Frontiers in Environmental Research
ISSN: 2960-026X
DOI: https://doi.org/10.61784/fer3011

THE CONVERGENCE RATE ANALYSIS OF CONJUGATE GRADIENT METHOD IN TYPHOON FORECASTING

HaiFeng Liu
School of Mathematics (Zhuhai), Sun Yat-sen University, Zhuhai 519082, Guangdong, China.
Corresponding Email: liuhf27@mail.sysu.edu.cn

Abstract: During the process of typhoon forecasting, numerious symmetric positive definite linear systems are needed to be solved. They are often solved by conjugate gradient method with preconditioning technique. This paper focuses on the convergence rate analysis of conjugate gradient method. The properties of Chebyshev polynomial and Krylov subspace are utilized. The effect of the right-hand-side vector are considered. Several convergence rate estimations are given. Compared with the existing estimation results, the presented results are more exact. This enable us to construct more efficient preconditioners to forecast typhoon more quickly.
Keywords: Typhoon forecasting; Conjugate gradient method; Convergence rate; Eigenvalue

1 INTRODUCTION

Typhoons often bring inconvenience and even disasters to our daily lives. Thus, typhoon forecasting is crucial, especially for the coastal areas [1-2]. Typhoon forecasting always involves a large number of symmetric positive definite (SPD) linear systems. These systems are often solved by the conjugate gradient (CG) method with preconditioning technique [3-4], which is called preconditioned CG method. The key of preconditioned CG method is the construction of the preconditioner [5-8]. For constructing efficient preconditioners during the process of typhoon forecast, the theoretical analysis of CG method is studied in this paper. Consider the following SPD linear system

$$Ax = b, \tag{1}$$

and solving equation (1) by preconditioned CG method. The base of preconditioned CG method is CG method. The main results regarding the convergence rate of the CG method are given in the following [9, 10].

Theorem 1. Assume $A \in R^{n\times n}$ is SPD, its eigenvalues are $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n > 0$. The convergent rate of CG method for solving problem (1) can be estimated by the following inequality

$$\|x_k - x^*\|_A \leq 2\left(\frac{\sqrt{\lambda_1}-\sqrt{\lambda_n}}{\sqrt{\lambda_1}+\sqrt{\lambda_n}}\right)^k \|x_0 - x^*\|_A \tag{2}$$

where $\|x\|_A = (Ax,\ x)^{\frac{1}{2}}$, x_k, x_0 and x^* represent the kth approximate solution, the initial guess and the exact solution, respectively.

By theorem 1, the smaller the spectral condition number $\mathrm{cond}_2(A) = \lambda_1/\lambda_n$ of coefficient matrix A is, the more fast the convergence rate of CG method is.

Theorem 2. Assume $A \in R^{n\times n}$ is SPD, the CG method for solving problem (1) converges after n iterations at most.

By theorem 2, when the CG method was initially proposed, it was initially regarded as a direct method for solving SPD linear algebraic equations. Until 1971, Reid pointed out that viewing the CG method as an iterative approach can effectively solve large-scale sparse SPD linear systems.

Theorem 3. Assume $A \in R^{n\times n}$ is SPD and has l distinct eigenvalues, the CG method for solving problem (1) converges after l iterations at most.

Theorem 3 shows that even if the condition number of coefficient matrix A is large, if the number of its distinct eigenvalues is small, the CG method will also converge fast. It can be seen that, the above three theorem does not consider the influence of the right hand side vector b on the convergence rate of the CG method. This paper study the effect of the right hand side vector b on the convergence rate of the CG method.

2 MAIN RESULTS

Theorem 4. Assume $A \in R^{n\times n}$ is SPD, its eigenvalues are $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n > 0$. The corresponding orthonormal eigenvectors are $u_1, u_2, \cdots, u_n$, the right hand side vector $b \in \mathrm{span}\{u_{i_1}, u_{i_2}, \cdots, u_{i_s}\}$. Let the initial guess $x_0 = 0$, then the convergent rate of CG method for solving problem (1) can be estimated by the following inequality

$$\|x_k - x^*\|_A \leq 2\left(\frac{\sqrt{\lambda_{i_1}}-\sqrt{\lambda_{i_s}}}{\sqrt{\lambda_{i_1}}+\sqrt{\lambda_{i_s}}}\right)^k \|x_0 - x^*\|_A. \tag{3}$$

where $\|x\|_A = (Ax,\ x)^{\frac{1}{2}}$, x_k and x^* represent the kth approximate solution and the exact solution, respectively.

Note that $\frac{\sqrt{\lambda_{i_1}}-\sqrt{\lambda_{i_s}}}{\sqrt{\lambda_{i_1}}+\sqrt{\lambda_{i_s}}} \le \frac{\sqrt{\lambda_1}-\sqrt{\lambda_n}}{\sqrt{\lambda_1}+\sqrt{\lambda_n}}$ due to $\frac{\lambda_{i_1}}{\lambda_{i_s}} \le \frac{\lambda_1}{\lambda_n}$. Thus, theorem 4 is an efficient improvement on theorem 1.

Theorem 5. Assume $A \in R^{n\times n}$ is SPD, its eigenvalues are $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_n > 0$. The corresponding orthonormal eigenvectors are $u_1, u_2, \cdots, u_n$, the right hand side vector $b \in \text{span}\{u_{i_1}, u_{i_2}, \cdots, u_{i_s}\}$. Let the initial guess $x_0 = 0$, then the CG method for solving problem (1) converges after s iterations at most.

Note that $s \le n$. Thus, the result of theorem 5 is more precisely than that of theorem 2.

Theorem 6. Assume $A \in R^{n\times n}$ is SPD, its eigenvalues are $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_n > 0$. The corresponding orthonormal eigenvectors are $u_1, u_2, \cdots, u_n$, the right hand side vector $b \in \text{span}\{u_{i_1}, u_{i_2}, \cdots, u_{i_s}\}$. If $\lambda_{i_1}, \lambda_{i_2}, \cdots, \lambda_{i_s}$ are t distinct eigenvalues of A $(t \le s)$. Let the initial guess $x_0 = 0$, then the CG method for solving problem (1) converges after t iterations at most.

Note that $\{\lambda_{i_1}, \lambda_{i_2}, \cdots, \lambda_{i_s}\}$ is a subset of the set that includes all eigenvalues of A, thus t in theorem 6 is less than l in theorem 3. Theorem 6 is an efficient improvement on theorem 3.

3 PRELIMINARIES

The CG method can be implemented as follows [9, 10]

Algorithm 1 (CG method)

x_0 is the initial guess, usually selected as zero vector. ε is a specified error tolerance

$$r_0 = b - Ax_0\ ;\ k = 0 \tag{4}$$

While

$$\begin{gathered} \| r_k\|_2 \ge \varepsilon \| b \|_2 \\ k = k + 1 \\ \text{if } k = 1 \\ p_0 = r_0 \end{gathered} \tag{5}$$

else

$$\begin{gathered} \beta_{k-2} = r_{k-1}^T r_{k-1} / r_{k-2}^T r_{k-2} \\ p_{k-1} = r_{k-1} + \beta_{k-2} p_{k-2} \end{gathered} \tag{6}$$

end

$$\begin{gathered} \alpha_{k-1} = r_{k-1}^T r_{k-1} / p_{k-1}^T A p_{k-1} \\ x_k = x_{k-1} + \alpha_{k-1} p_{k-1} \\ r_k = r_{k-1} - \alpha_{k-1} A p_{k-1} \end{gathered} \tag{7}$$

end

$x = x_k$.

It can be seen from the above algorithm, the main computational cost of CG method in one iteration is a matrix- vector multiplication. Thus, the sparsity of the coefficient matrix can be made full use of in order to reduce the computational cost. If the coefficient matrix is dense but have special structure, the computational cost may also be reduced.

Lemma 1. [10] The kth approximate solution x_k produced by CG method satisfies

$$\|x_k - x^*\|_A = \min\{\|x - x^*\|_A\ ;\ x \in x_0 + \mathcal{K}(A, r_0, k)\}, \tag{8}$$

where $\mathcal{K}(A, r_0, k) = \text{span}\{r_0, Ar_0, \cdots, A^{k-1}r_0\}$, it is called Krylov subspace.

Lemma 2. [10] Let $\mathcal{P}_k$ denote all real coefficient polynomial with degree not exceeding k, and $P_k(0) = 1$. $b > a > 0$, then the following optimization problem

$$\min_{P_k\in\mathcal{P}_k} \max_{a\le\lambda\le b} |P_k(\lambda)|, \tag{9}$$

has a unqiue solution

$$\widetilde{P_k}(\lambda) = \frac{T_k\left(\frac{b+a-2\lambda}{b-a}\right)}{T_k\left(\frac{b+a}{b-a}\right)}, \tag{10}$$

where $T_k(\lambda)$ is Chebyshev polynomial of degree k, and

$$\min_{P_k\in\mathcal{P}_k} \max_{a\le\lambda\le b} |P_k(\lambda)| = \max_{a\le\lambda\le b} \left|\widetilde{P_k}(\lambda)\right| = \frac{1}{T_k\left(\frac{b+a}{b-a}\right)} \le 2\left(\frac{\sqrt{b}-\sqrt{a}}{\sqrt{b}+\sqrt{a}}\right)^k. \tag{11}$$

Lemma 3. [10] The vectors $r_i, p_i\ (0 \le i \le k)$ produced by CG method satisfy

$$\text{span}\{r_0, \cdots, r_k\} = \text{span}\{p_0, \cdots, p_k\} = \mathcal{K}(A, r_0, k+1). \tag{12}$$

Lemma 4. Assume $A \in R^{n\times n}$ is SPD, its eigenvalues are $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_n > 0$. The corresponding orthonormal eigenvectors are $u_1, u_2, \cdots, u_n$. $y \in R^n$, the vector $y \in \text{span}\{u_{i_1}, u_{i_2}, \cdots, u_{i_s}\}$. $P_k(\lambda)$ is a polynomial of λ with degree not exceeding k. Then,

$$\|P_k(A)y\|_A \le \max_{1\le j\le s} \left|P_k\left(\lambda_{i_j}\right)\right| \|y\|_A. \tag{13}$$

Proof: By the known condition, there exist $c_1, c_2, \cdots, c_s \in R$ such that $y = c_1u_{i_1} + c_2u_{i_2} + \cdots + c_su_{i_s}$. Thus,

$$\|P_k(A)y\|_A^2 = (AP_k(A)y,\ P_k(A)y)$$

$$
\begin{aligned}
&= \left(AP_k(A)\left(c_1u_{i_1} + c_2u_{i_2} + \cdots + c_su_{i_s}\right),\ \ P_k(A)\left(c_1u_{i_1} + c_2u_{i_2} + \cdots + c_su_{i_s}\right)\right) \\
&= \left(\lambda_{i_1}P_k(\lambda_{i_1})c_1u_{i_1} + \cdots + \lambda_{i_s}P_k(\lambda_{i_s})c_su_{i_s},\ \ P_k(\lambda_{i_1})c_1u_{i_1} + \cdots + P_k(\lambda_{i_s})c_su_{i_s}\right) \\
&= P_k^2(\lambda_{i_1})\left(\lambda_{i_1}c_1u_{i_1}, c_1u_{i_1}\right) + \cdots + P_k^2(\lambda_{i_s})\left(\lambda_{i_s}c_su_{i_s}, c_su_{i_s}\right) \\
&\le \max_{1\le j\le s}\left|P_k\left(\lambda_{i_j}\right)\right|^2\left[\left(\lambda_{i_1}c_1u_{i_1}, c_1u_{i_1}\right) + \cdots + \left(\lambda_{i_s}c_su_{i_s}, c_su_{i_s}\right)\right] \\
&= \max_{1\le j\le s}\left|P_k\left(\lambda_{i_j}\right)\right|^2\left(\lambda_{i_1}c_1u_{i_1} + \cdots + \lambda_{i_s}c_su_{i_s}, c_1u_{i_1} + \cdots + c_su_{i_s}\right) \\
&= \max_{1\le j\le s}\left|P_k\left(\lambda_{i_j}\right)\right|^2\left(A\left(c_1u_{i_1} + \cdots + c_su_{i_s}\right), c_1u_{i_1} + \cdots + c_su_{i_s}\right) \\
&= \max_{1\le j\le s}\left|P_k\left(\lambda_{i_j}\right)\right|^2(Ay, y) \\
&= \max_{1\le j\le s}\left|P_k\left(\lambda_{i_j}\right)\right|^2\|y\|_A^2.
\end{aligned}
\tag{14}
$$

4 THE PROOF OF THE MAIN RESULTS

Proof of theorem 4: Note that the initial guess $x_0 = 0$, we have

$$r_0 = b - Ax_0 = b = Ax^*. \tag{15}$$

Thus, $x^* = A^{-1}r_0$. By lemma 3, for any $x \in x_0 + \mathcal{K}(A, r_0, k)$, the following relationship holds true

$$
\begin{aligned}
x^* - x = x^* + d_{k0}r_0 + d_{k1}Ar_0 + \cdots + d_{k\,k-1}A^{k-1}r_0 = \\
A^{-1}\left(r_0 + d_{k0}Ar_0 + d_{k1}A^2r_0 + \cdots + d_{k\,k-1}A^kr_0\right) = A^{-1}P_k(A)r_0,
\end{aligned}
\tag{16}
$$

where $P_k(\lambda) = 1 + \sum_{j=1}^{k} d_{k\,j-1}\lambda^j$. Note that $r_0 = b \in \operatorname{span}\{u_{i_1}, u_{i_2}, \cdots, u_{i_s}\}$, thus

$$A^{-1}r_0 \in \operatorname{span}\{u_{i_1}, u_{i_2}, \cdots, u_{i_s}\}. \tag{17}$$

Lemmas 2 and 3 indicate that

$$
\begin{aligned}
\|x_k - x^*\|_A &= \min\left\{\|x - x^*\|_A\ ;\ x \in x_0 + \mathcal{K}(A, r_0, k)\right\} \\
&= \min_{P_k\in\mathcal{P}_k}\|A^{-1}P_k(A)r_0\|_A \\
&= \min_{P_k\in\mathcal{P}_k}\|P_k(A)A^{-1}r_0\|_A \\
&\le \min_{P_k\in\mathcal{P}_k}\max_{a\le\lambda\le b}\left|P_k\left(\lambda_{i_j}\right)\right|\|A^{-1}r_0\|_A \\
&\le 2\left(\frac{\sqrt{\lambda_{i_1}} - \sqrt{\lambda_{i_s}}}{\sqrt{\lambda_{i_1}} + \sqrt{\lambda_{i_s}}}\right)^k\|x_k - x^*\|_A.
\end{aligned}
\tag{18}
$$

Similar to the above proof, the following result can be easily obtained

Corollary 1. Assume $A \in R^{n\times n}$ is SPD, its eigenvalues are $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_n > 0$, the corresponding orthonormal eigenvectors are $u_1, u_2, \cdots, u_n$. If the initial guess x_0 satisfies $b - Ax_0 \in \operatorname{span}\{u_{i_1}, u_{i_2}, \cdots, u_{i_s}\}$, then the kth approximate solution x_k produced by CG method for solving problem (1) satisfies

$$\|x_k - x^*\|_A \le 2\left(\frac{\sqrt{\lambda_{i_1}} - \sqrt{\lambda_{i_s}}}{\sqrt{\lambda_{i_1}} + \sqrt{\lambda_{i_s}}}\right)^k\|x_0 - x^*\|_A. \tag{19}$$

Proof of theorem 5: Note that $b \in \operatorname{span}\{u_{i_1}, u_{i_2}, \cdots, u_{i_s}\}$. Thus, there exist $c_1, c_2, \cdots, c_s \in R$ such that

$$b = c_1u_{i_1} + c_2u_{i_2} + \cdots + c_su_{i_s}. \tag{20}$$

We just need to proof that for any positive integer k, the dimension of Krylov subspace $\mathcal{K}(A, r_0, k) = \operatorname{span}\{r_0, Ar_0, \cdots, A^{k-1}r_0\}$ is not exceeding s. Note that the initial guess $x_0 = 0$, thus

$$
\begin{aligned}
r_0 = b - Ax_0 = b = c_1u_{i_1} + c_2u_{i_2} + \cdots + c_su_{i_s}, \\
Ar_0 = c_1Au_{i_1} + c_2Au_{i_2} + \cdots + c_sAu_{i_s} = c_1\lambda_{i_1}u_{i_1} + c_2\lambda_{i_2}u_{i_2} + \cdots + c_s\lambda_{i_s}u_{i_s}.
\end{aligned}
\tag{21}
$$

It can be calculated that for any positive integer k, the following relationship holds true

$$A^{k-1}r_0 = c_1\lambda_{i_1}^{k-1}u_{i_1} + c_2\lambda_{i_2}^{k-1}u_{i_2} + \cdots + c_s\lambda_{i_s}^{k-1}u_{i_s}. \tag{22}$$

This shows that the vector group $r_0, Ar_0, \cdots, A^{k-1}r_0$ can be linearly represented by the vector group $u_{i_1}, u_{i_2}, \cdots, u_{i_s}$ and its rank is not exceeding s. Thus,

$$\dim\mathcal{K}(A, r_0, k) = \dim\left(\operatorname{span}\{r_0, Ar_0, \cdots, A^{k-1}r_0\}\right) \le s. \tag{23}$$

By Lemma 1,

$$\|x_k - x^*\|_A = \min_{x\in x_0 + \mathcal{K}(A, r_0, k)}\|x - x^*\|_A. \tag{24}$$

The above equation indicates that the CG method can achieve the exact solution of the linear system $Ax = b$ after s iterations at most.

Simply modify the above proof, the following corollary can be obtained.

Corollary 2. Assume $A \in R^{n\times n}$ is SPD, its eigenvalues are $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n > 0$, the corresponding orthonormal eigenvectors are $u_1, u_2, \cdots, u_n$. If the initial guess x_0 satisfies $b - Ax_0 \in \operatorname{span}\{u_{i_1}, u_{i_2}, \cdots, u_{i_s}\}$, then the CG method for solving problem (1) can achieve the exact solution after s iterations at most.

Proof of theorem 6: Let's assume $\lambda_{i_1} = \lambda_1$, $\lambda_{i_2} = \lambda_2$, $\cdots$, $\lambda_{i_s} = \lambda_s$, i.e, $\lambda_{i_1}, \lambda_{i_2}, \cdots, \lambda_{i_s}$ is the first s eigenvalues of matrix A. At the same time, they are t distinct eigenvalues of A. Thus, we can assume

$$\underbrace{\lambda_1 = \cdots = \lambda_{s_1}}_{s_1}, \underbrace{\lambda_{s_1+1} = \cdots = \lambda_{s_1+s_2}}_{s_2}, \cdots, \underbrace{\lambda_{s_1+\cdots+s_{t-1}+1} = \cdots = \lambda_{s_1+\cdots+s_{t-1}+s_t}}_{s_t} \tag{25}$$

where $s_1 + s_2 + \cdots + s_t = s$. The corresponding orthonormal eigenvectors are

$$u_1, \cdots, u_{s_1}, u_{s_1+1}, \cdots, u_{s_1+s_2}, \cdots, u_{s_1+\cdots+s_{t-1}+1}, \cdots, u_{s_1+\cdots+s_{t-1}+s_t}. \tag{26}$$

Note that the initial guess $x_0 = 0$, thus $r_0 = b - Ax_0 = b$. By the known condition, we can assume

$$\begin{gathered} r_0 = c_1u_1 + \cdots + c_{s_1}u_{s_1} + c_{s_1+1}u_{s_1+1} + \cdots + c_{s_1+s_2}u_{s_1+s_2} + \cdots + \\ c_{s_1+\cdots+s_{t-1}+1}u_{s_1+\cdots+s_{t-1}+1} + \cdots + c_{s_1+\cdots+s_{t-1}+s_t}u_{s_1+\cdots+s_{t-1}+s_t}. \end{gathered} \tag{27}$$

Let

$$\begin{gathered} z_1 = c_1u_1 + \cdots + c_{s_1}u_{s_1}, \\ z_2 = c_{s_1+1}u_{s_1+1} + \cdots + c_{s_1+s_2}u_{s_1+s_2}, \\ \cdots \\ z_t = c_{s_1+\cdots+s_{t-1}+1}u_{s_1+\cdots+s_{t-1}+1} + \cdots + c_{s_1+\cdots+s_{t-1}+s_t}u_{s_1+\cdots+s_{t-1}+s_t}. \end{gathered} \tag{28}$$

Then,

$$\begin{gathered} r_0 = z_1 + z_2 + \cdots + z_t, \\ Ar_0 = \lambda_1z_1 + \lambda_{s_1+1}z_2 + \cdots + \lambda_{s_1+\cdots+s_{t-1}+1}z_t. \end{gathered} \tag{29}$$

It can be calculated that for any positive integer k, the following relationship holds true

$$A^{k-1}r_0 = \lambda_1^{k-1}z_1 + \lambda_{s_1+1}^{k-1}z_2 + \cdots + \lambda_{s_1+\cdots+s_{t-1}+1}^{k-1}z_t, \tag{30}$$

The above equation shows that the vector group $r_0, Ar_0, \cdots, A^{k-1}r_0$ can be linearly represented by the vector group $z_1, z_2, \cdots, z_t$ and its rank is not exceeding t. Thus,

$$\dim \mathcal{K}(A, r_0, k) = \dim\left(\operatorname{span}\{r_0, Ar_0, \cdots, A^{k-1}r_0\}\right) \leq t. \tag{31}$$

By lemma 1, the CG method can achieve the exact solution of the linear system $Ax = b$ after t iterations at most.

Simply modify the above proof, the following corollary can be obtained.

Corollary 3. Assume $A \in R^{n\times n}$ is SPD, its eigenvalues are $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n > 0$, the corresponding orthonormal eigenvectors are $u_1, u_2, \cdots, u_n$. If the initial guess x_0 satisfies $b - Ax_0 \in \operatorname{span}\{u_{i_1}, u_{i_2}, \cdots, u_{i_s}\}$, and $\lambda_{i_1}, \lambda_{i_2}, \cdots, \lambda_{i_s}$ are t distinct eigenvalues of A. Then, the CG method for solving problem (1) can achieve the exact solution after t iterations at most.

5 CONCLUSIONS

Three efficient improvements on the convergence rate of CG method are presented in this paper. These improvements consider the effect of the right hand side vector on the convergence rate of the CG method. They can provide directions for constructing preconditioners during the process of typhoon forecasting and accelerate the process of typhoon forecasting.

COMPETING INTERESTS

The author has no relevant financial or non-financial interests to disclose.

FUNDING

This research was supported by Guangdong Basic and Applied Basic Research Foundation (Grant Number 2021A1515310003).

REFERENCES

[1] Xiao Z X, Wang Z Q, Luo X L, et al. Ensemble seasonal forecasting of typhoon frequency over the western north Pacific using multiple machine learning algorithms. Environmental Research Letters, 2024, 19(10): 104007.

[2] Liu C, Deng H Q, Qiu X X, et al. A new post-processing method for improving track and rainfall ensemble forecast for typhoons over eastern China. Atmoshpere, 2024, 15(8): 874.

[3] Wang H, Gao L, Zhu L, et al. Changes in the typhoon intensity under a warming climate: a numerical study of typhoon mangkhut. Journal of climate, 2024, 37(14): 3703-3722.

[4] Gordon D. Conjugate gradient acceleration of coordinate descent for linear systems. Journal of Scientific Computing, 2023, 96(3): 86.

[5] Shan X J, Van Gijzen M B. Deflated preconditioned conjugate gradient methods for noise filtering of low-field MR Images. Journal of Computational and Applied Mathematics, 2022, 400: 113730.

[6] Xia J L. Robust and effective eSIF preconditioning for general dense SPD matrices. SIAM Journal on Scientific Computing, 2021, 43(5): S767-S790.

[7] Agullo E, Darve E, Giraud L, et al. Low-rank factorizations in data sparse hierarchical algorithms for preconditioning symmetric positive definite matrices. SIAM Journal on Matrix Analysis and Applications, 2018, 39(4): 1701-1725.

[8] Li R, Wang Z Q. Restrictively preconditioned conjugate gradient method for a series of constantly augmented least squares. SIAM Journal on Matrix Analysis and Applications, 2020, 41(2): 838-851.

[9] Golub G H, Van Loan C F. Matrix Computations. 3rd ed. Baltimore: The Johns Hopkins University Press, 1996.

[10] Xu S F, Gao L, Zhang P W. Numerical Linear Algebra. 2nd ed. Beijing: Peking University Press, 2013.