Re-implementation of CG vs MINRES: An Empirical Comparison

Xingri Geng

October 15, 2024

Abstract

This report is devoted to the re-implementation of the methods discussed in the paper "CG vs MINRES: An Empirical Comparison." We detail our approach, the experiments performed, and the insights gained through this process.

1 Introduction

The paper "CG vs MINRES: An Empirical Comparison", presents a comparative analysis of two Krylov subspace methods, namely, Conjugate Gradient (CG) and Minimum Residual (MINRES), which are used for iterative solutions of symmetric linear systems $Ax = b$.

CG is widely used for symmetric positive-definite (SPD) systems, while MINRES is applied to indefinite systems. The paper also shows the preference of MINRES to CG for SPD systems if early termination of the iterative process is required. Particularly, it focus on three norms for both methods and shows that:

- The approximate solutions $||x_k||$ are monotonically increasing for both CG and MINRES when A is SPD.
- The errors $||x^* - x_k||$ decrease monotonically for MINRES.
- The backward errors, that measure how closely the current iterate satisfies a perturbed system, decrease monotonically for MINRES but not always for CG, which makes MINRES more stable in practice for problems requiring less precision.

This report focus on the re-implementation of the original work by Fong and Saunders, aiming to replicate and verify the results discussed in the paper with another programming language *Python*.

1.1 Objective

The objective of this work is to implement the numerical experiments using *Python* and validate the convergence behavior of CG and MINRES across various test matrices.

1.2 Outline

The report is structured as follows: Section 2 discusses the methodology; Section 3 presents the experimental results; Section 4 provides a discussion on the findings.

2 Methodology

In this section, we describe the algorithms and experimental setup used for the re-implementation. The original paper's focus was on comparing CG and MINRES based on their convergence properties.

2.1 Algorithm Overview

The CG and MINRES algorithms were implemented to solve the following linear system:

$$Ax = b$$

where A is a symmetric positive-definite matrix for CG and possibly indefinite for MINRES.

We also solve an indefinite system, illustrating how MINRES behaves with such matrices. The example system is:

$$\begin{pmatrix} 2 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}. \tag{1}$$

Moreover, the indefinite systems are also considered as

$$(A - \delta I)x = b,$$

where A is SPD and $\delta = 0.5$.

3 Experimental Results

This section shows the results obtained from re-implementing the numerical experiments described in the original paper. Each experiment reproduces the results for specific matrices using either the CG or MINRES algorithm.

The performance of these algorithms is evaluated based on convergence behavior, iteration count, residual norms. The implementation and additional resources are available at $https://github.com/gengxingri/CG-vs-MINRES$.

3.1 Experiment 1: Figure 4.1

In this experiment, we analyze the distribution of condition numbers for various matrices before and after diagonal preconditioning. The matrices are from *the University of Florida Sparse Matrix Collection*, and our goal is to illustrate the effect of preconditioning on convergence.

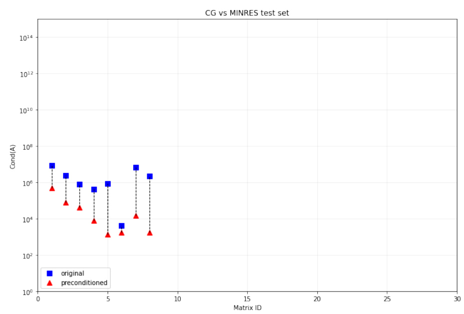
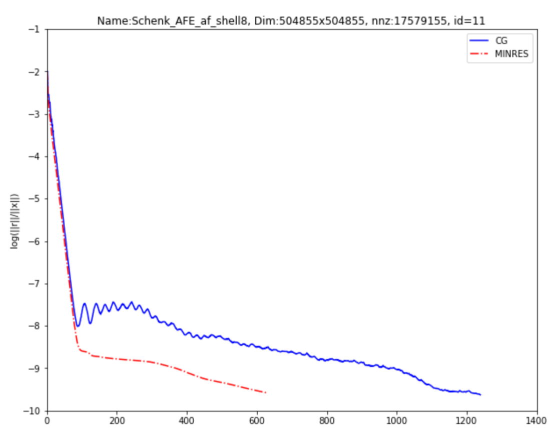


FIG. 4.1: Distribution of condition number for matrices used for CG vs MINRES comparison, before and after diagonal preconditioning

FIG. 4.1. is different from the original paper "CG VS MINRES: An Empirical Comparison" because the author of this report can not find the data about Matrix ID $1-26$. Following the reference of the original paper, the author can only find 8 symmetric positive-definite matrices among 26 matrices. However, if the 26 matrices are provided, the same codes can be also used to plot FIG. 4.1.

3.2 Experiment 2: Figures 4.2

FIG. 4.2. is the same as the graphs listed in the original paper. The first experiment reproduces the results using the Schenk_AFE_af_shell8 matrix, which has dimensions 504855×504855. The convergence behavior of CG and MINRES is compared. We examine both residual norms and error norms.



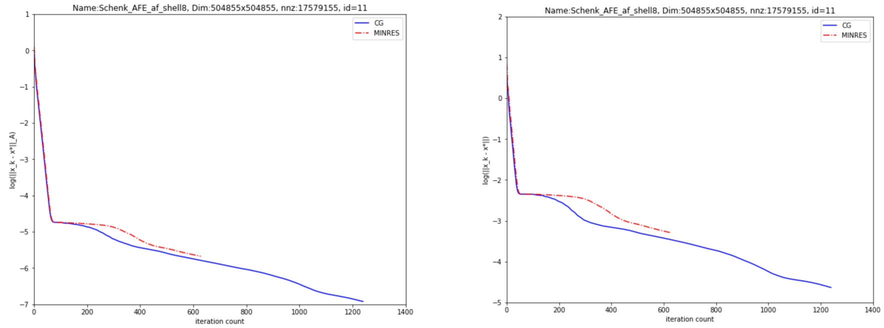


FIG. 4.2. Left: Schenk_AFE_af_shell8

Next experiment reproduces the results using the Cannizzo_sts4098 matrix, which has dimensions 4098×4098.

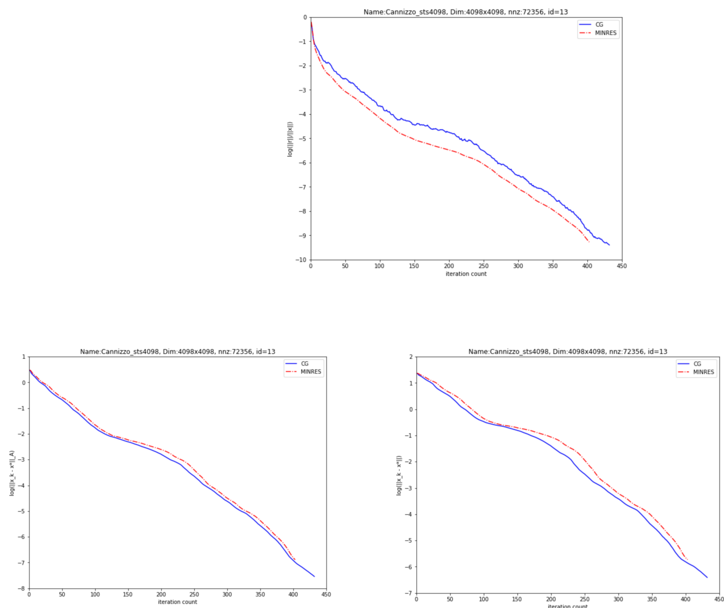
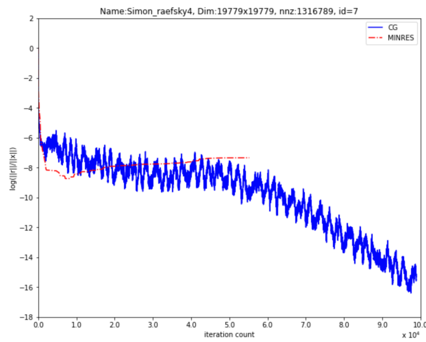


FIG. 4.2 Right: Cannizzo_sts4098

3.3 Experiment 3: Figures 4.3

This experiment uses the Simon_raefsky4 matrix, which has dimensions 19779×19779 and the BenElechi_BenElechi1 matrix, which has dimensions 245874×245874. We compare the backward and forward error for CG and MINRES, along with the solution norms across iterations.



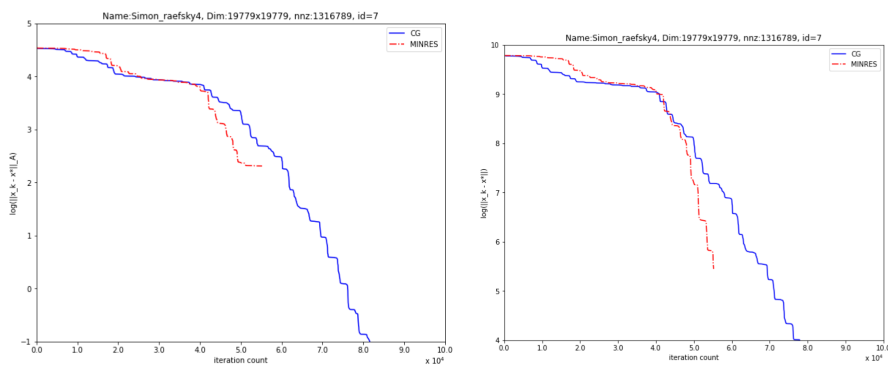
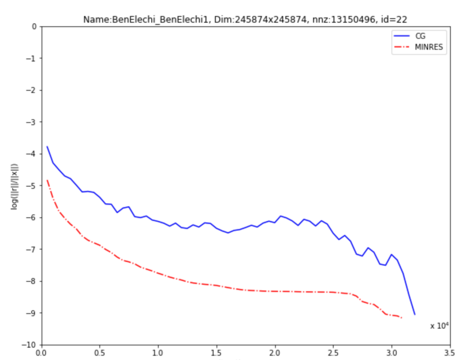


FIG. 4.3 Left: Simon_raefsky4



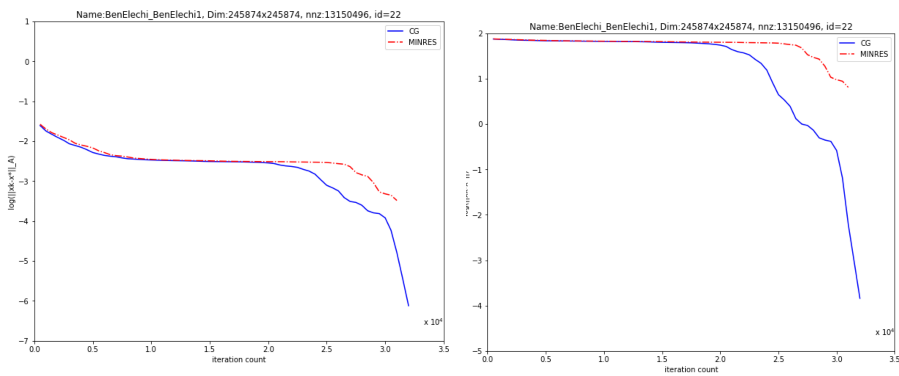


FIG. 4.3 Right: BenElechi_BenElechi1

FIG. 4.3. here is almost the same as FIG. 4.3. listed as in the original paper except that our iterations by MINRES are fewer than Fong and Saunders's work. These iterations can increase if we improve the tolerance in codes, in which case, more time will be needed.

3.4 Experiment 4: Figures 4.4

FIG. 4.4. is the same as the graphs listed in the original paper.

This experiment reproduces the results using the Simon_olafu_af matrix, which has dimensions 16146×16146 and the Simon_olafu_af matrix, which has dimensions 16146×16146.

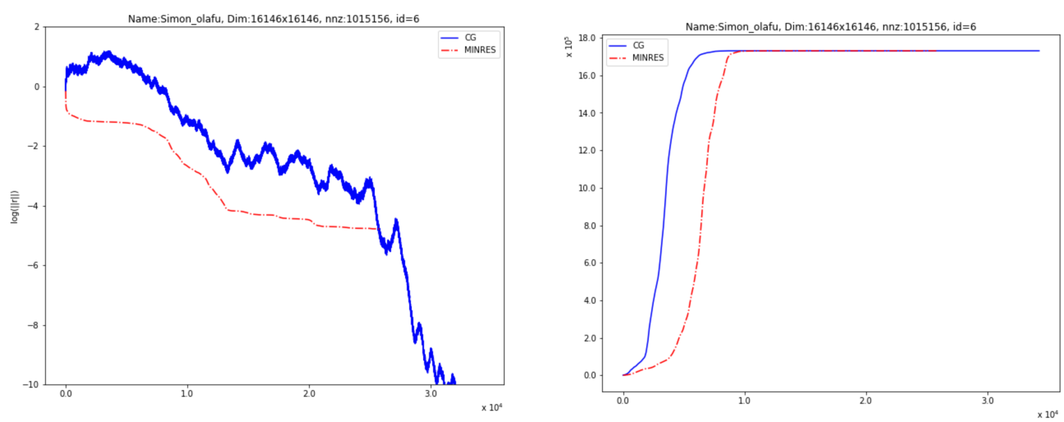


FIG. 4.4 Left: Simon_olafu

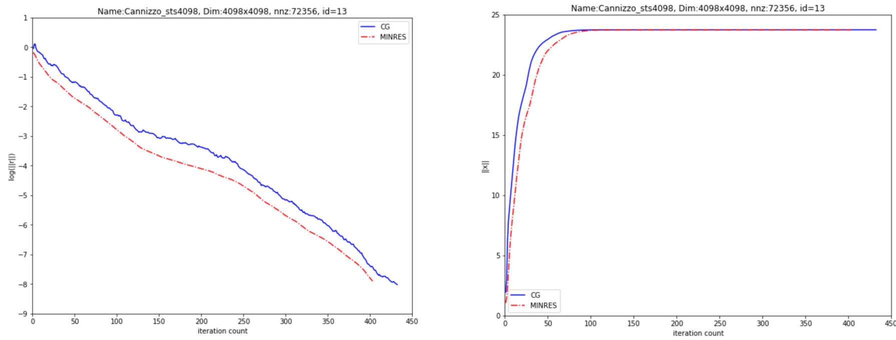


FIG. 4.4 Right: Cannizzo_sts4098

3.5 Experiment 5: Figures 4.5

FIG. 4.5. here is the same as the graphs listed in the original paper.

This experiment uses the Schmid_thermal1 matrix with size 19779×19779 to explore convergence behavior with a high condition number.

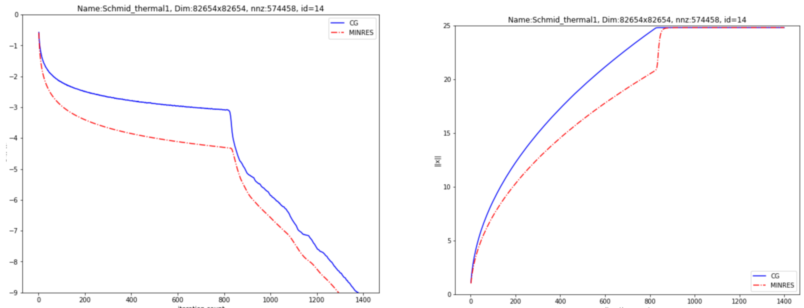


FIG. 4.5 Left: Schmid_thermal1

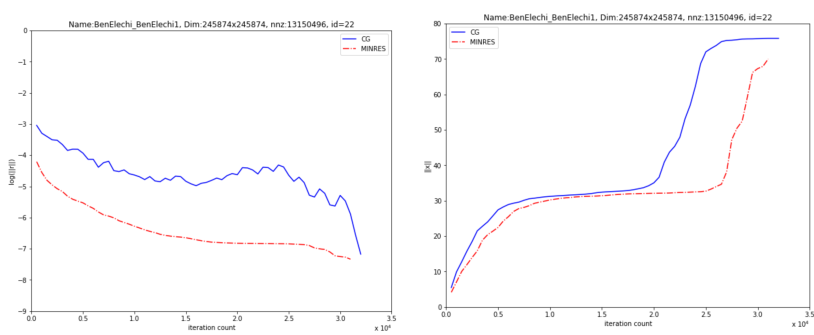


FIG. 4.5 Right: BenElechi_BenElechi1

3.6 Experiment 6: Figure 4.6

FIG. 4.6. here is the same as the graphs listed in the original paper.

This experiment uses the matrix according to (1).

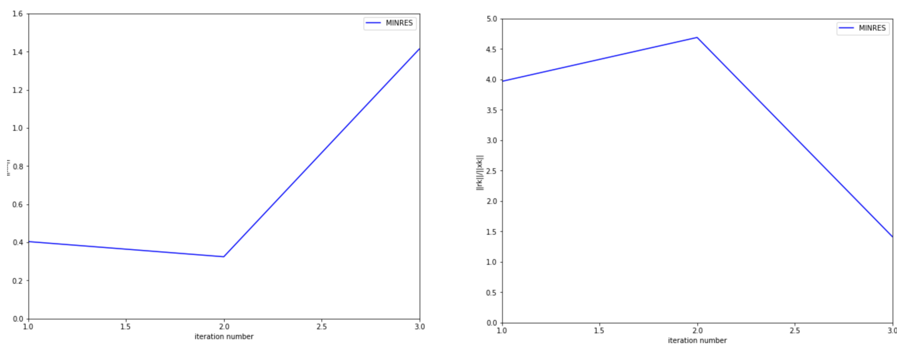


FIG. 4.6 MINRES on the indefinite problem (4.2)

3.7 Experiment 7: Figure 4.7

FIG. 4.7. here is the same as the graphs listed in the original paper.

The first experiment uses the Simon_olafu matrix with size 19779×19779.

The second experiment uses the Cannizzo_sts4098 matrix with size 4098×4098.

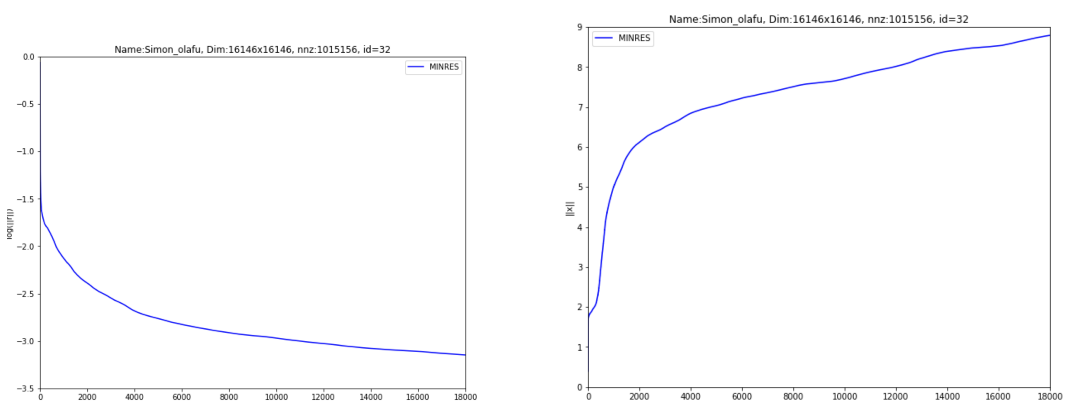


FIG. 4.7 Left: Simon_olafu

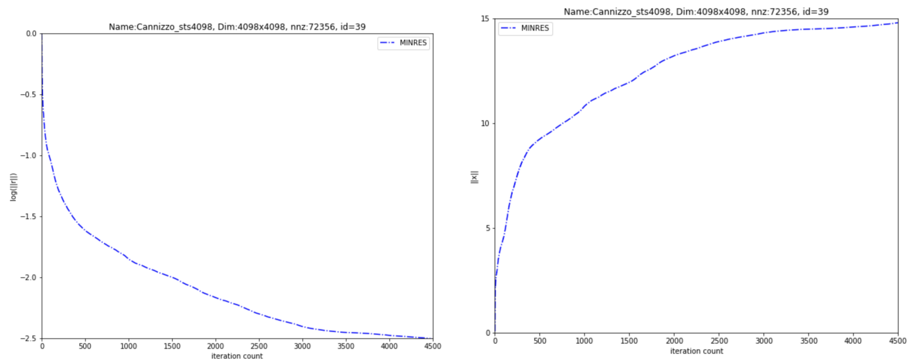


FIG. 4.7 Right: Cannizzo_sts4098

3.8 Experiment 8: Figure 4.8

FIG. 4.8. Left here is the same as the graphs listed in the original paper.

FIG. 4. 8. Right is slightly different form the original paper because the size of the given matrix and the corresponding condition number is drastically large, leading to computational burden. In our case, we plot the graph by taking solutions in each 500 iterations. Moreover, the iterations in FIG. 4.8. Right is less than those in the original paper which can be improved by changing the tolerance in our codes with more time cost.

This experiment uses the Schmid_thermal1 matrix with size 82654×82654 and the BenElechi_BenElechi1 matrix with size 245874×245874.

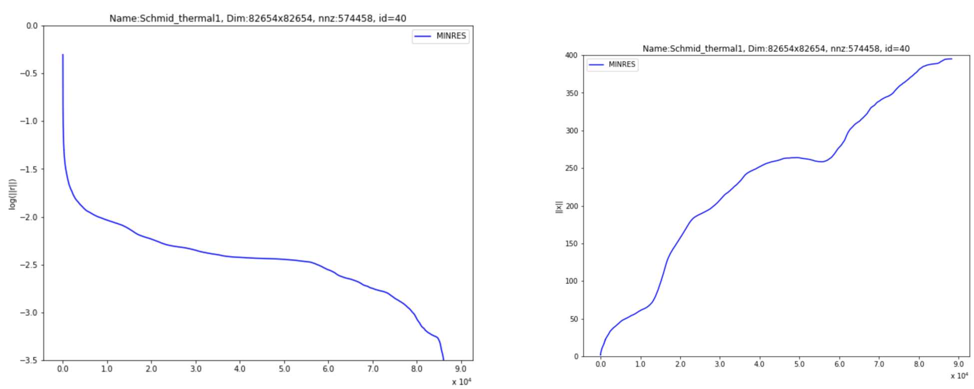


FIG. 4.8 Left: Schmid_thermal1

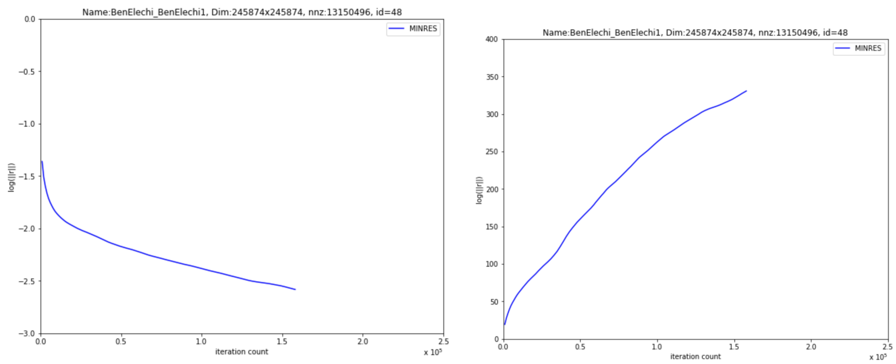


FIG. 4.8 Right: BenElechi_BenElenchi1

4 Discussion

The above experiments confirm the theoretical results outlined in the original paper. MINRES exhibits faster convergence than CG for positive-definite matrices. Moreover, MINRES demonstrates robustness in indefinite cases. This rethinking opens up potential applications where MINRES can outperform CG.

5 Conclusion

The re-implementation validates the original findings and provides additional insights into the behavior of CG and MINRES algorithms. These methods complement each other well, with MINRES offering advantages in cases where early termination is required.