Shift-Invert Lanczos Algorithm for Generalized Eigenvalue Problems

by

Ayobami Adebesin

Under the Direction of Michael Stewart, Ph.D.

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

Master of Science

in the College of Arts and Sciences

Georgia State University

2025

ABSTRACT

This thesis presents numerical results for the shift-invert lanczos algorithm applied to a generalized eigenvalue problem involving a dense matrix and analyzing the accuracy of the results for eigenvalues near and farther away from the shift.

INDEX WORDS: eigenvalues, eigenvectors, lanczos algorithm, ritz values, krylov subspaces, orthogonality

Shift-Invert Lanczos Algorithm for Generalized Eigenvalue Problems

by

Ayobami Adebesin

Committee Chair:	Michael Stewart
Committee:	Committee Member
	Committee Member
	Committee Member
	Committee Member

Electronic Version Approved:

Office of Graduate Studies

College of Arts and Sciences

Georgia State University

May 2025

DEDICATION

I dedicate this project to God.

ACKNOWLEDGMENTS

Again, this page is optional but also highly recommended. You may use this section to express acknowledgement of those who have helped you with this document (your committee members and instructors) and your academic career. This section should be double spaced.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1
INTRODUCTION

1.1 Background

The problem of computing eigenvalues and eigenvectors of matrices in numerical linear algebra is a well-studied one. The computation of eigenvalues and eigenvectors plays a central role in scientific computing with applications in structural analysis, quantum mechanics, data science and control theory. However, eigenvalue problems(standard and generalized) involving dense and sparse matrices present significant computational challenges, especially as the size of the matrices increases. These problems are fundamental in many scientific and engineering disciplines where the underlying mathematical models are often expressed in terms of eigenvalue equations. Historically, methods for solving eigenvalue problems date back to the early 20th century with foundational contributions from David Hilbert, Erhard Schmidt, and John von Neumann, who laid the groundwork for understanding linear operators and their spectral properties.

With the advent of digital computing in the mid-20th century, numerical methods for eigenvalue problems began to flourish. Classical iterative methods, such as the power iteration and inverse iteration, were among the first to be employed due to their simplicity and effectiveness for small-scale problems. However, as computational requirements grew, particularly with the need to solve larger sparse systems, researchers turned to more sophisticated algorithms. The Lanczos method, introduced by Cornelius Lanczos in 1950, represented a significant advancement for efficiently solving eigenvalue problems for large symmetric ma-

trices. The method exploits the sparsity of matrices and reduces the dimensionality of the problem by constructing a tridiagonal matrix whose eigenvalues approximate those of the original matrix.

1.2 Mathematical Preliminaries

In this section, we attempt to introduce some notations and the key mathematical concepts underlying the eigenvalue problems that will be used in this study .

1.2.1 Notation

Throughout this study, we make use of the following notations:

$A \in \mathbb{C}^{m \times n}$: denotes square or rectangular matrices

$Q \in \mathbb{C}^{m \times m}$: denotes unitary or orthogonal matrices

a_{ij} : denotes element (i, j) of A

$x, q \in \mathbb{C}^m$: denotes column vectors

Greek letters $\alpha, \beta...$: denotes scalars in $\mathbb{C}$

A^T : denotes the transpose of matrix A

$\| \cdot \|$: denotes a vector or matrix norm

$\otimes$: denotes the Kronecker product of two matrices

$A_{i:i',j:j'}$: denotes the $(i' - i + 1) \times (j' - j + 1)$ submatrix of A

$A^{(k)}$: denotes the matrix A at the kth step of an iteration

1.2.2 Floating Point Arithmetic

We define a *floating point* number system, $\mathbf{F}$ as a bounded subset of the real numbers $\mathbb{R}$, such that the elements of $\mathbf{F}$ are the number 0 together with all numbers of the form

$$x = \pm(m/\beta^t)\beta^e$$

where m is an integer in the range $1 \leq m \leq \beta^t$ known as the significand, $\beta \geq 2$ is known as the *base* or *radix* (typically 2), e is an arbitrary integer known as the exponent and $t \geq 1$ is known as the precision.

To ensure that a nonzero element $x \in \mathrm{F}$ is unique, we can restrict the range of $\mathbf{F}$ to $\beta^{t-1} \leq m \leq \beta^t - 1$. The quantity $\pm(m/\beta^t)$ is then known as the *fraction* or *mantissa* of x. We define the number $u \coloneqq \frac{1}{2}\beta^{1-t}$ as the *unit roundoff* or *machine epsilon*. In a relative sense, the *unit roundoff* is as large as the gaps between floating point numbers get.

Let $fl : \mathbb{R} \to \mathbf{F}$ be a function that gives the closest floating point approximation to a real number, then the following theorem gives a property of the unit roundoff.

Theorem 1.2.1. *If $x \in \mathbb{R}$ is in the range of $\mathbf{F}$, then $\exists\ \epsilon$ with $|\epsilon| \leq u$ such that $fl(x) = x(1+\epsilon)$.*

One way we could think of this is that, the difference between a real number and its closest floating point approximation is always smaller than u in relative terms.

1.2.3 Vector Norms

Norms are generally used to capture the notions of size and distance in a vector space. A norm is a function $\|\cdot\| : \mathbb{C}^m \to \mathbb{R}$ satisfying the following properties $\forall$ vectors x and y and scalars $\alpha \in \mathbb{C}$:

$$(1)\ \|x\| \geq 0, \text{ and } \|x\| = 0 \text{ only if } x = 0,$$

$$(2)\ \|x + y\| \leq \|x\| + \|y\|,$$

$$(3)\ \|\alpha x\| = |\alpha|\|x\|$$

The most important class of vector norms are the p-norms and are defined as follows:

$$\|x\|_1 = \sum_{i=1}^{m} |x_i| = 1,$$

$$\|x\|_2 = \Big(\sum_{i=1}^{m} |x_i|^2\Big)^{1/2},$$

$$\|x\|_\infty = \max_{1 \leq i \leq m} |x_i|,$$

$$\|x\|_p = \Big(\sum_{i=1}^{m} |x_i|^p\Big)^{1/p}, \qquad (1 \leq p < \infty)$$

1.2.4 General Matrix Norms

Similar to a vector norm, a matrix norm is a function $\|\cdot\| : \mathbb{C}^{m\times n} \to \mathbb{R}$ satisfying the following properties $\forall$ matrices A and B and scalars $\alpha \in \mathbb{C}$:

$$(1)\ \|A\| \geq 0, \text{ and } \|A\| = 0 \text{ only if } A = 0,$$

$$(2)\ \|A + B\| \leq \|A\| + \|B\|,$$

$$(3)\ \|\alpha A\| = |\alpha|\|A\|$$

1.2.5 Conditioning

1.2.6 Forward and Backward Errors

1.2.7 The Standard Eigenvalue Problem

Let $A \in \mathbb{C}^{m\times m}$ be a square matrix. A nonzero vector $x \in \mathbb{C}^m$ is an *eigenvector* of A, and $\lambda \in \mathbb{C}$ is the corresponding *eigenvalue* if,

$$Ax = \lambda x \tag{1.1}$$

The (multi)set of all eigenvalues of A is called the *spectrum* of A and is denoted by $spec(A)$. Equation (1.1) can be written as $(A - \lambda I)x = 0$. Since $x \neq 0$, this is equivalent to $\det(A - \lambda I) = 0$ and the solutions are the roots of the degree n polynomial

$$p_A(\lambda) = det(\lambda I - A) \tag{1.2}$$

known as the *characteristic polynomial* of A.

We define the eigenspace, E_λ of A corresponding to an eigenvalue $\lambda \in spec(A)$ as the set of all eigenvectors associated with λ as follows:

$$E_\lambda = \mathcal{N}(A - \lambda I) = \{x \mid (A - \lambda I)x = 0\} \tag{1.3}$$

If A is Hermittian, the eigenvalues of A are real and the

1.2.8 The Generalized Eigenvalue Problem

1.2.9 Lanczos Algorithm

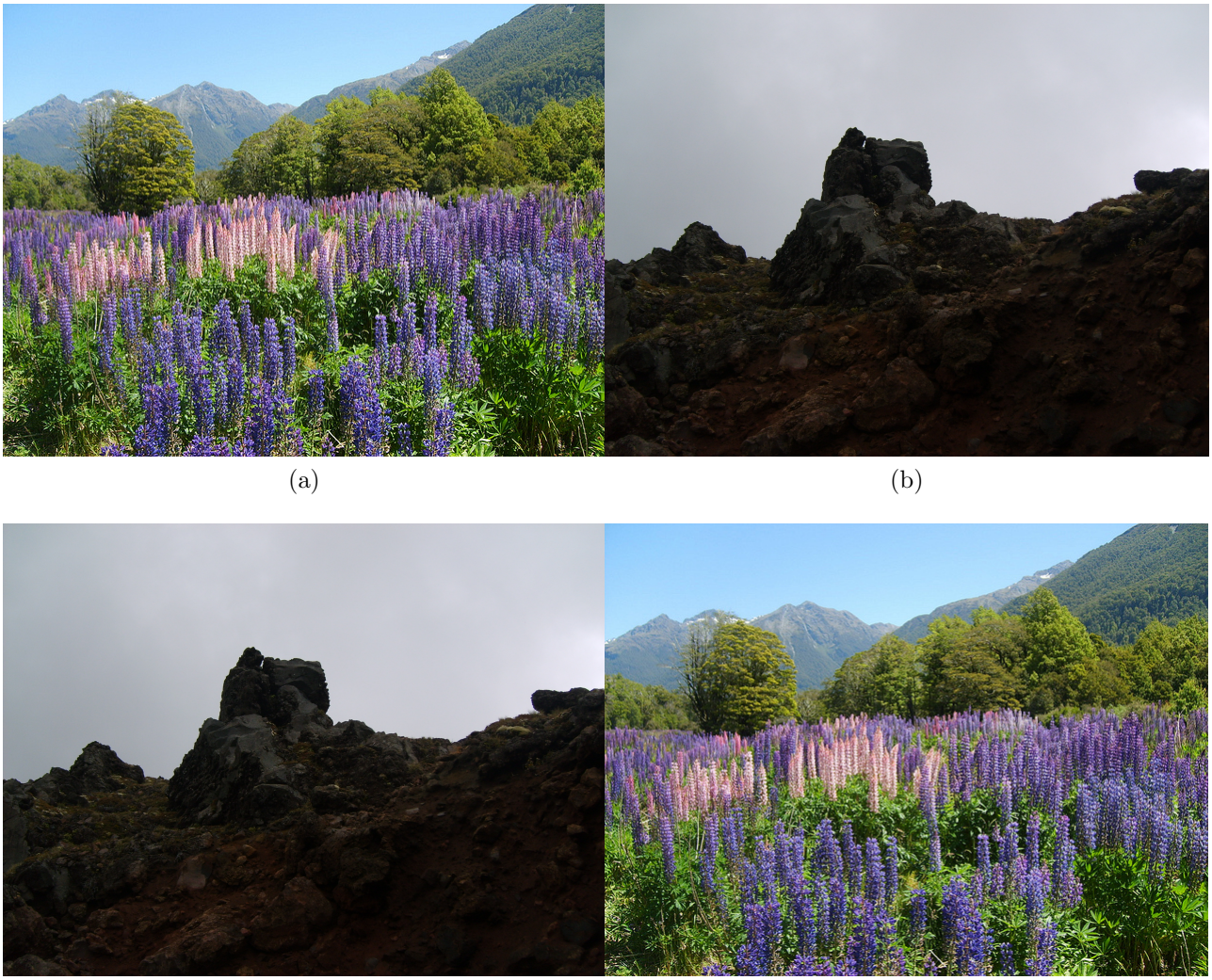
CHAPTER 2
CHAPTER 2 TITLE

Hooray for Chapter 2!!!

Sample Figures and Tables below.

And as an example of citing things, I'm going to cite a brilliant paper - (Jones et al. 2016). See bibliography.bib for doing references.

Figure 2.1 An individual figure!



(a) (b)

(c) (d)

Figure 2.2 Multiple figures!

Table 2.1: Insert Table Caption here

Blah	Blah	Blah
Stuff	Things	etc.
...	...	...

Appendices

A Something

This is the appendix!

B Something Else

Another appendix!

REFERENCES

Jones, J., White, R. J., Quinn, S., Ireland, M., Boyajian, T., Schaefer, G., & Baines, E. K. 2016, ApJ, 822, L3