

IRR: An Algorithm for Computing the Smallest Singular Value of Large Scale Matrices

Hongbin Guo

Institute of Mathematics, Fudan University, Shanghai, 200433, P.R. China.

e-mail: 970008@fudan.edu.cn

Abstract

For large-scale matrix A , we try to improve the computed smallest singular value with a more satisfied accuracy. In practical applications, the widely used Golub-Kahan-Lanczos-bidiagonalization method nearly always compute the largest singular value with a good relative accuracy, but not the smallest one. In this paper we project $(A^T A)^{-1}$ onto a Krylov subspace, then compute the Rayleigh-Ritz values. We refer this algorithm as Inverse-Rayleigh-Ritz (IRR) method. The technique of computing quadratic form plays an important role in IRR. IRR method is based on Golub-Kahan-Lanczos-bidiagonalization procedure, and gives more accurate results with no more flop cost and storage than Golub-Kahan-Lanczos-bidiagonalization method.

Keywords : Quadratic form; Golub-Kahan-Lanczos-bidiagonalization process; Inverse-Rayleigh-Ritz; The smallest singular value; Large scale matrix.

AMS(MOS) subject classifications : 65F15

1 Introduction

The smallest singular value plays an important role in many applications. In control theory, when we compute the distance to instability (or uncontrollability) and H_∞ norm of a linear control system, we confront the smallest singular value [12] [2][11]. Estimating spectrum condition number also requires computing the smallest singular value. If the scale of matrix is not large, the bidiagonal SVD can compute the smallest singular value accurately enough [5], but computing the smallest singular value of large-scale matrices is a difficult problem. General speaking, “we have no right to ask this problem” ([16]Ch2.7). However, we can improve the existing algorithms to get a better accuracy.

Let $A \in R^{l \times n}$ be nonsingular and $l \geq n$. Golub and Kahan[6] pointed out that the Lanczos recursion together with the matrix

$$C = \begin{bmatrix} 0 & A \\ A^T & 0 \end{bmatrix}$$

could be used to obtain a Lanczos bidiagonalization algorithm for computing all singular values of matrix A . From the view of theory, this algorithm is equal to Lanczos on $A^T A$. Cullum, Willoughby and Lake [3] developed an algorithm following Golub and Kahan’s idea. They [3] claimed that their algorithm can be used to compute the ends portions of the

spectrum. However, numerical results show that the smallest singular value converges too slowly.

table 1 is a example coming from Cullum and Willoughby's paper [3].

Table 1: The computed smallest singular values by Lanczos method

Order: 411×411 , $\sigma_1 = 2.365$, $\sigma_n = 1.46 \times 10^{-8}$

steps	20	412	822	1234	1644
$\bar{\sigma}_1$	2.36524				
$\bar{\sigma}_n$		2.236e-4	4.159e-5	2.33e-7	1.457e-8

This phenomena can be expected, because replacing A by the symmetric matrix C , the smallest singular value of A becomes the middle of spectrum of C .

We know Lanczos method can be used to compute the largest eigenvalues with good accuracy, how about computing the largest eigenvalue of $(A^T A)^{-1}$ on Krylov subspace $K((A^T A)^{-1}, v_1, m)$? Ruhe's rational Krylov sequence method [17] is a projection on $K(S^{-1}, v_1, m)$, S is used to denote a symmetric positive definite (S.P.D.) matrix in this paper. But solving m large-scale linear equations in rational Krylov sequence method is too expensive.

Let us examine another existing algorithm, the so-called harmonic Ritz method [15], harmonic Ritz values are reciprocal of the Ritz values of S^{-1} on space $SK(S, v_1, m)$. By Lanczos process we have :

$$SV_m = V_m T_m + \beta_{m+1, m} v_{m+1} e_m^T = V_{m+1} \bar{T}_m,$$

where $V_{m+1} = [v_1, v_2, \dots, v_{m+1}]$, $V_{m+1}^T V_{m+1} = I$, I denotes the identity matrix, $v_i \in R^n$, $T_m \in R^{m \times m}$ is a tridiagonal matrix, $e_m^T = [0, 0, \dots, 0, 1] \in R^{1 \times m}$, $\beta_{m+1, m} \in R^1$, $\text{span}(V_m) = K(S, v_1, m)$. The Galerkin condition on $W_m = SV_m$ gives:

$$(SV_m)^T (SV_m y - \theta V_m y) = 0,$$

$$W_m^T \left(\frac{1}{\theta} W_m y - S^{-1} W_m y \right) = 0.$$

Solving the generalized eigenvalues problem:

$$W_m^T W_m y = \theta W_m^T V_m y,$$

or

$$\bar{T}_m^T \bar{T}_m y = \theta T_m y,$$

we get the harmonic Ritz values θ_i . We may guess that the smallest harmonic Ritz values (in absolute) would be a better approximation to S 's smallest eigenvalues than the Ritz values. Unfortunately, this is not borne out by examples in [15].

In this paper, we present the IRR method which compute the Ritz values of $(A^T A)^{-1}$ from space V_m , i.e. using the condition:

$$V_m^T ((A^T A)^{-1} V_m y - \theta V_m y) = 0.$$

To compare IRR with harmonic Ritz method, we assume $A^T A$ has spectrum decomposition $A^T A = U \Sigma^2 U^T$, where $U^T U = I, U = (u_1, u_2, \dots, u_n), \Sigma = \text{diag}(\sigma_1, \dots, \sigma_n), \sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n > 0, \sigma_n < 1$ and $V_m = UH, H \in R^{m \times m}$. Then :

$$\|u_n^T V_m\| = \|e_n H\|,$$

$$\|u_n^T W_m\| = \|u_n^T (A^T A) V_m\| = \|u_n^T U \Sigma^2 H\| = \sigma_n^2 \|e_n H\|.$$

where $\|\cdot\|$ denote the 2-norm $\|\cdot\|_2$. Because V_m contains more u_n information than W_m , it is understandable that IRR method is more accurate than harmonic Ritz method.

The key trick in IRR is the technique of computing quadratic form $x^T (A^T A)^{-1} x$ which was studied in [1], [8] and [10]. We review the algorithms for computing quadratic form and give some properties in the following section.

2 Quadratic form

2.1 Basic algorithm

In this section we review some known results and algorithm in [8][1]. Let $S \in R^{n \times n}$ be a S.P.D. matrix, f be a smooth function. We will transform the quadratic form $\mathbf{u}^T f(S) \mathbf{u}$ to Riemann-Stieltjes integral and compute the integral by Gaussian type quadrature rules. Assume $S = Q^T \Lambda Q$, where Q is an orthonormal matrix and Λ is a diagonal matrix with increasingly ordered diagonal elements $\lambda_i \in [a, b]$. Then

$$\chi[f] \equiv \mathbf{u}^T f(S) \mathbf{u} = \mathbf{u}^T Q^T f(\Lambda) Q \mathbf{u} = \tilde{\mathbf{u}}^T f(\Lambda) \tilde{\mathbf{u}} = \sum_{i=1}^n f(\lambda_i) \tilde{u}_i^2 = \int_a^b f(\lambda) d\mu(\lambda),$$

where

$$\mu(\lambda) = \begin{cases} 0 & a \leq \lambda < \lambda_1, \\ \sum_{j=1}^i \tilde{u}_j^2 & \lambda_i \leq \lambda < \lambda_{i+1}, \\ \sum_{j=1}^n \tilde{u}_j^2 & \lambda_n \leq \lambda \leq b. \end{cases}$$

By using Gauss quadrature rule [13][7], we have

$$\chi[f] = \sum_{j=1}^N \omega_j f(\theta_j) + R[f], \quad (2.1)$$

where the weights ω_j and the nodes θ_j are unknown. We denote the error by $R[f]$.

To determine ω_j and θ_j , we construct a sequence of orthonormal polynomials $\{p_k(\lambda)\}_{k=0}^N$, i.e. ,

$$\int_a^b d\mu(\lambda) = 1, \quad (2.2)$$

$$\int_a^b p_i(\lambda) p_j(\lambda) d\mu(\lambda) = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases} \quad (2.3)$$

$$\beta_{k-1} p_k(\lambda) = (\lambda - \alpha_{k-1}) p_{k-1}(\lambda) - \beta_{k-2} p_{k-2}(\lambda) \quad k = 2, 3, \dots, N+1 \quad (2.4)$$

2.2 Properties

Property 2.1 [10] *Let \mathbf{x}_1 be a normalized vector and compute $\mathbf{x}_1^T f(S)\mathbf{x}_1$ by Algorithm 1, then:*

$$\sum_{j=1}^m \omega_j f(\theta_j) = e_1^T f(T_m) e_1.$$

□

If $f(x) = 1/x$, we have $\mathbf{x}_1^T S^{-1}\mathbf{x}_1 = (T_m^{-1})_{11} + R[f]$ and it is easy to get recurrence of $(T_m^{-1})_{11}$ about m .

Property 2.2 [8] *For a S.P.D. matrix S , solving $Sx = b$ by conjugate gradient (CG) method with initial direction $r_0 = b - Sx_0$ and computing $\chi = r_0^T S^{-1} r_0$ by Gauss-type quadrature rules have the same convergence ratio:*

$$\|x^* - x_m\|_S = \|r_0\|_2 (\chi - (T_m^{-1})_{11}).$$

□

We know that the convergence of CG method depends on the condition number of S , initial vector r_0 and the distribution of eigenvalues of S . Now we give a preconditioning method for quadratic form to improve the condition number of S . Let $M = \text{diag}(1/\sqrt{s_{11}}, \dots, 1/\sqrt{s_{nn}})$, where s_{ii} is the i -th diagonal element of S , we configure the preconditioning form:

$$x_1^T S^{-1} x_1 = (Mx_1)^T (MSM)^{-1} (Mx_1).$$

The theory is based on the following theorem ([18], Th4.3)

Theorem 2.3 [18] *Let S be a S.P.D. matrix, all diagonal elements are equal and has at most c non-zero elements in any row. Denote the set of n -by- n diagonal matrices by \mathcal{D}_n . Then:*

$$\kappa(S) \leq c \cdot \min_{D \in \mathcal{D}_n} \kappa(DSD)$$

where $\kappa(S) = \|S\|_2 \|S^{-1}\|_2$.

□

3 Algorithm

In this section, we develop an algorithm to compute the smallest singular value of a nonsingular matrix A . We project $(A^T A)^{-1}$ onto Krylov subspace $K(A^T A, q_1, m)$ to approximate $\|A^{-1}\|_2$.

First, we perform m -step Golub-Kahan-Lanczos-bidiagonalization process

$$\begin{aligned} AQ_m &= U_m B_m, \\ A^T U_m &= Q_m B_m^T + b_m q_{m+1} e_m^T, \end{aligned}$$

where $U_m^T U_m = I, Q_m^T Q_m = I, B_m \in R^{m \times m}$ is an upper bidiagonal matrix

$$B_m = \begin{pmatrix} a_1 & b_1 & & & \\ & \ddots & \ddots & & \\ & & a_{m-1} & b_{m-1} & \\ & & & a_m & \end{pmatrix}. \quad (3.6)$$

To generate the Krylov information of $A^T A$, we deduce the following relations:

$$\begin{aligned} A^T A Q_m &= Q_m B_m^T B_m + a_m b_m q_{m+1} e_m^T, \\ &= Q_m T_m + \beta_m q_{m+1} e_m^T. \end{aligned} \quad (3.7)$$

where tridiagonal matrix $T_m = B_m^T B_m$ has the same form as (2.5), but we do not compute it explicitly. In this paper, when we refer Lanczos method for the smallest singular value problem of an unsymmetric matrix A , we mean Golub-Kahan-Lanczos-bidiagonalization method on A and its result is the smallest singular value of B .

From the former equations, we have

$$\begin{aligned} Q_m^T (A^T A)^{-1} Q_m &= T_m^{-1} - \beta_m Q_m^T (A^T A)^{-1} q_{m+1} e_m^T T_m^{-1} \\ &= T_m^{-1} - \beta_m (Q_m T_m^{-1} - \beta_m (A^T A)^{-1} q_{m+1} e_m^T T_m^{-1})^T q_{m+1} e_m^T T_m^{-1} \\ &= T_m^{-1} + \beta_m^2 q_{m+1}^T (A^T A)^{-1} q_{m+1} T_m^{-1} e_m (T_m^{-1} e_m)^T. \end{aligned} \quad (3.8)$$

Theorem 3.1 *Let A be a nonsingular matrix, $AQ_m = U_m B_m$ and $A^T U_m = Q_m B_m^T + b_m q_{m+1} e_m^T$ is the Lanczos bidiagonalization equations. we have the following projection:*

$$Q_m^T (A^T A)^{-1} Q_m = T_m^{-1} + \beta_m^2 q_{m+1}^T (A^T A)^{-1} q_{m+1} T_m^{-1} e_m (T_m^{-1} e_m)^T. \quad (3.9)$$

where $\beta_m = b_m a_m$ and $T_m = B_m^T B_m$. □

Now we get the projection of $(A^T A)^{-1}$ onto Q_m . Noticing $q_{m+1} \perp Q_m$, if $\text{span}\{Q_m\}$ contains enough direction information corresponding to small eigenvalues of $A^T A$, then $\chi(A, q_{m+1}) = q_{m+1}^T (A^T A)^{-1} q_{m+1}$ will be small. If the whole second term of the right hand of (3.9) is small, our IRR method will not have remarkable effect. This situation take place when Krylov subspace $\text{span}\{Q_m\}$ approach a invariant subspace of $A^T A$ (i.e. β_m small) and Q_m contains enough direction information corresponding to small eigenvalues of $A^T A$. But When the eigenvectors corresponding small eigenvalues convergence too slowly or we have not chosen a proper initial vector, we can expect IRR method will substantial improve the result of Golub-Kahan-Lanczos-bidiagonalization method.

To compute the Ritz value we do singular values decomposition $B = \bar{U} \Sigma \bar{V}^T$ and obtain $T_m = \bar{V} D \bar{V}^T$, where \bar{V} and D satisfy $\bar{V}^T \bar{V} = I, D = \text{diag}(d_1, d_2, \dots, d_m), d_i = \sigma_i^2$ and $d_1 > d_2 > \dots > d_m > 0$. We have:

$$Q_m^T (A^T A)^{-1} Q_m = \bar{V} (D^{-1} + \rho u u^T) \bar{V}^T,$$

where $\rho = \beta_m^2 q_{m+1}^T (A^T A)^{-1} q_{m+1}, u = D^{-1} \bar{V}^T e_m = [\bar{v}_{m1}/d_1, \bar{v}_{m2}/d_2, \dots, \bar{v}_{mm}/d_m]^T$. At this stage we only need to compute $\chi(A, q_{m+1}) = q_{m+1}^T (A^T A)^{-1} q_{m+1}$ and the largest eigenvalue of $D^{-1} + \rho u u^T$. Let ζ denote the largest eigenvalue, then $1/\sqrt{\zeta}$ is an approximation to $\sigma_n(A)$.

We can not compute the quadratic form exactly, but estimate it. To save flop cost in computing χ , of course, we do not run Lanczos process with initial vector q_{m+1} , we can use the following three recurrences instead:

$$\begin{aligned}
\xi_{11} &= q_1^T (A^T A)^{-1} q_1, \\
\xi_{12} &= q_1^T (A^T A)^{-1} q_2 = \frac{1}{\beta_1} q_1^T (A^T A)^{-1} (A^T A q_1 - \alpha_1 q_1) = \frac{1}{\beta_1} (1 - \alpha_1 \xi_{11}), \\
\xi_{22} &= q_2^T (A^T A)^{-1} q_2 = \frac{-\alpha_1}{\beta_1} \xi_{12}, \\
\xi_{i,i+2} &= \frac{1}{\beta_{i+1}} (-\alpha_{i+1} \xi_{i,i+1} - \beta_i \xi_{ii}), \\
\xi_{i+1,i+2} &= \frac{1}{\beta_{i+1}} (1 - \alpha_{i+1} \xi_{i+1,i+1} - \beta_i \xi_{i,i+1}), \\
\xi_{i+2,i+2} &= \frac{1}{\beta_{i+1}} (-\alpha_{i+1} \xi_{i+1,i+2} - \beta_i \xi_{i,i+2}), \quad i = 1, 2, \dots, m-1.
\end{aligned} \tag{3.10}$$

Combining these recurrences with Algorithm 1 gives a method to compute $\chi_m = \xi_{mm}$. Let $R[f] = \chi - \chi_m$ denote the error. It is natural to ask how the error of χ_m affects the error of eigenvalues of $D^{-1} + \chi_m \beta_m^2 uu^T$?

Theorem 3.2 *Let $M(\rho) = D^{-1} + \rho uu^T$, D, u and $\rho = \chi \beta_m^2$ be defined as above and $\zeta(\rho)$ denote the largest eigenvalue of $M(\rho)$. Then $\zeta(\rho)$ is a monotonically increasing function to ρ and,*

$$\left. \frac{d\zeta(\rho)}{d\rho} \right|_{\rho=0} = \left(\frac{\bar{v}_{mm}}{d_m} \right)^2$$

Proof: Denote the eigenvector corresponding to $\zeta(\rho)$ by $\mathbf{z}(\rho)$, after a little induction we have:

$$\frac{d\zeta(\rho)}{d\rho} = \mathbf{z}^T(\rho) \frac{dM(\rho)}{d\rho} \mathbf{z}(\rho) = \mathbf{z}^T uu^T \mathbf{z} \geq 0.$$

Setting $\rho = 0$, we get $\mathbf{z}(0) = e_m$ and

$$\left. \frac{d\zeta(\rho)}{d\rho} \right|_{\rho=0} = |u^T e_m|^2 = \left(\frac{\bar{v}_{mm}}{d_m} \right)^2$$

□

Property 3.3 *Let T_m be defined as above and $T_m = \bar{V} D \bar{V}^T$, $\bar{V} = [\bar{v}_1, \bar{v}_2, \dots, \bar{v}_m]$, where $\bar{v}_m^T = [\bar{v}_{1m}, \bar{v}_{2m}, \dots, \bar{v}_{mm}]$ and $d_1 > d_2 > \dots > d_m > 0$. Then*

$$\frac{1}{d_m} + \chi \left(\frac{\beta_m \bar{v}_{mm}}{d_m} \right)^2 \leq \zeta \leq \frac{1}{d_m} + \chi \left(\frac{\beta_m}{d_m} \right)^2,$$

where $\chi = \chi(A, q_{m+1})$.

proof. By theorem 3.1, we have

$$Q_m^{-T}(A^T A)^{-1}Q_m = T_m^{-1} + \beta_m^2 q_{m+1}^T (A^T A)^{-1} q_{m+1} T_m^{-1} e_m (T_m^{-1} e_m)^T.$$

Denote $Q_m^{-T}(A^T A)^{-1}Q_m$ by M , which is symmetric, then we have

$$\begin{aligned} \zeta &= \max_{x \in R^m} x^T M x \\ &\leq \|T_m^{-1}\|_2 + \chi \beta_m^2 \|T_m^{-1}\|_2^2 \\ &= \frac{1}{d_m} + \chi \left(\frac{\beta_m}{d_m} \right)^2, \end{aligned}$$

and

$$\begin{aligned} \zeta &= \max_{x \in R^m} x^T M x \\ &\geq \|T_m^{-1}\|_2 + \chi \beta_m^2 (\bar{v}_m^T T_m^{-1} e_m)^2 \\ &= \frac{1}{d_m} + \chi \left(\frac{\beta_m \bar{v}_{mm}}{d_m} \right)^2. \end{aligned}$$

□

From Property 3.3 and Cauchy's interlace theorem ([16] Ch10.1) we know $\|(A^T A)^{-1}\|_2 = \|A^{-1}\|_2^2 \geq \zeta \geq 1/d_m$. Thus, computed $\|A^{-1}\|_2$ from IRR is always more accurate than that from Lanczos method. Considering $A^T A Q_m = Q_m T_m + \beta_m q_{m+1} e_m^T$, if $\beta_m \bar{v}_{mm}$ is not small, d_m will not be a good approximation to the smallest eigenvalue of $A^T A$, $\beta_m \bar{v}_{mm}/d_m$ and χ will not be small in turn. Thus, $1/\sqrt{\zeta}$ will improve $\sqrt{d_m}$ surely !

To avoid large condition number, we do not compute ζ and $\sqrt{\zeta}$ directly, but the largest singular value of $(B_m^{-1}, \sqrt{\rho} T_m^{-1} e_m)$.

Algorithm 2

Let $A \in R^{l \times n}$ be nonsingular. This algorithm computes the smallest singular value through Krylov subspace method.

1. Choose an initial vector q_1 and normalize it $q_1 = q_1/\|q_1\|$.
2. Use Golub-Kahan-Lanczos-bidiagonalization process to generate q_i, B_m and $\beta_m = a_m b_m$.
3. Compute $\chi(A, q_{m+1}) = q_{m+1}^T (A^T A)^{-1} q_{m+1}$ by Algorithm 1 and the three term recurrences (3.10). We get $\bar{\chi}$.
4. Compute the largest singular value of $(B_m^{-1}, \sqrt{\rho} T_m^{-1} e_m), \bar{s}_m$, by *qdqs* algorithm ([4]Ch5.4).
5. Set $\bar{\sigma}_n = 1/\bar{s}_m$.

In step 3, $\bar{\chi} = \chi - R[f]$. Due to the error, we can not obtain ζ but $\bar{\zeta}$. This algorithm refers to the matrix A only by matrix-vector computation. Generally speaking, it requires $\mathcal{O}(3l+k)$ memory storage and $\mathcal{O}(jm)$ floating point operations in case A is large scale, where k is the required memory for storing the matrix and forming matrix-vector products, j is the cost of matrix-vector product.

Property 3.4 On the quantity $\overline{s_m}$ in Algorithm 2, we have relation

$$\overline{s_m}^2 \leq \zeta \leq \|A^{-1}\|_2.$$

□

Noticing Theorem 3.2 and error analysis of quadratic form in section 2.1, the property is straightforward.

If $\overline{\sigma}_n$ of Algorithm 2 does not satisfy required accuracy, a restarting strategy can be considered.

4 Numerical examples

In this section, we test five examples. These examples come from Matrix Market, they are all difficult eigenvalue problems. All computations are carried out on PC pentium MMX200 using Matlab 5.0. By Algorithm 2 we perform ‘Mv’ steps Lanczos process with initial vector $q_1^T = (1, 1, \dots, 1)$. Numerical results are listed in table 2:

Table 2: Comparing the computed smallest singular values

Matrix	Order	Cond	Svds	Mv	Lanczos	Harmonic	IRR
cdde4	961	4.03e+4	1.16e-01	100	1.19e-1	3.27e-1	1.16e-1
cdde6	961	1.07e+5	4.39e-02	80	5.89e-2	2.93e-1	4.92e-2
tols340	340	2.03e+5	9.95e-01	40	3.49e+2	5.81e+3	2.89e+0
tols4000	4000	2.36e+7	9.95e-01	100	1.84e+4	3.21e+5	3.91e+0
rw5151	5151	7.30e+37	2.77e-34	60	2.24e-3	7.23e-1	7.21e-7

We compute the ‘exact’ smallest singular value ‘svds’ by Matlab function $svds(A, 1, 0)$. Lanczos results σ_m are filled in column ‘Lanczos’. ‘Harmonic’ stands for the results of harmonic Ritz method based on (3.7) and ‘Cond’ for condition number of A . The error of χ_m affect the results of IRR method. In any case, the results of IRR method are more accurate or converge faster than Golub-Kahan-Lanczos-bidiagonalization method. Thus it can be use as a basic part of a iteration algorithm.

5 Conclusions

In this paper, we consider the problem of computing the smallest singular value of a large scale matrix. A new algorithm IRR is introduced to approximate the smallest singular value. The largest Ritz value based on Krylov subspace $K((A^T A)^{-1}, b, m)$ is derived from Golub-Kahan-Lanczos-bidiagonalization process on A by computing a quadratic form. The cost is cheap enough and the numerical results are more accurate than Lanczos method.

Acknowledgment The author thanks Prof. Zhaojun Bai for helpful suggestions and Prof. Yangfeng Su for careful check.

References

- [1] Z. Bai, Mark Fahey and Gene Golub, Some large-scale matrix computation problems, *J. Comput. Appl. Math.* 74 (1996) 71-89.
- [2] S. Boyd and V. Balakrishnan, A regularity result for the singular values of a transfer matrix and a quadratically convergent algorithm for computing its L_∞ -norm. *System Control Letters*, 1990, 15, 1-7.
- [3] J. Cullum, R. A. Willoughby and M. Lake, A Lanczos algorithm for computing singular values and vectors of large matrices. *SIAM J. Sci. Stat. Comput.*, 1983, vol.4, no.2,197-215.
- [4] J.W. Demmel, *Applied Numerical Linear Algebra*, SIAM press, Philadelphia,1997.
- [5] K. Fernando and B. Parlett, Accurate singular values and differential qd algorithm. *Numer. Math.*, 67 (1994), 191-229.
- [6] G. Golub and W. Kahan, Calculating the singular values and pseudoinverse of a matrix. *SIAM J. Numer. Anal.(Series B)*,1965, 2:205-224.
- [7] G. Golub, Some modified matrix eigenvalue problems, *SIAM Rev.*, 15(1973)318-334.
- [8] G. Golub and Z. Strakös, Estimates in quadratic formulas. *Numer. Algor.*,1994, 8, 241-268.
- [9] G. Golub and C. Van Loan, *Matrix Computation*, John Hopkins University press, Baltimore, MD, 3rd edition, 1996.
- [10] Hongbin Guo, Computing trace of function of matrix. *Numerical Math., A Journal of Chinese Universities, English series*,No.2, 2000.
- [11] C. He, Estimating the distance to uncontrollability: a fast method and a slow one. *System Control Letters*, 1995, 26, no.4, 275-281.
- [12] C. He and G.A.Watson, An algorithm for computing the distance to instability. *SIAM J. Matrix Anal. Appl.*, 1999 vol.20,no.1,101-116(electronic).
- [13] E. X. Jiang, F. Zhao, *Numerical Approximation* (Chinese edition) 1996, Fudan University press.
- [14] E. X. Jiang, *The Symmetric Matrices Computation*. (Chinese edition) 1984, Shanghai Science and Technology press.
- [15] C. Paige, Approximate solutions and eigenvalue bounds from Krylov subspaces. *Numer. Linear Algebra Appl.*, 1995,vol.2(2),115-133.
- [16] B. N. Parlett, *The symmetric eigenvalue problem*. SIAM press, Philadelphia, 1997.
- [17] A. Ruhe, Rational Krylov sequence methods for eigenvalue computation. *Linear Algebra Appl.*, 1984,vol.58, 391-405.
- [18] A. Van Der Sluis, Condition number and equilibration of matrices. *Numer. Math.*, 1969,14,14-23.