mathcal{N}(0, 1)$.

Example 5.4 has computed ϕ_τ for the nonsymmetric matrix pair (A, B) within $\mathcal{N}_{(A,B)}$ and $\mathcal{C}_{(A,B)}$, respectively. Both of the results within $\mathcal{N}_{(A,B)}$ and $\mathcal{C}_{(A,B)}$ are increasing with the increase of n except $n = 12,000$. That is to say that the condition estimation will not always increase with n , whatever we apply $\mathcal{N}_{(A,B)}$ or $\mathcal{C}_{(A,B)}$ in the nonsymmetric Example 5.4.

Table 7**Example 5.5** (Condition numbers of the deflating subspaces (P_1, Q_1)).

n	1000	2000	4000	8000
m	4	4	4	4
$\mathcal{N}_{(A,B)}$	1.9641e+06	1.0733e+07	8.4230e+07	5.2200e+08
$\mathcal{C}_{(A,B)}$	1.4967e+03	4.2256e+03	1.3652e+04	2.9941e+04

Table 8**Example 5.6** (Condition numbers of the deflating subspaces (P_1, Q_1)).

n	1000	2000	4000	8000
m	3	3	3	3
$\mathcal{N}_{(A,B)}$	1.0300e+07	4.6711e+07	1.4659e+08	3.7181e+08
$\mathcal{C}_{(A,B)}$	7.4136e+00	9.8513e+00	7.6554e+00	9.5381e+00

Example 5.5. Example 5.5 for the GEP using SCE in Table 7 is symmetric. The symmetric matrix A has zero components except on the three diagonals, which are $a_{i,i} = 2$ on the main diagonal entries and $a_{i,i+1} = a_{i+1,i} = -1$ on the first superdiagonal and subdiagonal entries, respectively; the other matrix B also has nonzero elements on the three diagonals with the main diagonal entries $b_{i,i} = 2$, which is the same as $a_{i,i}$ and the first superdiagonal and subdiagonal elements $b_{i,i+1} = b_{i+1,i} = 1$, for $i = 1, 2, \dots, n$. The matrix pair (A, B) is defined as below

$$A = \begin{pmatrix} 2 & -1 & \cdots & 0 \\ -1 & 2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & -1 \\ 0 & \cdots & -1 & 2 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 2 & 1 & \cdots & 0 \\ 1 & 2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 0 & \cdots & 1 & 2 \end{pmatrix}.$$

Example 5.5 discusses the sensitivities of the deflating subspaces (P_1, Q_1) for the symmetric matrix pair (A, B) and the numerical results have been represented in Table 7. The condition estimations under $\mathcal{N}_{(A,B)}$ and $\mathcal{C}_{(A,B)}$ are increasing with the increase of n and the increase rate of ψ_ρ with $\mathcal{N}_{(A,B)}$ is more fast than that with $\mathcal{C}_{(A,B)}$ since the condition numbers are consisted in $[1.96 \times 10^6, 5.22 \times 10^8]$ and $[1.50 \times 10^3, 2.99 \times 10^4]$ applying $\mathcal{N}_{(A,B)}$ and $\mathcal{C}_{(A,B)}$ respectively.

Example 5.6. Example 5.6 for the GEP applying SCE in Table 8 is the modification of Example 5.2. The elements of the matrix A have nonzero elements on the two diagonals with the main diagonal entries $a_{i,i} = 1$ and the first superdiagonal entries $a_{i,i+1} = 2$, for $i = 1, 2, \dots, n$ and the matrix B is the same as the matrix A in Example 5.2. The matrix pair (A, B) is defined as below

$$A = \begin{pmatrix} 1 & 2 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 2 \\ 0 & \cdots & 0 & 1 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 & 1 & \cdots & 0 \\ 0 & 2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 0 & \cdots & 0 & n \end{pmatrix}.$$

In Table 8, we discuss the condition estimation of the deflating subspaces (P_1, Q_1) for the nonsymmetric matrix pair (A, B) using $\mathcal{N}_{(A,B)}$ and $\mathcal{C}_{(A,B)}$. For $\mathcal{N}_{(A,B)}$, the sensitivities will be increasing with the increase of n but the condition numbers under $\mathcal{C}_{(A,B)}$ are between 7 and 10. In summary, the computation of the condition estimation of the deflating subspaces (P_1, Q_1) using $\mathcal{C}_{(A,B)}$ for the nonsymmetric Example 5.6 is more stable than that using $\mathcal{N}_{(A,B)}$ when the size of n becomes larger.

6. Conclusions

The SCE method for the sensitivity of the large-scale GEPs under different perturbations has been proposed. This approach can be also applied to the large-scale SEPs with a much wider class of perturbations. The most important advantage of the SCE is to compute the condition number of multiple eigenvalues rather than that of a single eigenvalue at no extra cost. Our technique solves (couple) Sylvester equation (13) or (33) by the quadratically convergent Newton's method, inverting unstructured matrices $A_{22} - \gamma I_{n-m} (A_{22}^\top - \gamma I_{n-m})$ or $A_{22} - \gamma B_{22} (A_{22}^\top - \gamma B_{22}^\top)$ via structures in A or (A, B) . The computation of the inversion of $A - \gamma I (A^\top - \gamma I)$ or $A - \gamma B (A^\top - \gamma B^\top)$ can be calculated efficiently through the vector multiplications by P_1, P_2 or Q_1, Q_2 in $O(n)$ flops.

In 1994, Kenney and Laub [38] firstly proposed the method called small-sample statistical condition estimates using the probability analysis to estimate the condition numbers of general matrix functions. After three years, Gudmundsson et al. [1] applied the method SCE to compute the sensitivity of small-size SEPs under three different kinds of perturbations, norm-bounded perturbations, componentwise relative perturbations and structured perturbations. For the sensitivity of

small-size GEPs, please refer to Stewart [43]. A lot of applications for the condition estimation of eigenvalue problems have been discussed in [29–31], but this paper is possibly the first on condition estimation of large-scale SEPs and GEPs.

Acknowledgements

This work was supported by (a) Career Development Award of Academia Sinica (Taiwan) grant number 103-CDA-M04 for Phoa and Weng, (b) Ministry of Science and Technology (Taiwan) grant number 102-2628-M-001-002-MY3 for Phoa, (c) Ministry of Science and Technology (Taiwan) grant number 103-2811-M-001-166 for Weng and (d) Ministry of Science and Technology (Taiwan) grant number 104-2118-M-001-016-MY2 for Phoa and Weng.

References

- [1] T. Gudmundsson, C. Kenney, A.J. Laub, Small-sample statistical estimates for the sensitivity of eigenvalue problems, *SIAM J. Matrix Anal. Appl.* 18 (1997) 868–886.
- [2] G.H. Golub, C.F. Van Loan, *Matrix Computations*, third ed., Johns Hopkins University Press, Baltimore, MD, 1996.
- [3] G.W. Stewart, Error and perturbation bounds for subspaces associated with certain eigenvalue problems, *SIAM Rev.* 15 (1973) 727–764.
- [4] B. Datta, *Numerical Methods for Linear Control Systems*, Elsevier Academic Press, 2004.
- [5] U. Baur, P. Benner, Cross-gramian based model reduction for data-sparse systems, *Electron. Trans. Numer. Anal.* 31 (2008) 256–270.
- [6] D. Calvetti, L. Reichel, Application of ADI iterative methods to the restoration of noisy images, *SIAM J. Matrix Anal. Appl.* 17 (1996) 165–186.
- [7] W. Enright, Improving the efficiency of matrix operations in the numerical solution of stiff ordinary differential equations, *ACM Trans. Math. Software* 4 (1978) 127–136.
- [8] A. El Guennouni, K. Jbilou, J. Riquet, Block Krylov subspace methods for solving large Sylvester equations, *Numer. Algorithms* 29 (2002) 75–96.
- [9] D.Y. Hu, L. Reichel, Krylov-subspace methods for the Sylvester equation, *Linear Algebra Appl.* 172 (1992) 283–313.
- [10] A. Lu, E. Wachspress, Solution of Lyapunov equations by ADI iteration, *Comput. Math. Appl.* 21 (1991) 43–58.
- [11] E.L. Wachspress, Iterative solution of the Lyapunov matrix equation, *Appl. Math. Lett.* 1 (1988) 87–90.
- [12] F. Chatelin, Simultaneous Newton's iteration for the eigenproblem, *Comput. Suppl.* 5 (1984) 67–74.
- [13] J.J. Dongarra, C.B. Moler, J.H. Wilkinson, Improving the accuracy of computed eigenvalues and eigenvectors, *SIAM J. Numer. Anal.* 20 (1983) 46–58.
- [14] J.W. Demmel, Three methods for refining estimates of invariant subspaces, *Computing* 38 (1987) 43–57.
- [15] H.-Y. Fan, P.C.-Y. Weng, E.K.-W. Chu, Refining estimates of invariant and deflating subspaces for large and sparse matrices and pencils, *BIT* 54 (2014) 147–169.
- [16] A.K. Cline, C.B. Moler, G.W. Stewart, J.H. Wilkinson, An estimate for the condition number of a matrix, *SIAM J. Numer. Anal.* 16 (1979) 368–375.
- [17] W.W. Hager, Condition estimators, *SIAM J. Sci. Stat. Comput.* 5 (1984) 311–316.
- [18] N.J. Higham, Algorithm 674: Fortran codes for estimating the one-norm of a real or complex matrix, with applications to condition estimation, *ACM Trans. Math. Software* 14 (1988) 381–396.
- [19] G.W. Stewart, On the perturbation of pseudo-inverses, projections and linear least squares problems, *SIAM Rev.* 19 (1977) 634–662.
- [20] J. Hammersley, *Monte Carlo Methods*, Methuen and Wiley, London, 1964.
- [21] H. Rief, A synopsis of Monte Carlo perturbation algorithms, *J. Comput. Phys.* 111 (1994) 33–48.
- [22] I. Dimov, A. Karaivanova, Parallel computations of eigenvalues based on a Monte Carlo approach, *J. Monte Carlo Methods Appl.* 4 (1998) 33–52.
- [23] J. Dufek, Development of new monte carlo methods in reactor physics (Ph.D. thesis), Royal Institute of Technology, 2009.
- [24] S. Matsuura, F.B. Brown, R.N. Blomquist, Parallel Monte Carlo eigenvalue calculations, *Trans. Am. Nucl. Soc.* 71 (1994) 199–202.
- [25] H. Niederreiter, P. Hellekalek, G. Larcher, P. Zinterhof, *Monte Carlo and Quasi-Monte Carlo Methods*, Springer, New York, 1996.
- [26] H. Rief, Generalized Monte Carlo perturbation algorithms for correlated sampling and a second-order Taylor series approach, *Ann. Nucl. Energy* 11 (1984) 455–476.
- [27] C. Kenney, A.J. Laub, M.S. Reese, Statistical condition estimation for linear systems, *SIAM J. Sci. Comput.* 19 (1998) 566–583.
- [28] C. Kenney, A.J. Laub, M.S. Reese, Statistical condition estimation for linear least squares problems, *SIAM J. Matrix Anal. Appl.* 19 (1998) 906–923.
- [29] J.D. Dixon, Estimating extremal eigenvalues and condition numbers of a matrix, *SIAM J. Numer. Anal.* 20 (1983) 812–814.
- [30] R. Fletcher, Expected conditioning, *IMA J. Numer. Anal.* 5 (1985) 247–273.
- [31] G.W. Stewart, Stochastic perturbation theory, *SIAM Rev.* 32 (1990) 579–610.
- [32] N.J. Higham, A survey of componentwise perturbation theory in numerical linear algebra, in: *Proc. Symposia in Applied Mathematics*, vol. 1943–1993, 1994, pp. 49–77.
- [33] N.J. Higham, *Accuracy and Stability of Numerical Algorithms*, Society for Industrial and Applied Mathematics, Philadelphia, 2002.
- [34] J.-G. Sun, Invariant subspaces and generalized invariant subspaces I. Existence and uniqueness theorems, *Math. Numer. Sin.* 2 (1980) 1–14.
- [35] J.-G. Sun, Invariant subspaces and generalized invariant subspaces II. Perturbation theorems, *Math. Numer. Sin.* 2 (1980) 113–123.
- [36] T.F. Chan, D.C. Resasco, Generalized deflated block-elimination, *SIAM J. Numer. Anal.* 23 (1986) 913–924.
- [37] W. Govaerts, J.D. Pryce, Mixed block elimination for linear systems with wider borders, *IMA J. Numer. Anal.* 13 (1993) 161–180.
- [38] C. Kenney, A.J. Laub, Small-sample statistical condition estimates for general matrix functions, *SIAM J. Sci. Comput.* 15 (1994) 36–61.
- [39] T. Gudmundsson, C. Kenney, A.J. Laub, Small-sample statistical estimates for matrix norms, *SIAM J. Matrix Anal. Appl.* 16 (1995) 776–792.
- [40] A.T. James, Normal multivariate analysis and the orthogonal group, *Ann. Math. Statist.* 25 (1954) 40–75.
- [41] J.R. Rice, A theory of condition, *SIAM J. Numer. Anal.* 3 (1966) 287–310.
- [42] Mathworks *MATLAB User's Guide*, 2010.
- [43] G.W. Stewart, On the sensitivity of the eigenvalue problem $Ax = \lambda Bx$, *SIAM J. Numer. Anal.* 9 (1972) 669–686.