COMPONENTWISE EIGENVALUE CONDITION NUMBERS FOR DIFFERENT REPRESENTATIONS OF TRIDIAGONAL AND BANDED MATRICES

September 10, 2014

MARTA GÓMEZ SANCHO

UNDERGRADUATE THESIS PROJECT

DEGREE IN INDUSTRIAL TECHNOLOGIES ENGINEERING

ADVISOR: Dr. FROILÁN MARTÍNEZ DOPICO

SUBJECT: NUMERICAL LINEAR ALGEBRA

DEPARTMENT OF MATHEMATICS

CARLOS III UNIVERSITY

Abstract

Sensitivity analysis is a tool for the validation of the quality of computed eigenvalues, having wide ranges of applications in natural and applied science. For modelling strenght, elasticity, deformations, PID controller (Proportional-Integral-Derivative controller) and many other problems, there is a need for computing reliably eigenvalues. Componentwise relative perturbation analysis can provide good estimations of the accuracy of these eigenvalues. In this undergraduate thesis project we provide an expression for the 2-norm of an special vector. This vector is given by $relgrad(\lambda) = \left(\frac{p_j}{\lambda}\frac{\partial\lambda}{\partial p_j}\right)$ where p_j are the entries of certain representations of tridiagonal matrices. It occurs that $relgrad(\lambda)$ determines the relative componentwise eigenvalue condition number of a tridiagonal matrix (unsymmetric , unreduced, real tridiagonal) and its factored forms (with parameters $\mathfrak{l}_j, u_j, l_j, \ldots$ instead of p_j). This allows us to derive the expressions of the eigenvalue condition numbers in the 2-norm, $relcond_2(\lambda)$, and compare the entrywise tridiagonal with the factored form eigenvalue condition number . The derivations of the condition numbers were already shown in the reference [12], in which the authors use only 1-norm for the computations . We also obtain expressions for relative componentwise eigenvalue condition numbers of banded matrices.

Abstract

El análisis de sensibilidad es una herramienta para la validación de la calidad de los autovalores computados. Tiene amplios rangos de aplicación en ciencias naturales y aplicadas. Para modelar fuerzas, elasticidad, deformaciones, controladores PDI (Proportional-Integral-Derivative controller) y otros muchos parámetros, hay una necesidad de confiabilidad en el cómputo de autovalores. El análisis de perturbación por componentes relativo puede proveer buenas estimaciones de la precisión de esos autovalores. En este proyecto fin de grado aportamos una expresión para la 2-norma de un vector especial. Este vector viene dado por $relgrad(\lambda) = \left(\frac{p_j}{\lambda}\frac{\partial\lambda}{\partial p_j}\right)$ donde p_j son las entradas de ciertas representaciones de una matriz tridiagonal. Ocurre que $relgrad(\lambda)$ determina el numero de condición relativo , por componentes, de una matriz tridiagonal (no simétrica , irreducible y real) y sus formas factorizadas (con parámetros $\mathfrak{l}_j, u_j, l_j, \ldots$ en vez de p_j). Esto nos permite derivar la expresión de los números de condición usando la 2-norma, $relcond_2(\lambda)$, y comparar los numeros de condición de autovalores por componentes de las matrices tridiagonales con los de las formas factorizadas. La obtención de los números de condición se puede ver en la referencia [12], en el cual los autores usan sólo 1-norma para los cálculos. También hemos obtenido expresiones para el número de condición relativo por componentes de matrices de banda.

Contents

1 Introduction **6**

2 Structure of the undergraduate thesis **8**

3 A note on applications of eigenvalue sensitivity analysis **10**

4 Basic definitions **11**
4.1 Eigenvalues, eigenvectors 11
4.2 Vector and matrix norms 12
4.3 Algebraic and geometric multiplicities 13
4.4 Similarity transformations 14
4.5 Accuracy analysis: Floating point arithmetic, Sensitivity and Stability 14
4.5.1 Floating point arithmetic 14
4.5.2 Floating point operations 16

5 Notation and additional concepts **16**
5.1 Summary of notation 17

6 Tridiagonal and banded matrices **18**
6.1 Balancing 19

7 Factored forms of tridiagonal matrices **19**

8 LU factorization of banded matrices **21**

9 Gradients and condition numbers of eigenvalues of tridiagonal matrices **21**
9.1 Sensitivity analysis 22
9.2 Representation 1 - entries of C 24
9.3 Representation 2 - $\mathcal{L}, \mathcal{U}$ representation of J 26
9.4 Other representations 28
9.4.1 Representation 3 - L, D (Δ) representation of ΔT 28
9.4.2 Representation 4 - $\bar{L}, \Omega$ representation of T 30
9.5 Equivalence of $relcond_2(\lambda; \mathcal{L}, \mathcal{U})$, $relcond_2(\lambda; L, D)$, $relcond_2(\lambda; \bar{L})$ 32

10 Sensitivity analysis for banded matrices **35**
10.1 Condition numbers for banded matrices 37
10.1.1 Representation 2 - LU factorization 38

11 Conclusions and future work **41**

References **42**

1 Introduction

Eigenvalues and eigenvectors of matrices arise in many applications. Vibration problems in classical mechanics, acoustics, computation of energy levels in quantum mechanics, linear stability of flows in fluid mechanics, and the computation of the Google page-rank vector are just a few among many applications where eigenvalues and eigenvectors play an essential role [37]. On the other hand, it is not possible to compute the eigenvalues of an $n \times n$ matrix A through simple formulas, since the eigenvalues are the roots of the characteristic polynomial of A and Abel proved in 1824 that for $n \geq 5$ the roots of a polynomial cannot be expressed in terms of additions, subtractions, multiplications, quotients, and kth roots of the coefficients of the polynomial [38, p. 192]. Therefore, the computation of eigenvalues of matrices must be performed numerically via iterative methods implemented in a computer.

The development of numerical algorithms for computing eigenvalues of matrices has been, and still is, one of the most active areas of research inside Scientific Computing since modern digital electronic computers were invented in the late 1940s. One of the reasons of this intense activity is that this problem is extremely difficult, because the naive approach of computing first the characteristic polynomial of the matrix and, then, to compute its roots by using any numerical method for computing solutions of nonlinear equations is *unstable* [38, p. 92]. In fact, essentially none of the numerical algorithms for computing eigenvalues that were used before modern computers appeared has been implemented in a computer. The reason is simple: those old algorithms are unstable in floating point arithmetic.

As a consequence of the intense research effort performed during the years 1950-1965 for developing numerical algorithms for computing eigenvalues of matrices, several reliable methods were proposed and the best ones are thoroughly analyzed in the monumental treatise published by Wilkinson in 1965 [39]. Among all the methods presented in this book one has become the "superstar" numerical matrix eigenvalue algorithm: the *Francis QR eigenvalue algorithm*. This is invariably the method implemented in any professional software for computing *all* eigenvalues of an $n \times n$ matrix A and it is explained in all books on *Numerical Linear Algebra* and *Matrix Computations* (see for instance [14, 38] and the references therein). The Francis QR eigenvalue algorithm is an extremely sophisticated method that has been polished and improved over decades for becoming today a fully reliable a very efficient procedure. In fact, it was selected as one of the top-10 algorithms of the whole 20th Century [8]. Very recent techniques that have improved considerably the performance of the Francis QR eigenvalue algorithm can be found in [2, 3, 13] and the references therein.

The two most prominent features of the Francis QR eigenvalue algorithm for computing all the eigenvalues of an $n \times n$ matrix A are [14, Chapter 7]:

(a) It requires a computational cost of $O(n^3)$ flops (i.e., floating point operations) and $O(n^2)$ storage.

(b) It is *backward stable*. This means that the eigenvalues computed by the algorithm are the exact eigenvalues of a nearby matrix $A + E$, where in double precision IEEE arithmetic $\|E\|_2 = O(10^{-16})\|A\|_2$ and $\|\cdot\|_2$ is the spectral or 2-norm of a matrix [16, Chapter 6].

These two properties imply that, essentially, the Francis QR eigenvalue algorithm reaches the best limits of any possible eigenvalue algorithm since the cost of multiplying two $n \times n$ real matrices is $2n^3$ flops, to storage a matrix on a computer requires n^2 floating point numbers, and backward stability is the most that can be expected from any computation implemented on a computer subjected to roundoff errors.

Taking into account the previous discussion, it might seem that the Francis QR eigenvalue algorithm has solved completely the problem of computing all the eigenvalues of an $n \times n$ matrix.

However, this is not the case, since the computational cost and the storage requirements of the Francis QR eigenvalue algorithm limit its use to matrices of sizes $n \leq 20000$ in modern computers (approximately, since the exact bound depends on each particular computer) but in many present applications arise much larger matrices, which in most cases have many of its entries equal to zero, that is, they are *sparse matrices*. This has motivated the development of other types of eigenvalue algorithms for large sparse matrices, which are more efficient from the points of view of number of operations and storage requirements but are not guaranteed *a priori* to be backward stable.

Loosely speaking modern methods for computing eigenvalues of large sparse matrices can be classified into two main classes: (1) projection methods that project the large problem into a much smaller one whose eigenvalues are approximations of just a few eigenvalues of the original matrix A [34]; and (2) methods that compute *all* eigenvalues of matrices that have very particular and simple structures [14]. This undergraduate thesis project is related with the second type of methods, more precisely, with the computation of all eigenvalues of an *unsymmetric tridiagonal matrix*.

An $n \times n$ matrix A is said to be tridiagonal if its entries a_{ij} satisfy $a_{ij} = 0$ whenever $|i - j| > 1$. Therefore among the n^2 entries of A at most $3n - 2$ are nonzero, so when n is large we can say that most of the entries of A are zero. This property allows us to expect that all the eigenvalues of A can be computed with a computational cost of $O(n^2)$ flops (instead of $O(n^3)$ as in Francis QR algorithm) and by storing just $O(n)$ floating point numbers (instead of $O(n^2)$ as in Francis QR algorithm). If n is large ($n > 10^4$, for instance), these would be impressive improvements with respect to the computational cost and storage requirements of Francis QR algorithm. In fact, there are at least two algorithms that reach these goals for unsymmetric tridiagonal matrices by exploiting carefully this structure. One is based on the Ehrlich-Aberth iteration [1] and the other belongs to the family of *dqds* algorithms [11, 29]. However, none of these algorithms is guaranteed a priori to be backward stable, as neither is any other algorithm for the unsymmetric tridiagonal eigenvalue problem. In fact, to find a *backward stable* algorithm for computing all eigenvalues of an unsymmetric $n \times n$ tridiagonal matrix with $O(n^2)$ computational cost and $O(n)$ storage is a classical open problem in Matrix Computations, whose solution seems to be very far from the present state of the art in this field of research. This open problem is known as *the unsymmetric tridiagonal eigenvalue problem*.

In this context, since the stability of the currently available algorithms for the unsymmetric tridiagonal eigenvalue problem [1, 11] is not guaranteed a priori, it is fundamental to provide together with each computed eigenvalue a bound on its relative forward error that can be efficiently computed from the data of the problem and from the computed eigenvalues and eigenvectors. *This undergraduate thesis project is a contribution towards solving this challenging long-term project.*

Estimation of bounds on the forward errors of any computed magnitude in a reliable and sharp way is, in general, a very difficult task for any problem in Numerical Analysis. The obvious reason for this difficulty is that the exact value of the magnitude is not known. In Matrix Computations the estimation of bounds on the forward errors is traditionally obtained by computing the backward error from the residual of the computed magnitude and by multiplying this backward error by the condition number of the problem we want to solve. This is encoded is the famous rule of thumb [14] [16]:

$$\text{forward error} \leq \text{condition number} \times \text{backward error}. \tag{1}$$

The condition number of a certain magnitude is a measure of the maximum variation of that magnitude under perturbations of the input data, i.e., the condition number reflects the *sensitivity* of the magnitude under perturbations.

The use of the traditional eigenvalue condition number and the traditional method for computing eigenvalue backward errors [14] in the expression (1) for the eigenvalues provided by the currently available algorithms for computing eigenvalues of tridiagonal unsymmetric matrices in $O(n^2)$ flops and with $O(n)$ storage [1, 11] often leads to very pessimistic estimations of the forward errors of

the computed eigenvalues. This is particularly frequent for computed eigenvalues that are very tiny with respect the norm of the matrix. The main reason of this pessimistic overestimation is that the traditional eigenvalue condition number and the traditional method for computing eigenvalue backward errors do not take into account the tridiagonal structure that is carefully preserved in the algorithms in [1, 11]. To be more precise, for instance the traditional Wilkinson eigenvalue condition number presented in [14], and in many other standard books on Matrix Computations, measures the sensitivity of eigenvalues under *general unstructured perturbations* that destroy the tridiagonal structure of the input matrix. Therefore, a sharp estimation of the forward errors for the eigenvalues computed by the algorithms in [1, 11] (or by any other algorithm that takes advantage of the tridiagonal structure) requires first to develop formulas for eigenvalue condition numbers under perturbations that preserve the tridiagonal structure, to compute efficiently such condition numbers, and, finally, to estimate efficiently structured backward errors.

The study of eigenvalue condition numbers of tridiagonal matrices under perturbations that preserve this structure has been started very recently in the publication [12]. In this work, the authors have developed formulas for the structured eigenvalue condition numbers under perturbations of the parameters defining different representations of tridiagonal matrices that are used in different algorithms, but only when the size of the perturbations are measured via the *infinite norm* [16, Chapter 6] of the vector of relative changes of all parameters. In addition, the authors of [12] have shown that these eigenvalue condition numbers can be computed in $O(n)$ flops. However, current ongoing research projects (see [9, 31]) for estimating structured backward errors for eigenvalues of tridiagonal matrices have shown that it would lead to sharper forward error bounds to measure the size of the perturbations via the Euclidean 2-norm of the vector of relative changes of all parameters. Therefore, it is necessary to deduce formulas for the structured eigenvalue condition numbers of tridiagonal matrices under relative perturbations of the parameters measured via the Euclidean 2-norm of the vector of relative changes and also to develop efficient ways to compute them in $O(n)$ flops. *This is the main purpose of this undergraduate thesis project.* In addition, we will show that the techniques developed in this work and in [12] can be extended to the study of structured eigenvalue condition numbers of general low banded matrices.

In the next section, the structure of this undergraduate thesis project is discussed. The specific notations we use will be introduced in the places where they are used for first time. Here we only recall very basic mathematical notations that are employed throughout this thesis: $\mathbb{R}$ denotes the set of real numbers, $\mathbb{C}$ denotes the set of complex numbers, $\mathbb{R}^n$ denotes the set of column real vectors with n entries, $\mathbb{C}^n$ denotes the set of column complex vectors with n entries, $\mathbb{R}^{m\times n}$ denotes the set of real matrices with m rows and n columns, and $\mathbb{C}^{m\times n}$ denotes the set of complex matrices with m rows and n columns.

2 Structure of the undergraduate thesis

This undergraduate thesis extends in two directions some of the results presented in the paper of C. Ferreira, B. Parlett and F.M. Dopico cited in the reference [12]. The first extension consists in using the spectral norm of the vector of relative variations of the input parameters instead of the infinite norm, while the second extension considers general low-banded matrices instead of just tridiagonal matrices. The reader can find the main original contributions of this thesis in sections 9 and 10. In general terms, this document is organized in four parts: preliminaries, original contributions, general applications, and conclusions. The preliminaries are presented from section 1 to section 8 and include the introduction and a summary of important definitions and standard results that are needed to understand the original contributions of this work, which are presented in sections 9 and

10. Some general applications of sensitivity analysis of eigenvalues to Engineering are presented in Section 3, although we emphasize that, as explained in the introduction, the main applications of the original results we have developed are found in the estimation of errors of numerical algorithms. Finally, the conclusions and some lines of future work are presented in Section 11.

This undergraduate thesis has eleven sections, whose contents are now outlined.

1. Introduction. In this section the state of the art, origins, motivations, and applications for the problems studied in this undergraduate thesis are discussed.

2. Structure of the undergraduate thesis. Here, we explain the organization and structure of this document.

3. A note on applications of eigenvalue sensitivity analysis: This section discusses some general applications of sensitivity analysis of eigenvalues which are interesting in Engineering.

4. Basic definitions . This section is formed by 5 subsections. First, subsection 4.1 introduces the fundamental concepts, eigenvalues and eigenvectors, analyzed in this work. Subsection 4.2 is reserved for vector norms, which are fundamental for the definitions of the eigenvalue condition numbers studied in this undergraduate thesis. Then, we include two subsections for the definitions of algebraic and geometric multiplicities and similarity transformations. Finally, subsection 4.5 is devoted to the discussion of the main concepts of floating point arithmetic, rounding errors and their effect on magnitudes computed in a modern computer, paying special attention to the sensitivity of a problem and the stability of an algorithm. This section is based on the general references [4, 16, 17, 21, 26, 38].

5. Notation and additional concepts . In this section we introduce a few more advanced notions that are not usually found in standard textbooks on Matrix Computations. Moreover, we remind the classical Householder's notation for matrices and vectors, since it is the one employed in this work. We also list the principal classes of matrices and vectors appearing in this undergraduate thesis and the particular notation we use for them.

6. Tridiagonal and banded matrices. Here, we review some definitions and properties about this class of matrices in order to make easier the understanding of the subsequent sections. An extra subsection describes balanced tridiagonal matrices and another particular type of tridiagonal matrices that is useful in numerical computations.

7. Factored forms of tridiagonal matrices: The purpose of this section is to study those factorizations in which it is worth decomposing a tridiagonal matrix before trying to compute its eigenvalues. One of the factored forms is related to balanced matrices, which are matrices close to be symmetric in a certain sense. Another factored form is related to tridiagonal matrices that use the fewer possible parameters, denoted as J.

8 LU **factorization of banded matrices**. In this section, we present the main features of the standard LU factorization of banded matrices.

9. Gradients and condition numbers of eigenvalues of tridiagonal matrices. This section introduces formulas and efficient methods for the computation of the structured eigenvalue condition numbers of tridiagonal matrices using the spectral or 2-norm to measure the relative entrywise vari-

ation of the input parameters. This section is divided in three parts: the first one is devoted to the definition of the ECN (eigenvalue condition number) under perturbations of a set of parameters, the second one includes the derivation of ECNs for four different representations of tridiagonal matrices, and in the last part we prove certain equivalence relations between ECNs in different representations. For convenience, we always denote ECNs as $relcond_2(\lambda; T)$, where T is one (or two) square matrix (tridiagonal, banded or factored). It is clarifying to mention that section 9 will exclusively consider eigenvalue condition numbers for tridiagonal matrices with respect to the 2-norm of the vector of relative variations of the input parameters, while section 10 will consider general banded matrices but with respect the ∞-norm of the vector of relative variations. We emphasize that general banded matrices are not studied in [12].

10. Sensitivity analysis for banded matrices: Together with Section 9, this section form the core of this work. Here, the eigenvalue condition numbers $relcond(\lambda; C)$ and $relcond(\lambda; \mathcal{L}, \mathcal{U})$ are computed for banded matrices with respect the ∞-norm of the vector of relative variations of the input parameters.

11. Conclusions and future work: This section summarizes the main original contributions of this undergraduate thesis and proposes some lines of future work.

3 A note on applications of eigenvalue sensitivity analysis

There are some problems in engineering that require knowledge of the behaviour of eigenvalues under perturbation of data or discretization. Two cases are reviewed: application of eigenvalues in control engineering and eigenpairs in connectomics. In control engineering we may model the behaviour of a device, with an ODE (Ordinary Differential Equation), and it is important to linearize this equation. To reduce noise it is added a PID, that are also described by a mathematical model differential equation with properties to cancell noise errors. The ODE is called *Transfer Function* and is converted to the Laplace domain, where s is the unknown, to operate easily. A *system*, including more devices, thus, many transfer functions, may be simplified to a single transfer function by block operations or/and other methods. The transfer function is often a fraction, with polynomials in the numerator and the denominator,

$$H(s) = \frac{N(s)}{D(s)} = K\frac{(s-z_1),(s-z_2),\ldots,(s-z_m)}{(s-p_1),(s-p_2),\ldots,(s-p_n)}$$

The z_i's are the roots of the polinomial $N(s) = 0$, and are called *zeros* of the system.The p_i's are the roots of the polinomial $D(s) = 0$, and are called *poles* of the system.
The transfer function represent a system of differential equations and the homogeneous solution is defined by the poles.The homogeneous solutions are of the form:

$$y_h = \sum_{i=1}^{n} C_i e^{\lambda_i t}$$

where the constants C_i are determined by the initial conditions and λ_i are the poles, that is: $\lambda_i = p_i$ or the system **eigenvalues** (system of differential equations). To converge to a solution, as time goes to infinity (or before), we need that all λ lie in the negative semi-plane. That is, all λ must be

negative, so that, for increasing time, $t \to \infty$, the solution $e^{\lambda}t$ does not grow to infinity.
Finally, we give the last example on eigenvalue-eigenvectors applications. Connectomics is the science of analyzing and forming the map of neural connections in the brain. Recently, there has been a big effort to map the brain and their connections (synapses). Mapping tiny slices of the brain in a computer is a research area of connectommics. A tool for connectomics research is diffusion MRI (magnetic resonance imaging). "It allows the mapping of the difussion process of molecules, mainly water, in biological tissues" [40]. For example, if we want to estimate the axonal fiber orientation and in what proportions are each of the particular orientations in a volume element, we can modelize it by a problability function called fODF (fiber orientation density function). This fODF function may be approximated by a 3×3 matrix called Diffusion Tensor Imaging (DTI). DTI provides an "ellipsoid representation of the water-diffusion profile for a given voxel" [4] (voxel: volume element used in computer imaging - like pixel, voxel). Eigenvectors of matrix DTI "determine the direction of maximum and minimum water motion .The direction of the maximal diffusion is the best estimate of fiber orientation. Similarly, eigenvalues determine the amount of diffusion produced in each direction. Estimation of fODF at each voxel is the first step in estimating structural connectivity" (extracted from [4]).

4 Basic definitions

We set up standard notation and basic terminology. Bibliographic research has been done on [4], [16],[17], [21],[26], [38], to elaborate the text.

4.1 Eigenvalues, eigenvectors

Any linear transformation maps $\boldsymbol{x} \to \lambda \boldsymbol{x}$ for a finite set of vectors $\boldsymbol{x}$ called eigenvectors . This has a meaning : eigenvectors stay in the same line (direction) once the linear transformation is applied on them.
The following definitions are adapted to the conditions of this work : real matrices and eigenvalues of algebraic multiplicity one .

Definition 4.1. *In $\mathbb{C}^n$, two vectors $\boldsymbol{x} \neq 0$ and $\boldsymbol{y} \neq 0$ are said to be a right eigenvector and a left eigenvector of a matrix $C \in \mathbb{R}^{n \times n}$, respectively, if there exists an scalar λ , called eigenvalue of C, such that*

$$C\boldsymbol{x} = \lambda \boldsymbol{x} \text{ and } \quad \boldsymbol{y}^* C = \lambda \boldsymbol{y}^*$$

where $\boldsymbol{y}^* = \bar{\boldsymbol{y}}^T$, or the transpose of the conjugate of $\boldsymbol{y}$. The pair $(\lambda, \boldsymbol{x})$ is called a right eigenpair and, in the same line, $(\lambda, \boldsymbol{x}, \boldsymbol{y}^*)$ is known as an eigentriple of C.

For us, C and any other matrix that appear in this text, is a real matrix. However, in other contexts complex matrices may appear. A real matrix can have eigenpairs and eigenvectors that appear in conjugate pairs. One further definitions are useful,

Definition 4.2. *The spectrum of $C \in \mathbb{R}^{n \times n}$ is the set of all eigenvalues of C, i.e. the set of all the roots of the characteristic polynomial $P(\lambda) = det(C - \lambda I)$. The spectrum of C is denoted by $\Lambda(C)$.*

Spectrums of C and C^T should be the same in exact arithmetic but in floating point it turns out that they are distinct, due to machine arithmetic errors.

4.2 Vector and matrix norms

Vector norms are measurements of the length of a vector. We use them in perturbation theory for measuring perturbations' magnitudes and producing perturbation bounds (local bounds). However, by using normwise analysis we do not appreciate how sparse is a matrix or how large turn out to be part of its entries with respect to the others. This is the negative aspect of norms. This means that to have a more rigorous exam of the behaviour of eigenvalues we need to search for more detailed expressions, which include more information. These expressions are measurements in the components of the matrix, rather than a value that comprises all the elements of the matrix in one value, as the matrix and vector norm does. This is called componentwise analysis. For the purpose of rigurosity and to preserve the structure of tridiagonal matrices and banded matrices, in this project we work in componentwise analysis. In addition, we measure the relative perturbation vector of the input entries through a vector norm. In other traditional condition numbers, as Wilkinson condition number, matrix norms are key operators applied widely in normwise analysis for unstructured matrices. Loosely speaking, componentwise analysis is open to be more rigorous than absolute and to give another more detailed perspective. Usually, vectors and matrices are measured componentwise when the matrix is sparse. This affect directly to our work with tridiagonal matrices that, as we mentioned in the introduction, are sparse matrices. In this section vector and matrix norms are defined conveniently. In addition important bounds are described as well as some of the better advantages of the 2-norm.

Definition 4.3. *A matrix $Q \in \mathbb{C}^{n\times n}$ is said to be unitary if $Q * Q = I$. If, in addition, $Q \in \mathbb{R}^{n\times n}$, Q is said to be real orthogonal.*

Definition 4.4. *A vector norm is a function $\|.\| : \mathbb{C}^m \to \mathbb{R}$ that satisfy three properties, for all vectors $\boldsymbol{x}$ and $\boldsymbol{y} \in \mathbb{C}^m$ and for all scalars $\alpha \in \mathbb{C}$,*

$$1.\|\boldsymbol{x}\| \geq 0, \text{ and } \|\boldsymbol{x}\| = 0 \text{ only if } \boldsymbol{x} = 0$$
$$2.\|\boldsymbol{x} + \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|,$$
$$3.\|\alpha\boldsymbol{x}\| = |\alpha|\|\boldsymbol{x}\|.$$

A general vector norm which includes other norms as special cases is the p-norm,

$$\|\boldsymbol{y}\|_p = \left(\sum_{i=1}^{m} |y_i|^p\right)^{\frac{1}{p}} \quad (1 \leq p \leq \infty)$$

where y_i is the i-th entry of $\boldsymbol{y}$. For instance, we often use 2-norm or Euclidean lenght,

$$\|\boldsymbol{y}\|_2 = \left(\sum_{i=1}^{m} |y_i|^2\right)^{\frac{1}{2}} = \sqrt{\boldsymbol{y}^*\boldsymbol{y}}$$

because it is invariant under unitary and orthogonal transformations, that is ,

$$\|Q\boldsymbol{x}\|_2^2 = \boldsymbol{x}^*Q^*Q\boldsymbol{x} = \boldsymbol{x}^*\boldsymbol{x} = \|\boldsymbol{x}\|_2^2$$

where Q is unitary and/or othogonal matrix, so that $Q * Q = I$. In banded matrices, for norms to be consistent , we will use,

$$\|\boldsymbol{y}\|_\infty = \max_i |y_i| \qquad \text{but} \qquad \|\boldsymbol{y}^*\|_\infty = \|\boldsymbol{y}^T\|_\infty = \|\boldsymbol{y}\|_1 = \sum_i |\boldsymbol{y}_i|.$$

which do not give problems in the sense that it is straightforward calculated. Inner products can be bounded using p-norms by means of Hölder inequality.

Lemma 4.1. *(Hölder inequality) Let p and q satisfy $\frac{1}{p} + \frac{1}{q} = 1$, with $1 < p,\ q \leq \infty$. Then it is satisfied, for any vectors $\boldsymbol{x}$ and $\boldsymbol{y}$,*

$$|\boldsymbol{x}^*\boldsymbol{y}| \leq \|\boldsymbol{x}\|_p \|\boldsymbol{y}\|_q$$

The special case for $p = q = 2$ is the Cauchy- Schwarz inequality:

$$|\boldsymbol{x}^*\boldsymbol{y}| \leq \|\boldsymbol{x}\|_2 \|\boldsymbol{y}\|_2$$

Matrix norms are functions $\|\cdot\| : \mathbb{C}^{m\times n}$ that satisfy the three properties defined for vector norms but substituting the vector by the matrix. The important subordinate norm (matrix norms subordinate to vector norms) is defined as follows:

$$\|A\| = max_{x\neq 0} \frac{\|Ax\|}{\|x\|} = max_{\|x\|=1} \|Ax\|.$$

One case of subordinate matrix is the spectral norm, which is the only one used in this work (used to define Wilkinson condition number):

$$\|A\|_2 = (\rho(A * A))^{1/2} = \sigma_{max}(A)$$

where ρ is the spectral radius, defined as

$$\rho(B) = max\{|\lambda| : det(B - \lambda I) = 0\}$$

4.3 Algebraic and geometric multiplicities

Definition 4.5. *Let $\{\lambda_1, \ldots, \lambda_n\}$ be the eigenvalues of a matrix $A \in \mathbb{R}^{n\times n}$. For each i, we call Algebraic multiplicity of the eigenvalue λ_i, to the higher exponent α_i, for which the factor $(\lambda_i - \lambda)^{\alpha_i}$ appear in the factorization of the characteristic polynomial $p(\lambda)$.*

Definition 4.6. *Let $\{\lambda_1, \ldots, \lambda_n\}$ be the eigenvalues of a matrix $A \in \mathbb{R}^{n\times n}$. We call Geometric multiplicity of λ_i to the dimension d_i of the eigenspace V_{λ_i} of $A \in \mathbb{R}^{n\times n}$, that is,*

$$d_i = V_{\lambda_i} = n - rg(A - \lambda_i I)$$

4.4 Similarity transformations

The key tool to transform a matrix into another and to preserve all eigenvalue information is the so-called *Similarity transformation.*

Definition 4.7. *Two matrices $A \in \mathbb{R}^{m\times m}$ and $B \in \mathbb{R}^{m\times m}$ are similar if there exists a regular (nonsingular) matrix $P \in \mathbb{R}^{m\times m}$, such that $B = P^{-1}AP$. The transformation $B = P^{-1}AP$ is said to be a similarity transformation.*

Theorem 4.1. *(from [38]) If P is non singular, then A and $P^{-1}AP$ have the same characteristic polynomial, eigenvalues, and algebraic and geometric multiplicities.*

Proof. That both matrices have the same characteristic polynomials is shown by,

$$p_{P^{-1}AP}(\lambda) = det(P^{-1}AP - \lambda I) = det(P^{-1}(A - \lambda I)P) =$$

$$= det(P^{-1})det(A - \lambda I)det(P) = det(A - \lambda I) = p_A(\lambda)$$

Since the polinomials are equal , their factorizations are the same, thus , the algebraic multiplicity is the same. In the same line, we have proved in theorem 4.2, the relation between eigenvectors, this means that, if V_{λ_i} is the eigenspace of A, then $P^{-1}V_{\lambda_i}$ is an eigenspace of $P^{-1}AP$. □

Theorem 4.2. *Let $B = P^{-1}AP$ be a similarity transformation. If $\boldsymbol{x}$ is an eigenvector of B, then $P\boldsymbol{x}$ is an eigenvector of A.*

Proof. The proof is straightforward. Let us call $\boldsymbol{x}_A$ to the eigenvector of A and $\boldsymbol{x}_B$ to the eigenvector of B corresponding to the eigenvector λ, being $B = P^{-1}AP$, then,

$$A\boldsymbol{x}_A = \lambda\boldsymbol{x}_A \qquad \text{and} \quad B\boldsymbol{x}_B = \lambda\boldsymbol{x}_B$$

Substituting B by $P^{-1}AP$,

$$P^{-1}AP\boldsymbol{x}_B = \lambda\boldsymbol{x}_B$$

Therefore,

$$\boldsymbol{x}_A = P\boldsymbol{x}_B$$

□

4.5 Accuracy analysis: Floating point arithmetic, Sensitivity and Stability

4.5.1 Floating point arithmetic

Computers store a finite set of real numbers and operate on them. This limitation leads to a *model of representation* that consist of a discrete subset of $\mathbb{R}$, which is called $\mathbb{F}$, the system of floating point numbers. Let $\mathbb{F} \subset \mathbb{R}$ be the set of machine numbers, considering $0 \in \mathbb{F}$. Numbers of $\mathbb{F}$ are normalize so that, all floating point numbers follow this expression:

$$x = \pm m \times \beta^{e-t}$$

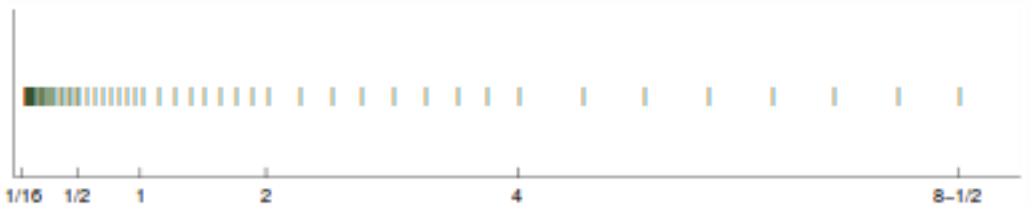

Figure 1: Floating point arithmetic representation with $t = 3, emax = 3, emin = -4$, from the programm $floatgui$ of [26].

where m is the mantissa of x, for e to be the exponent, β the radix or base, and t the precision. If t is the precision , m is an integer within $1 \leq m \leq \beta^t$. Similarly, since e is on the range, $emin \leq e \leq emax$, the range of floating point numbers is $\beta^{emin-1} \leq |x| \leq \beta^{emax}(1 - \beta^t)$. In double-precision IEEE arithmetic: $\beta = 2$, $t = 52$, $emin = -1022$ and $emax = +1023$.
There are two restrictions to representation of real numbers in this arithmetic: Boundness and discretization. Boundness means that there is one largest and one smallest number. The largest number, $L \in \mathbb{F}$, such that $|x| \leq L$, with $|x| \in \mathbb{F}$, is $L = 1.79 \times 10^{308}$. Similarly, the smallest number, $l \in \mathbb{F}$, that for $|x| \in \mathbb{F}$, satisfies $|x| \geq l$, is $l = 2.23 \times 10^{-308}$. If, for example, we simply need to compute a^p with $p >> 1$ the power of $a >> 1$, even if a can be represented, we may be given a huge number that fall out of the range, it means, such that $|x| > L$. In such case, operations will break down, that is, it leads to overflow. If $|x| < l$, is called underflow. The second restriction is the discrete condition of $\mathbb{F}$. There appear gaps among numbers and these gaps do not posses the same width. Usually, books show a very illustrative distribution of the gaps with figures. As an example look at Figure 1 extracted from [26], where it is represented the output of Matlab program $floatgui$, with $t = 3, emax = 3, emin = -4$. [Figure 1]. In [38], the exact position of each floating point number mapped into the real line is explained in a very clear form. It shows the general case of IEEE double precision arithmetic. That is, the interval [1,2] is represented by the discrete subset,

$$1, 1 + 2^{-52}, 1 + 2 \times 2^{-52}, 1 + 3 \times 2^{-52}, \ldots, 2. \tag{2}$$

In general, the entries of $\mathbb{F}$ in the interval $[2^j, 2^{j+1}]$ is represented by (2) times 2^j. For j=1, we have the interval [2,4],

$$2, 2 + 2^{-51}, 2 + 2 \times 2^{-51}, 2 + 3 \times 2^{-51}, \ldots, 4.$$

Note that within an interval, the size of the gaps is constant. Consecutive intervall gaps are different by a factor of 2. If a real number not in $\mathbb{F}$ lies within one of the gaps of one of the interval, it can be approximated by two nearby numbers and the standard choice is to select the closest one. The worst case is that the number would lie in the middle of the gap. This is the main reason to define the accuracy of $\mathbb{F}$ in terms:

$$\boxed{u = \frac{1}{2}\beta^{1-t}} \quad \text{that in double precision is} \quad u \approx \frac{1}{2} \times 2^{-52} = 1.11 \times 10^{-16}$$

This is the so-called *roundoff unit* and is "half the distance between 1 and the next larger floating point number" ([38]). In double precision arithmetic is half the distance between 1 and $1 + 2^{-52}$. Approximations have an error of less than u except for the number in the middle of the gap, in

which case there is a rule to break ties.
The quantity u has the following property:

$$\text{For all } \; x \in \mathbb{R}, \;\; \text{inside the range of } \mathbb{F}, \text{ there exists } \; x' \in \mathbb{F} \;\; \text{ such that } \; |x - x'| \leq u|x|. \tag{3}$$

(from [38]). It is interesting to express 3 in floating point terms (fl). That is ,

$$\text{For all } \; x \in \mathbb{R}, \;\; \text{inside the range of } \mathbb{F}, \text{ there exists } \; \varepsilon \; \text{ with } \; |\varepsilon| \leq u \;\; \text{ such that } \; fl(x) = x(1+\varepsilon). \tag{4}$$

4.5.2 Floating point operations

Elementary arithmetic operations are denoted by $+, -, \times, /$. In the same line, floating point operations are denoted by $\oplus, \ominus, \otimes, \oslash$.
The *Fundamental axiom of Floating point arithmetic* states,
Let $*$ be one of the operations $+, -, \times, /$ and let $\circledast$ be its floating point analogue, i.e.,

$$x \circledast y = fl(x * y).$$

Then for all $x, y \in \mathbb{F}$, there exist ε with $|\varepsilon| \leq \varepsilon_{machine}$ such that,

$$x \circledast y = (x * y)(1 + \varepsilon),$$

whenever underflow and overflow do not occur.

To end up, we briefly comment that sensitivity analysis works as a model of the behaviour of a problem under three factors that can perturb a matrix in a digital computer: truncation, rounding errors , data errors. The first two can be understood as discretization errors. Usually, computers may work with discrete tiny perturbations of the order of $u = 10^{-16}$ or higher in the entries of a matrix. The challenge of numerical analysts is to design stable numerical algorithm that converge to a solution with a tiny error of the order of the roundoff unit.

5 Notation and additional concepts

We would like to bring here, the notation used in this work, which is the same in [12]. We have taken as starting point based our computations in the results of that paper, so we follow a common notation in order to compare with the results in [12]. Therefore, we find that capital Roman letters $A, B, \ldots$ represent matrices, bolfaced lower case Roman letters $\boldsymbol{x}, \boldsymbol{y}, \ldots$ are used for vectors and lower case Greek letters $\alpha, \beta, \ldots$, for scalars. This is Householder's notation [18]. Among these vectors, letters $\boldsymbol{y}$ and $\boldsymbol{x}$ are reserved for left and right eigenvectors of matrices and , among scalars, λ for eigenvalues, to be applied in equations:

$$M\boldsymbol{x} = \boldsymbol{x}\lambda, \quad \boldsymbol{y}^* M = \lambda \boldsymbol{y}^*.$$

In that sense, eigenvectors can be complex although only real matrices are presented . So, we use $\boldsymbol{y}^T$ (transpose of the eigenvector) and $\boldsymbol{y}^* := \bar{\boldsymbol{y}}^T$, where $\bar{\alpha}$ is the conjugate of α.

For norms, in order to be consistent, we use the definitions

$$\|\boldsymbol{y}\|_2 = \sqrt{\boldsymbol{y}^* \boldsymbol{y}}$$

$$\|\boldsymbol{y}\|_\infty = \max_i |y_i| \qquad \text{but} \qquad \|\boldsymbol{y}^*\|_\infty = \|\boldsymbol{y}^T\|_\infty = \|\boldsymbol{y}\|_1 = \sum_i |\boldsymbol{y}_i|.$$

Now, we move to the concept of condition number. Conditioning estimates the changes in the outcomes in response to data perturbation. Our work is based in eigenvalues as outcomes. Formulation of eigenvalue condition numbers contains as a principal concept the notion of Wilkinson condition number. The definition connects the spectral projector onto λ 's eigenspace with the condition number for λ. For simple eigenvalues, λ, it is satisfied $\boldsymbol{y^*x} \neq 0$. Assuming simple eigenvalues, we arrive to the conclusion that the spectral projector onto λ's eigenspace is

$$P_\lambda = \boldsymbol{x}(\boldsymbol{y}^*\boldsymbol{x})^{-1}\boldsymbol{y}^* \tag{5}$$

as it is described in [12]. Then the spectral norm of P_λ is the *Wilkinson condition number* for λ, so

$$\kappa_\lambda := \|P_\lambda\|_2 = \frac{\|\boldsymbol{x}\|_2\|\boldsymbol{y}\|_2}{|\boldsymbol{y}^*\boldsymbol{x}|} = \frac{1}{|\cos\angle(x,y)|}. \tag{6}$$

Moreover, by applying to the general notion of condition number of any mathematical problem , the absolute (not relative) *Wilkinson condition number* can be defined also as:

$$\kappa_\lambda = \lim_{\eta\to 0} \sup\left\{ \frac{|\delta\lambda|}{\eta} \ : \ (\lambda+\delta\lambda) \text{ is an eigenvalue of } (M+\delta M),\ \|\delta M\|_2 \leq \eta \right\}. \tag{7}$$

(see proof in [15]). So, the corresponding relative *Wilkinson condition number*, named $BGT(\lambda; M)$, (used by Bini, Gemignani and Tisseur in [1]) is :

$$BGT(\lambda; M) := \kappa_\lambda \frac{\|M\|_2}{|\lambda|}, \quad \lambda \neq 0. \tag{8}$$

Note that κ_λ is invariant under translation $M \to M - \sigma I$, while BGT is not. BGT is interesting for comparing it to our relative condition number (relcond). Altough BGT is studied from the nomwise point of view, relcond is a componentwise analysis.

5.1 Summary of notation

[1] We will use the following letters:
B, C, G, J, M, T (Capital Roman letters) denote real tridiagonal n×n matrices.
A (Capital Roman letters) denotes a real banded n×n matrix.
$D, F, S, \mathring{R}$ (Capital Roman letters) denote real diagonal n×n matrices.
$\mathcal{L}, \bar{\mathcal{L}}, \mathcal{U}$ (Caligraphic and bar Roman letters) denote real bidiagonal n×n matrices.
L, U (Capital Roman letters) denote real bidiagonal n×n matrices.
Δ, Ω (Capital Greek letters) denote real sign diagonal n×n matrices.
H, N (Capital Roman letters) denote real 1's n×n matrices.
$\boldsymbol{x}, \boldsymbol{y}$, (boldfaced lower case Roman letters) denote column eigenvectors.
$\boldsymbol{v}, \boldsymbol{w}, \boldsymbol{h}, \boldsymbol{q}$ (boldfaced lower case Roman letters) denote column vectors.
$\beta, \delta, \eta, \lambda$ (lower case Greek letters) denote scalars.
q denotes the upper bandwith of an n×n banded matrix.
p denotes the lower bandwith of an n×n banded matrix.

[1] Terminology is inherited from article [12], since the present document extends many of its results to a more general setting.

6 Tridiagonal and banded matrices

The most important description, at a glance, of tridiagonal and/or banded matrices, could be that their definition is an identification of which entries are zero. Most of them belong to the category of sparse matrices and we can use sparse properties. We start with the definition and the main characteristics of tridiagonal matrices and then, continue with banded matrices. Let A be a square matrix, then A is tridiagonal if $a_{ij} = 0$ whenever $|i-j| > 1$. That means that if $|i-j| = 1$, it is permitted that some, a_{ij} are zero. If we want to avoid zeroes in the entries with $|i-j| = 1$, we should use another word to underline that <u>all</u> $a_{ij} \neq 0$ whenever $|i-j| = 1$, so that, the matrix is said to be tridiagonal *unreduced*. Tridiagonal unreduced matrices have the following structure:

$$C = \begin{pmatrix} x & x & & & \\ x & x & x & & \\ & \ddots & \ddots & \ddots & \\ & & x & x & x \\ & & & x & x \end{pmatrix} \tag{9}$$

following the notation used in [14], where x's are arbitrary nonzero entries. One outstanding property of simple eigenvalues λ is that if $E^i(\lambda) = Ker(C-\lambda I)^i; i = 1, 2, \ldots$, is the generalized subspace of λ, it turns out that $dim(E^1(\lambda)) = 1$ (geometric multiplicity of λ is 1), so there is only one independent eigenvector associated to each eigenvalue λ and only one Jordan block per eigenvalue with size 1×1. In fact, for simple eigenvalues $dim(E^i(\lambda)) = 1 \quad \forall i = 1, 2, \ldots$. Another interesting property is that, in unreduced tridiagonals , the entries of their eigenvectors are dependent among them by the recurrence relation,

$$C_{i,i-1}x_{i-1} + (C_{i,i} - \lambda)x_i + C_{i,i+1}x_{i+1} = 0, \quad i = 2, \ldots, n-1.$$

so, first and last entries are non-zero because otherwise the whole eigenvector would be zero. The category of tridiagonal matrices is, indeed, a specific case of banded matrix with $q = p = 1$, being q, p, the number of possibly nonzero diagonals above and under the main diagonal respectively. If A is a (p, q) banded matrix, then, for a number of diagonals q above the main diagonal, it will be satisfied $a_{ij} = 0$ whenever $j - i > q$ and for a number of diagonals p under the main diagonal, it will be satisfied $a_{ij} = 0$ whenever $i - j > p$. The parameter q is called *upper bandwidth* and the parameter p is said to be the *lower bandwidth*. The structure of a (p, q) banded matrix is

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1,q+1} & & 0 \\ a_{21} & & & & a_{2,q+2} & \\ \vdots & & & & \ddots & \\ a_{p+1,1} & & \ddots & & & a_{n-q,n} \\ & a_{p+2,2} & & & & \vdots \\ & & \ddots & & & \\ 0 & & & a_{n,n-p} & \cdots & a_{nn} \end{bmatrix} \tag{10}$$

where the entries and the dots delimit an hexagonal structure, with band form, which is thicker when p and q grow, being q , the upper bandwith and p, the lower bandwith. So A it is (p,q)-banded, with $q, p \geq 0$. As an advantage, this structure enables to reduce the cost of problems of the form of $A\boldsymbol{x} = b$ or $A\boldsymbol{x} = \lambda\boldsymbol{x}$. Usually the cost is reduced from $O(n^3)$ flops to $O((p+q)n^2)$ or

$O((p+q)n)$ flops (see [6], [7]). Tridiagonal and small width banded matrices belong to the category of sparse matrices, hence, sparse properties can be used. We shall use less space to store matrices, for example.

6.1 Balancing

Balancing is a method for making the ith row and ith column norms of a matrix equal, for all i, for norms $\|\cdot\|_1, \|\cdot\|_2$, or $\|\cdot\|_\infty$, by means of diagonal similarity transformations. For general matrices is an iterative process. For unreduced tridiagonal matrices it is possible to find explicitly a diagonal matrix D, such that $B = DCD^{-1}$ is balanced (to find the structures of C and B see below, in (12)). To avoid rounding errors, balancing, is often done by changes in the exponents (See [30]). In some rare cases balancing can spread this errors and make them more significant, but usually is a method that helps, for example, it helps to make unsymmetric matrices closer to symmetric. In addition, for unreduced tridiagonal matrices there are other similarity transformations that allow to reduce the number of parameters of the problem $J = \tilde{D}J\tilde{D}^{-1}$, but this is not balancing. This facts are summed up by saying that equivalent tridiagonal matrices which are linked by similarity transformations that were proposed in article [12] are now also useful for us,

$$C = \begin{pmatrix} a_1 & c_1 & & & \\ b_1 & a_2 & c_2 & & \\ & \ddots & \ddots & \ddots & \\ & & b_{n-2} & a_{n-1} & c_{n-1} \\ & & & b_{n-1} & a_n \end{pmatrix}, \quad J = \begin{pmatrix} a_1 & 1 & & & \\ b_1c_1 & a_2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & b_{n-2}c_{n-2} & a_{n-1} & 1 \\ & & & b_{n-1}c_{n-1} & a_n \end{pmatrix} \tag{11}$$

and

$$B = \begin{pmatrix} a_1 & \sqrt{|b_1c_1|} & & & & \\ \gamma_1\sqrt{|b_1c_1|} & a_2 & \sqrt{|b_2c_2|} & & & \\ & \ddots & \ddots & \ddots & & \\ & & \gamma_{n-2}\sqrt{|b_{n-2}c_{n-2}|} & a_{n-1} & \sqrt{|b_{n-1}c_{n-1}|} & \\ & & & \gamma_{n-1}\sqrt{|b_{n-1}c_{n-1}|} & a_n & \end{pmatrix} \tag{12}$$

where $\gamma_i = \text{sign}(b_ic_i)$, $i = 1, \ldots, n-1$. Although appear exceptions, matrices B are not symmetric. In some cases, the best we can achieve when balancing is *approximately* equal row and column norms. "For general applications, C is the input matrix and J uses the fewest parameters" [12].

7 Factored forms of tridiagonal matrices

Factored forms of tridiagonal matrices of [12] play an important role to analyse whether eigenvalues are more sensitive to the entries of tridiagonal matrices perturbations than to perturbations in its factored forms, a phenomenon that has been observed in practice. For that purpose, we present herein two special factored forms of tridiagonal matrices, extracted from the mentioned document [12] : $B = \Delta LDL^T$ and $J = \mathcal{LU}$, for B and J being those matrices introduced in section 6.1, it is

meant balanced and fewer parameter matrices. Tridiagonal matrices are very easy to balance (see (12)) and even easier to make them real symmetric. The trick consist of changing sings on strategic rows. This can be achieved using a special diagonal matrix, called *signature matrix* Δ,

$$\Delta = \text{diag}(\delta_1, \delta_2, \ldots, \delta_n), \qquad \delta_i = \pm 1.$$

So, from B (12) and Δ, we obtain a real symmetric tridiagonal matrix T,

$$\Delta B = T,$$

Note that this transformation does not conserve eigenvalues of C since the eigenvalues of B and T are different. So, if $(B - \lambda I)\boldsymbol{x} = \mathbf{0}$ we need to premultiply by Δ to find

$$(T - \lambda\Delta)\boldsymbol{x} = \mathbf{0}. \tag{13}$$

Observe that T admits, in general, triangular factorization that are symmetric as

$$T = LDL^T \tag{14}$$

with L, a lower bidiagonal matrix of the form,

$$L = \begin{pmatrix} 1 & & & & \\ l_1 & 1 & & & \\ & \ddots & \ddots & & \\ & & l_{n-2} & 1 & \\ & & & l_{n-1} & 1 \end{pmatrix},$$

and $D = \text{diag}(d_1, d_2, \ldots, d_n)$ (matrix of pivots). Then, we can express B in the following form,

$$B = \Delta T = \Delta LDL^T, \tag{15}$$

$$B^T = T\Delta = LDL^T\Delta. \tag{16}$$

As it is discussed in [12], elements of L are not necessary of the same order of D in factorization $T = LDL^T$, so, and may also happen that there can be large element growth, i.e.,

$$\|D\|_2 \gg \|T\|_2, \qquad \|L\|_2 \gg \|T\|_2.$$

In another line of explanation, recall that tridiagonal matrices B and C have common eigenvalues since they are related by a diagonal similarity transformation,

$$B = FCF^{-1}$$

Considering,

$$B\boldsymbol{x}_B = \boldsymbol{x}_B\lambda \quad \text{ and } C\boldsymbol{x}_C = \boldsymbol{x}_C\lambda$$

we have,

$$C\left(F^{-1}\boldsymbol{x}_B\right) = F^{-1}B\boldsymbol{x}_B = \left(F^{-1}\boldsymbol{x}_B\right)\lambda, \qquad F \text{ is diagonal.}$$

hence, the relation between eigenvectors is given by

$$\boldsymbol{x}_C = F^{-1}\boldsymbol{x}_B$$

This factorization allows the comparison of the factored form $\Delta T = \Delta L D L^T$ with B, itself, to find out if eigenvalues are more sensitive to componentwise perturbations in B entries or if it is more sensitive to perturbations in L and D. Another factorization appear for the J-form of C (see (11)). Suppose that J admits triangular factorization

$$J = \mathcal{L}\mathcal{U}$$

where $\mathcal{L}$ ($\neq L$ above) and $\mathcal{U}$ are lower and upper bidiagonals, respectively, of the form

$$\mathcal{L} = \begin{pmatrix} 1 & & & & \\ \mathfrak{l}_1 & 1 & & & \\ & \ddots & \ddots & & \\ & & \mathfrak{l}_{n-2} & 1 & \\ & & & \mathfrak{l}_{n-1} & 1 \end{pmatrix}, \qquad \mathcal{U} = \begin{pmatrix} \mathfrak{u}_1 & 1 & & & \\ & \mathfrak{u}_2 & 1 & & \\ & & \ddots & \ddots & \\ & & & \mathfrak{u}_{n-1} & 1 \\ & & & & \mathfrak{u}_n \end{pmatrix}. \tag{17}$$

For the $J = \mathcal{L}\,\mathcal{U}$ factorization it is also reasonable to find out if the eigenvalues in $\mathcal{L}\mathcal{U}$ are more sensitive to perturbations in $\mathcal{L}$ and $\mathcal{U}$ entries than the eigenvalues of J to perturbations in J. We wait to find that the relative condition numbers of C, J and B turn out to be equal (see Lemma 9.1). It also turns out that the relative eigenvalue condition numbers for the various factored forms ($J = \mathcal{L}\mathcal{U}, B = \Delta T = \Delta L D L^T$) are equivalent (see Section (9.5)). So, as in [12], we only present detailed derivations of relative eigenvalue condition numbers for C and for the factors $\mathcal{L}$, $\mathcal{U}$ and present just the formulae for the other factored forms.

8 LU factorization of banded matrices

Sensitivity analysis for banded matrices only accept standard LU factorization, being L a lower triangular matrix with $1's$ in its main diagonal and U an upper triangular matrix. For this special matrices the factorization LU without pivoting conserves the band structure. So L and U have this aspect.

$$L = \begin{pmatrix} 1 & & & & & \\ l_{21} & & & & & \\ \vdots & & & & & \\ l_{p+1,1} & & & \ddots & & \\ & l_{p+2,2} & & & & \\ & & \ddots & & & \\ & & & & l_{n,n-p} & \cdots & 1 \end{pmatrix} \quad \text{and} \quad U = \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1,q+1} & & \\ & u_{22} & & & & \\ & & & & \ddots & \\ & & & \ddots & & u_{n-q,n} \\ & & & & & \vdots \\ & & & & & u_{nn} \end{pmatrix}. \tag{18}$$

for a band matrix with upper bandwidth q and lower bandwidth p. Note that neither U, nor L are supposed to be singular. If they were, the form of the factorization LU might be very different.

9 Gradients and condition numbers of eigenvalues of tridiagonal matrices

This section is divided in three parts: the first is devoted to the definition of the ECN (eigenvalue condition number) under perturbations with respect parameters , the second includes the derivation

of the ECN and in the third one we set a proof of the equivalence between different ECNs. For convenience, we set the name of the ECNs to $relcond_2(\lambda, T)$, where T is an square matrix (tridiagonal, banded or factored) instead of the 1-norm used in [12]. The ECN, $relcond_2(\lambda, T)$, will indicate that we have used 2-norm for $relcond(\lambda, T)$ computation. It is important, to underline that Wilkinson condition number defined in (7), κ_λ , is a measurement of the absolute sensitivity i.e., quotation of [12] :" a measure of the absolute variation of an eigenvalue with respect to the norm of the matrix ". Whereas, the strategy of the reference work, [12] , was to consider relative variations of the eigenvalues in response to the largest relative perturbations of each of the entries of the associated matrix of factored forms. [2] The new results extending those in to paper [12] are in sections from 9.2 to 9.5, and 10.1. We start from the concepts in paper [12], but we continue by calculating the eigenvalue condition numbers in the 2-norm (of $relgrad(\lambda)$) and we analyse the results for such norm. In addition, it is clearifying to mention that, on the one hand, section 9 will exclusively consider tridiagonal matrices. The difference with the eigenvalue sensitivity analysis of tridiagonal matrices of [12] is that whereas the authors of that work measure the eigenvalue sensitivity by using 1-norm, here we use 2 norm, which can have some advantages (invariant under unitary and orthogonal transformations and derivable). On the other hand, section 10 will consider only general banded matrices, which are not included in [12].

9.1 Sensitivity analysis

According to [36], simple eigenvalues of C, defined in (11), are continuous functions of the entries of C in all its domain and differentiable (C^1 type). So that, we can obtain their gradients. Taking this into account, if we consider that diagonals of C are in arrays $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$ (see 11), then, for infinitesimal absolute changes,

$$(\delta a_1, \ldots, \delta a_n, \delta b_1, \ldots, \delta b_{n-1}, \delta c_1, \ldots, \delta c_{n-1})^T =: \delta C$$

we define,

$$rel\delta C = \left(\frac{\delta a_1}{a_1}, \ldots, \frac{\delta c_{n-1}}{c_{n-1}} \right)^T \tag{19}$$

which is the relative perturbation vector of C entries, where zero entries are supposed to remain as zero.
We can upper bound perturbations size such that the relative condition number of the simple eigenvalue λ is defined as

$$relcond_2(\lambda; C) = \lim_{\eta \to 0} \sup \left\{ \frac{|\delta\lambda|}{\eta|\lambda|} \ : \ (\lambda + \delta\lambda) \text{ is an eigenvalue of } \ (C + \delta C), \ \|rel\delta C\|_2 \leq \eta \right\}. \tag{20}$$

This is a general definition, so we need to assign the supremum to an specific value that should be computable and this should be a low cost computation.
The principal procedure to make the above definition (20), more practical is shown next.
For arrays $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$ of C diagonals and for simple eigenvalues λ, we gather all the partial derivatives $\left\{ \frac{\partial\lambda}{\partial a_i}, \frac{\partial\lambda}{\partial b_i}, \frac{\partial\lambda}{\partial c_i} \right\}$ and form the absolute gradient

$$\text{grad}_C(\lambda) = \left(\frac{\partial\lambda}{\partial a_1}, \ldots, \frac{\partial\lambda}{\partial a_n}, \frac{\partial\lambda}{\partial b_1}, \ldots, \frac{\partial\lambda}{\partial b_{n-1}}, \frac{\partial\lambda}{\partial c_1}, \ldots, \frac{\partial\lambda}{\partial c_{n-1}} \right)^T. \tag{21}$$

[2] Recall that the main concepts of this document are based on the article [12]. This means that some expressions are obtained from that article in order to lighten the main conclusions of this work.

Then, we relate it to the perturbation vector of C entries, δC, to have,

$$\delta\lambda = \mathrm{grad}_C(\lambda)^T \cdot \delta C + \textit{higher order terms (h.o.t.).} \tag{22}$$

This is not relative yet, so we need to relate $|\delta\lambda/\lambda|$ to $|\delta p_j/p_j|, j = 1,\ldots,3n-2$, $\boldsymbol{p} = \big(a_1,\ldots,a_n,b_1,\ldots,b_{n-1},c_1,\ldots,c_{n-1}\big)$. For no zeros, we can express the relative form of (22) as

$$\begin{aligned}\frac{\delta\lambda}{\lambda} &= \left(\frac{a_1}{\lambda}\frac{\partial\lambda}{\partial a_1},\ldots,\frac{c_{n-1}}{\lambda}\frac{\partial\lambda}{\partial c_{n-1}}\right)\cdot\left(\frac{\delta a_1}{a_1},\ldots,\frac{\delta c_{n-1}}{c_{n-1}}\right)^T + h.o.t.\\ &=: \textit{relgrad}_C(\lambda)^T \cdot \textit{rel}\,\delta C + h.o.t.,\end{aligned} \tag{23}$$

defining the *relative gradient* and the *relative perturbation*. Based on the following quotation of [12] : "When a parameter vanishes we should omit the corresponding term in the inner product", we follow the same convention here. Taking into account rounding errors, it is natural to have bounded perturbations of the form ,

$$\|\,\textit{rel}\,\delta C\|_2 \le \eta \quad \text{for} \quad 0 < \eta \ll 1, \tag{24}$$

Then, using $|\boldsymbol{u}^T\boldsymbol{v}| \le \|\boldsymbol{u}\|_2\|\boldsymbol{v}\|_2$ (Cauchy-Schwarz inequality), write ,

$$\left|\frac{\delta\lambda}{\lambda}\right| \le \|\,\textit{relgrad}_C(\lambda)\|_2\|\,\textit{rel}\,\delta C\|_2 + h.o.t. \le \eta\|\,\textit{relgrad}_C(\lambda)\|_2 + h.o.t. \tag{25}$$

where h.o.t are higher order terms, and we show that $\|\textit{relgrad}_C(\lambda)\|_2$ coincides with the *structured relative condition number* for λ as a function of C, defined in (20), in the next theorem:

Theorem 9.1. *For any unreduced real tridiagonal matrix C and any structured componentwise perturbation of it, δC, whose relative perturbation vector, $\textit{rel}\,\delta C$, defined in (19), satisfy $\|\,\textit{rel}\,\delta C\|_2 \le \eta$, if we define a relative condition number for a simple eigenvalue λ of C as in (20), then*

$$\boxed{\textit{relcond}_2(\lambda;C) := \|\,\textit{relgrad}_C(\lambda)\|_2, \quad \lambda \neq 0.}$$

Proof. We know that

$$\left|\frac{\delta\lambda}{\lambda}\right| \le \|\,\textit{relgrad}_C(\lambda)\|_2\|\,\textit{rel}\,\delta C\|_2 + h.o.t., \le \eta\|\,\textit{relgrad}_C(\lambda)\|_2 + h.o.t. \tag{26}$$

with equality attained only if there exist a constant β such that,

$$\textit{rel}\,\delta C = \beta\,\textit{relgrad}_C(\lambda) \tag{27}$$

by the properties of inner product (parallel vectors). Therefore, always exists a vector $\textit{rel}\,\delta C$, such that verifies (27), and satisfies perturbations of the size $\|\,\textit{rel}\,\delta C\|_2 \le \eta$. Taking the 2-norm of (27) yields,

$$\beta = \frac{\eta}{\|\,\textit{relgrad}_C(\lambda)\|_2}$$

That is:

$$\textit{rel}\,\delta C = \frac{\eta}{\|\,\textit{relgrad}_C(\lambda)\|_2}\,\textit{relgrad}_C(\lambda)^T$$

and , since we can always find that particular vector, the equality

$$\textit{relcond}(\lambda;C) := \|\,\textit{relgrad}_C(\lambda)\|_2, \quad \lambda \neq 0.$$

can always be achieved.
Observe that we take the largest perturbation attained for $\|\textit{rel}\delta C\|_2$, which is η. □

Singularity is one aspect of conditioning preferably avoided. Consider that $C \neq O$ is singular, fortunately, turns out that relative changes done to the entries almost always destroy singularity. Therefore, the only problem is with $\lambda = 0$, in which case we set $relcond(0; C) = \infty$. So we ask for high relative accuracy for all the eigenvalue condition number and it can be reached it in certain occasions, for quite small eigenvalues, except for $\lambda = 0$.
Again we quote from [12]: "We do not know in advance when η is small enough to warrant the neglect of *h.o.t.*. Our numerical examples shed light on this topic. We know of no other study that addresses it". So the allowed values of η in numerical practice may be not easy to be determined.

9.2 Representation 1 - entries of C

In this section we derive the relative condition number for C (11) with respect to the 2-norm according to definition (20) . We treat each entry of C as an independent variable. Thus, with $I = \big(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\big)$ which will act as a column vector,

$$\frac{\partial C}{\partial a_j} = \boldsymbol{e}_j \boldsymbol{e}_j^T, \quad \frac{\partial C}{\partial b_j} = \boldsymbol{e}_{j+1} \boldsymbol{e}_j^T \quad \text{and} \quad \frac{\partial C}{\partial c_j} = \boldsymbol{e}_j \boldsymbol{e}_{j+1}^T.$$

Consider $\lambda \neq 0$ as a simple eigenvalue of C,

$$C\boldsymbol{x} = \boldsymbol{x}\lambda, \quad \boldsymbol{y}^* C = \lambda \boldsymbol{y}^*.$$

Then, for $p_j = a_j, b_j, c_j$, we differentiate $C\boldsymbol{x} = \boldsymbol{x}\lambda$ to get

$$\frac{\partial C}{\partial p_j}\boldsymbol{x} + C\frac{\partial \boldsymbol{x}}{\partial p_j} = \frac{\partial \boldsymbol{x}}{\partial p_j}\lambda + \boldsymbol{x}\frac{\partial \lambda}{\partial p_j}.$$

Multiply by $\boldsymbol{y}^*$ and cancel equal terms to find

$$\frac{\partial \lambda}{\partial p_j}\boldsymbol{y}^*\boldsymbol{x} = \boldsymbol{y}^* \frac{\partial C}{\partial p_j}\boldsymbol{x}, \quad p_j = a_j, b_j, c_j.$$

Thus,

$$\frac{\partial \lambda}{\partial a_j} = \frac{\overline{y_j} x_j}{\boldsymbol{y}^*\boldsymbol{x}}, \quad \frac{\partial \lambda}{\partial b_j} = \frac{\overline{y_{j+1}} x_j}{\boldsymbol{y}^*\boldsymbol{x}}, \quad \frac{\partial \lambda}{\partial c_j} = \frac{\overline{y_j} x_{j+1}}{\boldsymbol{y}^*\boldsymbol{x}}$$

and

$$relgrad_C(\lambda) = \frac{1}{\lambda \boldsymbol{y}^*\boldsymbol{x}}\big(a_1\overline{y_1}x_1, \ldots, a_n\overline{y_n}x_n, b_1\overline{y_2}x_1, \ldots, b_{n-1}\overline{y_n}x_{n-1}, c_1\overline{y_1}x_2, \ldots, c_{n-1}\overline{y_{n-1}}x_n\big)^T.$$

consequently, if $C \circ C$ is the Hadamard product of C by itself, i.e., $(C \circ C)_{ij} = c_{ij}^2$, and we define

$$\boldsymbol{y'} = \boldsymbol{y} \circ \boldsymbol{y} = \big(y_1^2, \quad \ldots, \quad y_n^2\big)^T, \quad \boldsymbol{x'} = \boldsymbol{x} \circ \boldsymbol{x} = \big(x_1^2, \quad \ldots, \quad x_n^2\big)^T \tag{28}$$

we obtain,

$$\| relgrad_C(\lambda)\|_2 = \frac{1}{|\lambda||\boldsymbol{y}^*\boldsymbol{x}|}\left[\sum_{j=1}^{n} |a_j^2 y_j^2 x_j^2| + \sum_{j=1}^{n-1}\left(|b_j^2 y_{j+1}^2 x_j^2| + |c_j^2 y_j^2 x_{j+1}^2|\right)\right]^{1/2} = \frac{\big[|\boldsymbol{y'}|^T |C \circ C||\boldsymbol{x'}|\big]^{1/2}}{|\lambda||\boldsymbol{y}^*\boldsymbol{x}|}.$$

We can summarize the results above in a compact form in Theorem 9.2.

Theorem 9.2. *Let $\lambda \neq 0$ be a simple eigenvalue of an unreduced real tridiagonal matrix C with left eigenvector $\boldsymbol{y}$ and right eigenvector $\boldsymbol{x}$. Denote by $C \circ C$, the Hadamard product of C and denote by $\boldsymbol{x'}, \boldsymbol{y'}$ the modified left and right eigenvectors with each of its entries squared as defined in (28) . Let $rel\delta C$ be the relative perturbation vector (19). Then $relcond_2(\lambda; C) := \| relgrad_C(\lambda)\|_2$ is equal to*

$$\begin{aligned} relcond_2(\lambda; C) &= \lim_{\eta \to 0} \sup \left\{ \frac{|\delta\lambda|}{\eta|\lambda|} \, : \, (\lambda + \delta\lambda) \text{ is an eigenvalue of } (C + \delta C), \, \|rel\delta C\|_2 \leq \eta \right\} \\ &= \frac{\left[|\boldsymbol{y'}|^T |C \circ C| |\boldsymbol{x'}|\right]^{1/2}}{|\lambda||\boldsymbol{y}^* \boldsymbol{x}|}. \end{aligned}$$

The form of $relcond_2(\lambda; C)$ yields the following result.

Lemma 9.1. *For any scaling matrix S invertible and diagonal,*

$$relcond_2(\lambda; SCS^{-1}) = relcond_2(\lambda; C).$$

Proof. For any matrix G such that $G = SCS^{-1}$ we can define a matrix $G \circ G$ that satisfies $G \circ G = (S)^2 (C \circ C)(S^{-1})^2$. The proof of this fact is given by

$$(G \circ G)_{ij} = [g_{ij}]^2 = [s_{ii} c_{ij} s_{jj}^{-1}]^2 = [s_{ii}]^2 [c_{ij}]^2 [s_{jj}^{-1}]^2 = ((S)^2 (C \circ C)(S^{-1})^2)_{ij}$$

Also, the following relationship exists between the eigenvectors of G and C ,

$$\boldsymbol{y}^* = \boldsymbol{y}_G^* S \quad \text{and} \quad \boldsymbol{x} = S^{-1} \boldsymbol{x}_G.$$

Consequently, its corresponding $\boldsymbol{y'}_G, \boldsymbol{x'}_G$, $\boldsymbol{y'}, \boldsymbol{x'}$ are related in an analogous form,

$$\boldsymbol{y'}^* = \boldsymbol{y'}_G^* (S)^2 \quad \text{and} \quad \boldsymbol{x'} = (S^{-1})^2 \boldsymbol{x'}_G.$$

where $\boldsymbol{y'}, \boldsymbol{x'}$ are defined in (28) and $\boldsymbol{y'}_G, \boldsymbol{x'}_G$, are the vectors whose entries are the squares of the entries of the eigenvector of G.
Thus, $|\boldsymbol{y}^* \boldsymbol{x}| = |\boldsymbol{y}_G^* \boldsymbol{x}_G|$ and

$$\begin{aligned} |\boldsymbol{y'}^*||C \circ C||\boldsymbol{x'}| &= |\boldsymbol{y'}_G^* S^2||C \circ C||(S^{-1})^2 \boldsymbol{x'}_G| = |\boldsymbol{y'}_G^*||S^2||C \circ C||(S^{-1})^2||\boldsymbol{x'}_G| \\ &= |\boldsymbol{y'}_G^*||G \circ G||\boldsymbol{x'}_G| = |\boldsymbol{y'}_G|^T |G \circ G||\boldsymbol{x'}_G|, \end{aligned}$$

because S is diagonal and no additions occur in $S^2 (C \circ C)(S^{-1})^2$. □

This lemma indicates that $relcond_2(\lambda, C)$ is invariant under diagonal similarity transformations. Recall that J and B (see section 6.1) were obtained by similarity transformations in C and so

$$relcond_2(\lambda; C) = relcond_2(\lambda; J) = relcond_2(\lambda; B)$$

Thus, from the point of view of entrywise sensitivity, we do not need to balance the matrix C , neither to construct J, because there will not be improvement in λ sensitivity since the condition numbers are the same. Nevertheless, there is a need to obtain J and B to compute the factorizations $\Delta B = LDL^T$ and $J = \mathcal{LU}$.

9.3 Representation 2 - $\mathcal{L}$, $\mathcal{U}$ representation of J

We start by defining some matrices to avoid confusions and to let posterior calculation be more fluent. After that, we derive the relative gradient for $\mathcal{LU}$ as it is done in [12] and we derive a 2-norm for that vector, following the structure of the previous section. Finally we derive the eigenvalue condition number in the 2- norm for this representation. Recall that $\mathcal{LU}$ is the factorization of matrix J (11), that is $J = \mathcal{LU}$ and the definitions of $\mathcal{L}$ and $\mathcal{U}$ were shown in (17). Consider $\mathcal{L}$ as an addition of matrices, $\mathcal{L} = I + \mathring{\mathcal{L}}$, and consider $\mathcal{U}$, expressed as $\mathcal{U} = \text{diag}(\mathbf{u}) + N$,where $\mathring{\mathcal{L}}$ and N are,

$$\mathring{\mathcal{L}} = \begin{pmatrix} 0 & & & & \\ \mathfrak{l}_1 & 0 & & & \\ & \ddots & \ddots & & \\ & & \mathfrak{l}_{n-2} & 0 & \\ & & & \mathfrak{l}_{n-1} & 0 \end{pmatrix} \quad \text{and} \quad N = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix}. \tag{29}$$

Consider also the permutation matrix H,

$$H = \begin{pmatrix} 0 & & \cdots & & 1 \\ 1 & 0 & & & \\ & \ddots & \ddots & & \vdots \\ & & 1 & 0 & \\ & & & 1 & 0 \end{pmatrix} \tag{30}$$

Consider the perturbations vector,

$$rel\delta(\mathcal{L},\mathcal{U}) = \left(\frac{\delta\mathfrak{l}_1}{\mathfrak{l}_1}, \ldots, \frac{\delta\mathfrak{l}_{n-1}}{\mathfrak{l}_{n-1}}, \ldots, \frac{\delta\mathfrak{u}_1}{\mathfrak{u}_1}, \ldots, \frac{\delta\mathfrak{u}_n}{\mathfrak{u}_n} \right) \tag{31}$$

And to end up, define the eigenvector diagonal matrices as,

$$\text{diag}(\boldsymbol{x}) = \text{diag}(x_1, \cdots, x_n), \quad \text{diag}(\boldsymbol{y}) = \text{diag}(y_1, \cdots, y_n) \tag{32}$$

Thus, as we did in previous section for the eigenvalue condition number with respect the entries of C, we derive a relative gradient for $J = \mathcal{LU}$ assuming that J admits triangular factorization $\mathcal{LU}$: For $\mathfrak{u}_j$ we find

$$\frac{\partial\lambda}{\partial\mathfrak{u}_j}\boldsymbol{y}^*\boldsymbol{x} = \boldsymbol{y}^*\mathcal{L}\frac{\partial\mathcal{U}}{\partial\mathfrak{u}_j}\boldsymbol{x} = \boldsymbol{y}^*\mathcal{L}\boldsymbol{e}_j\boldsymbol{e}_j^T\boldsymbol{x} = (\boldsymbol{y}^*\mathcal{L})_j x_j, \quad j = 1, \ldots, n,$$

and for $\mathfrak{l}_j$,

$$\frac{\partial\lambda}{\partial\mathfrak{l}_j}\boldsymbol{y}^*\boldsymbol{x} = \boldsymbol{y}^*\frac{\partial\mathcal{L}}{\partial\mathfrak{l}_j}\mathcal{U}\boldsymbol{x} = \boldsymbol{y}^*\boldsymbol{e}_{j+1}\boldsymbol{e}_j^T\mathcal{U}\boldsymbol{x} = \overline{y_{j+1}}(\mathcal{U}\boldsymbol{x})_j, \quad j = 1, \ldots, n-1.$$

Then

$$\text{grad}_{\mathcal{L},\mathcal{U}}(\lambda) = \left(\frac{\partial\lambda}{\partial\mathfrak{u}_1}, \ldots, \frac{\partial\lambda}{\partial\mathfrak{u}_n}, \frac{\partial\lambda}{\partial\mathfrak{l}_1}, \ldots, \frac{\partial\lambda}{\partial\mathfrak{l}_{n-1}} \right)^T \tag{33}$$

and, including the parameters $\mathfrak{l}_j$ and $\mathfrak{u}_j$ appropriately,

$$\begin{aligned} \lambda(\boldsymbol{y}^*\boldsymbol{x})\, relgrad_{\mathcal{LU}}(\lambda) =& \big((\boldsymbol{y}^*\mathcal{L})_1\mathfrak{u}_1x_1, \ldots, (\boldsymbol{y}^*\mathcal{L})_n\mathfrak{u}_nx_n, \overline{y_2}\mathfrak{l}_1(\mathcal{U}\boldsymbol{x})_1, \ldots, \overline{y_n}\mathfrak{l}_{n-1}(\mathcal{U}\boldsymbol{x})_{n-1}\big)^T \\ =& \big((\boldsymbol{y}^*\mathcal{L})_1\mathfrak{u}_1x_1, \ldots, (\boldsymbol{y}^*\mathcal{L})_n\mathfrak{u}_nx_n, (\boldsymbol{y}^*\mathring{\mathcal{L}})_1(\mathcal{U}\boldsymbol{x})_1, \ldots, (\boldsymbol{y}^*\mathring{\mathcal{L}})_{n-1}(\mathcal{U}\boldsymbol{x})_{n-1}\big)^T. \end{aligned} \tag{34}$$

So, expressing (34) in a vector-matrix product form, we obtain the 2-norm of this gradient using the above definitions of H, $\mathrm{diag}(\mathfrak{u})$, $\mathrm{diag}(\boldsymbol{x})$ and $\mathrm{diag}(\boldsymbol{y})$,

$$\begin{aligned}
|\lambda||\boldsymbol{y}^*\boldsymbol{x}|\| \, relgrad_{\mathcal{L},\mathcal{U}}(\lambda)\|_2 &= \left[\sum_{j=1}^{n} \left|(\boldsymbol{y}^*\mathcal{L})_j^2\mathfrak{u}_j^2x_j^2\right| + \sum_{j=1}^{n-1} |(\boldsymbol{y}^*\mathring{\mathcal{L}})_j^2(\mathcal{U}\boldsymbol{x})_j^2|\right]^{1/2} \\
&= [|\boldsymbol{y}^*\mathcal{L}||\mathrm{diag}(\mathfrak{u})||\mathrm{diag}(\boldsymbol{x})||\mathrm{diag}(\boldsymbol{x})||\mathrm{diag}(\mathfrak{u})||(\boldsymbol{y}^*\mathcal{L})^T| \\
&\quad + |\mathcal{U}\boldsymbol{x}|^T|H|^T|\mathrm{diag}(\boldsymbol{y})||\mathring{\mathcal{L}}||\mathring{\mathcal{L}}^T||\mathrm{diag}(\boldsymbol{y})||H||\mathcal{U}\boldsymbol{x}|]^{1/2} \qquad (35) \\
&= [|(\boldsymbol{y}^*\mathcal{L})\mathrm{diag}(\mathfrak{u})\mathrm{diag}(\boldsymbol{x})||\mathrm{diag}(\boldsymbol{x})\mathrm{diag}(\mathfrak{u})(\boldsymbol{y}^*\mathcal{L})^T| \\
&\quad + |(\mathcal{U}\boldsymbol{x})^TH^T\mathrm{diag}(\boldsymbol{y})\mathring{\mathcal{L}}||\mathring{\mathcal{L}}^T\mathrm{diag}(\boldsymbol{y})H(\mathcal{U}\boldsymbol{x})|]^{1/2}
\end{aligned}$$

where we have used that no addition occur inside the absolute values. Note that the two terms of the sum contain a product of a vector by its transpose. Then, $\| \, relgrad_{\mathcal{L},\mathcal{U}}(\lambda)\|_2$ can be written in a more elegant form. This can be seen in next theorem, Theorem 9.3.

Theorem 9.3. *Let J be an unreduced real tridiagonal matrix with 1's in the first upper diagonal that permits a triangular factorization $J = \mathcal{L}\mathcal{U}$ with factors as in* (17)*. Denote $\mathring{\mathcal{L}} = \mathcal{L} - \mathcal{I}$,* $\mathrm{diag}(\mathfrak{u}) = \mathrm{diag}(\mathfrak{u}_1, \ldots, \mathfrak{u}_n)$*, and let* $\mathrm{diag}(\boldsymbol{x}), \mathrm{diag}(\boldsymbol{y}), H$ *be as defined in (32) and (30). Let $rel\delta(\mathcal{L},\mathcal{U})$ be defined in (31). Let $\lambda \neq 0$ be a simple eigenvalue of J with left eigenvector $\boldsymbol{y}$ and right eigenvector $\boldsymbol{x}$. And finally , denote $\boldsymbol{v}_{\mathcal{L}\mathcal{U}}^T = (\boldsymbol{y}^*\mathcal{L})\mathrm{diag}(\mathfrak{u})\mathrm{diag}(\boldsymbol{x})$ and denote $\boldsymbol{w}_{\mathcal{L}\mathcal{U}}^T = (\mathcal{U}\boldsymbol{x})^TH^T\mathrm{diag}(\boldsymbol{y})\mathring{\mathcal{L}}$. Then $relcond_2(\lambda;\mathcal{L},\mathcal{U}) := \| \, relgrad_{\mathcal{L},\mathcal{U}}(\lambda)\|_2$ is equal to*

$$relcond_2(\lambda;\mathcal{L},\mathcal{U}) = \lim_{\eta\to 0} \sup\left\{ \frac{|\delta\lambda|}{\eta|\lambda|} \, : \, (\lambda+\delta\lambda) \text{ is an eigenvalue of } (\mathcal{L}+\delta\mathcal{L})(\mathcal{U}+\delta\mathcal{U}), \right.$$
$$\left. \| \, rel\delta(\mathcal{L},\mathcal{U})\|_{\,2} \leq \eta \right\} = \frac{[|\boldsymbol{v}_{\mathcal{L}\mathcal{U}}^T||\boldsymbol{v}_{\mathcal{L}\mathcal{U}}| + |\boldsymbol{w}_{\mathcal{L}\mathcal{U}}^T||\boldsymbol{w}_{\mathcal{L}\mathcal{U}}|]^{1/2}}{|\lambda| \, |\boldsymbol{y}^*\boldsymbol{x}|} \, .$$

Observe that we set 1's and 0's perturbations equal to zero.

Next, we derive an expression of $relcond(\lambda;\mathcal{L},\mathcal{U})$ that is more convenient for numerical computations. So, consider,

$$\mathcal{U} = \mathrm{diag}(\mathfrak{u})\left(I + \mathring{\mathcal{U}}\right) \qquad (36)$$

where

$$\mathring{\mathcal{U}} = \begin{pmatrix} 0 & \mathfrak{u}_1^{-1} & & & \\ & 0 & \mathfrak{u}_2^{-1} & & \\ & & \ddots & \ddots & \\ & & & 0 & \mathfrak{u}_{n-1}^{-1} \\ & & & & 0 \end{pmatrix}.$$

and use,

$$\boldsymbol{y}^*\mathcal{L}\mathcal{U} = \lambda\boldsymbol{y}^*, \quad \mathcal{U}\boldsymbol{x} = \mathcal{L}^{-1}\boldsymbol{x}\lambda, \quad \lambda \neq 0, \qquad (37)$$

to cancel the factor $|\lambda|$ in the denominator of $relcond_2(\lambda;\mathcal{L},\mathcal{U})$. Substitute (36) in the first equation in (37), multiply both sides by $\mathrm{diag}(\boldsymbol{x})$ and take absolute values,

$$|(\boldsymbol{y}^*\mathcal{L})\mathrm{diag}(\mathfrak{u})\mathrm{diag}(\boldsymbol{x})| = |\lambda||\boldsymbol{y}^*\left(I + \mathring{\mathcal{U}}\right)^{-1}\mathrm{diag}(\boldsymbol{x})|, \quad \lambda \neq 0 \qquad (38)$$

Multiply both sides of the second equation of (37) by $\mathring{\mathcal{L}}^T \mathrm{diag}(\boldsymbol{y})H$, then obtain its absolute value,

$$|\mathring{\mathcal{L}}^T \mathrm{diag}(\boldsymbol{y})H(\mathcal{U}\boldsymbol{x})| = |\lambda||\mathring{\mathcal{L}}^T \mathrm{diag}(\boldsymbol{y})H\mathcal{L}^{-1}\boldsymbol{x}|, \quad \lambda \neq 0. \tag{39}$$

If $\mathcal{LU}$ exists, is unique and is J is singular, only $\mathfrak{u}_n$, may be zero. But, consider that $\mathfrak{u}_n$ does not appear in $\mathring{\mathcal{U}}$. Substitute the expressions in (38,39) into (35) and cancel $|\lambda|$ ($\neq 0$) to find

$$\begin{aligned} |\boldsymbol{y}^*\boldsymbol{x}|\| \, relgrad_{\mathcal{LU}}(\lambda)\|_2 = \big[& |\boldsymbol{y}^* \left(I + \mathring{\mathcal{U}}\right)^{-1} \mathrm{diag}(\boldsymbol{x})||\boldsymbol{y}^* \left(I + \mathring{\mathcal{U}}\right)^{-1} \mathrm{diag}(\boldsymbol{x})|^T + \\ & + |\mathring{\mathcal{L}}^T \mathrm{diag}(\boldsymbol{y})H\mathcal{L}^{-1}\boldsymbol{x}||\mathring{\mathcal{L}}^T \mathrm{diag}(\boldsymbol{y})H\mathcal{L}^{-1}\boldsymbol{x}|^T \big]^{1/2}. \end{aligned} \tag{40}$$

For the cost of solving two bidiagonal linear systems

$$\boldsymbol{z}^* \left(I + \mathring{\mathcal{U}}\right) = \boldsymbol{y}^* \text{ for } \boldsymbol{z}^* \quad \text{and} \quad \boldsymbol{x} = \mathcal{L}H^T\boldsymbol{s} \text{ for } \boldsymbol{s}$$

and for the cost of multiplying by two diagonal and one bidiagonal matrices,

$$\boldsymbol{v}^* = \boldsymbol{z}^* \mathrm{diag}(\boldsymbol{x}) \quad \text{and} \quad \boldsymbol{w} = \mathring{\mathcal{L}}^T \mathrm{diag}(\boldsymbol{y})\boldsymbol{s}$$

We end up by obtaining an expression of the relative condition number for $J = \mathcal{LU}$, more convenient for computations

$$\boxed{relcond_2(\lambda; \mathcal{L}, \mathcal{U}) := \frac{\left[|\boldsymbol{v}|^T|\boldsymbol{v}| + |\boldsymbol{w}|^T|\boldsymbol{w}|\right]^{1/2}}{|\boldsymbol{y}^*\boldsymbol{x}|}.} \tag{41}$$

9.4 Other representations

Section 7 of this work was devoted to two factorizations $J = \mathcal{LU}$ and $\Delta T = \Delta LDL^T$. Now, we present a close factorization, $T = \bar{L}\Omega\bar{L}^T$, which was also used in [12], that attemps to be as similar as possible to the widely known Cholesky factorization of symmetric positive definite matrices. We define Ω as a signature matrix.

9.4.1 Representation 3 - L, D (Δ) representation of ΔT

So, we show first the "*relcond*" case for $T = LDL^T$. Assume that the symmetric matrix T permit triangular factorization $T = LDL^T$ and recall that in the section "Factored forms of tridiagonal matrices" 7 we fixed T to be $T = \Delta B$, with B, the balanced matrix, defined in (12). Then, by taking the inverse of Δ (Δ, itself, i.e. $I = \Delta^2$), we have,

$$B = \Delta T = \Delta LDL^T$$

and this leads to obtain the eigenvalues and eigenvectos throughout the following traditional equations,

$$\Delta LDL^T\boldsymbol{x} = \boldsymbol{x}\lambda, \quad \boldsymbol{y}^*\Delta LDL^T = \lambda\boldsymbol{y}^*, \quad \text{where } \lambda \neq 0 \text{ is simple .}$$

Further developments in this section concern the study of the sensitivity of simple eigenvalues of B under perturbations of the type $\|rel\delta(L, D)\|_2 \leq \eta$, where

$$rel\delta(L, D) = \left(\frac{\delta l_1}{l_1}, \ldots, \frac{\delta l_{n-1}}{l_{n-1}}, \frac{\delta d_1}{d_1}, \ldots, \frac{\delta d_n}{d_n}\right) \tag{42}$$

where $l_1, \ldots, l_{n-1}$ are the subdiagonal entries of L and $d_1, \ldots, d_{n-1}$ are the diagonal entries of D. These developments run in parallel to those in section 9.2 for perturbations in the entries of the matrix and to those in section 9.3 for perturbations of the nontrivial entries of the factors $\mathcal{L}$ and $\mathcal{U}$. So, many details are omitted, since they are very similar to those in sections 9.2 and 9.3.
Some more ingredients for developing an expression of $relcond_2(\lambda; L, D)$ are still missed. Recall, from section 7, the definition $D = \mathrm{diag}(d_1, d_2, \ldots, d_n)$, and let us introduce the sum $L = I + \mathring{L}$ where,

$$\mathring{L} = \begin{pmatrix} 0 & & & & \\ l_1 & 0 & & & \\ & \ddots & \ddots & & \\ & & l_{n-2} & 0 & \\ & & & l_{n-1} & 0 \end{pmatrix}. \tag{43}$$

Regarding the eigenvectors, $\boldsymbol{x}$ and $\boldsymbol{y}^*$ of balanced matrices of section 6.1 (B-form), we present now their relation ($\boldsymbol{x}$ determine $\boldsymbol{y}^*$) . Hence, transposing $\Delta L D L^T \boldsymbol{x} = \boldsymbol{x}\lambda$ and inserting $I = \Delta^2$ yields

$$\left(\boldsymbol{x}^T \Delta\right)\left(\Delta L D L^T\right) = \lambda\left(\boldsymbol{x}^T \Delta\right)$$

so, equivalently,

$$\boldsymbol{y}^* \Delta L D L^T = \lambda \boldsymbol{y}^*$$

and get a relation between left and right eigenvectors for a balanced, unreduced, tridiagonal matrix of the form of B,

$$\boxed{\boldsymbol{y}^* = \boldsymbol{x}^T \Delta}$$

So, for λ simple, we have $0 \neq \boldsymbol{y}^* \boldsymbol{x} = \boldsymbol{x}^T \Delta \boldsymbol{x}$.
Following the analysis of (35) but using partial derivatives of λ with respect to $d_1, \ldots, d_n$ and $l_1, \ldots, l_{n-1}$ for this case, we find

$$\begin{aligned} |\lambda| \|\boldsymbol{x}^T \Delta \boldsymbol{x}\| \, relgrad_{L,D}(\lambda)\|_2 &= \left[\sum_{j=1}^{n} \left|d_j^2 [(L^T \boldsymbol{x})_j]^4\right| + 4 \sum_{j=1}^{n-1} |(\boldsymbol{x}^T \mathring{L})_j^2 (DL^T \boldsymbol{x})_j^2|\right]^{1/2} \\ &= \left[|(DL^T \boldsymbol{x})^T \mathring{R}| |\mathring{R}^T (DL^T \boldsymbol{x})| + 4 |(DL^T \boldsymbol{x})^T H^T \mathrm{diag}(\boldsymbol{x}) \mathring{L}| |\mathring{L}^T \mathrm{diag}(\boldsymbol{x}) H (DL^T \boldsymbol{x})|\right]^{1/2} \\ &= \left[|DL^T \boldsymbol{x}|^T \left(|\mathring{R}\mathring{R}^T| + 4 |H^T \mathrm{diag}(\boldsymbol{x}) \mathring{L} \mathring{L}^T \mathrm{diag}(\boldsymbol{x}) H|\right) |DL^T \boldsymbol{x}|\right]^{1/2}. \end{aligned} \tag{44}$$

where ,

$$\mathring{R} = \begin{pmatrix} x_1 + l_1 x_2 & & & \\ & \ddots & & \\ & & x_{n-1} + l_{n-1} x_n & \\ & & & x_n \end{pmatrix} = \mathrm{diag}(L^T x) \quad \text{and} \quad H^T \mathrm{diag}(\boldsymbol{x}) \mathring{L} = \mathrm{diag}(\boldsymbol{x}^T \mathring{L}) \tag{45}$$

So, we can summarize this result in Theorem 9.4.

Theorem 9.4. *Let B be a balanced unreduced real tridiagonal matrix that permits triangular factorization $B = \Delta L D L^T$ with factors as in* (15) *. Moreover, let $\mathring{L} = L - I$, $\mathring{R}, H$ and $\mathrm{diag}(\boldsymbol{x})$ be matrices defined as in (43),(45), (30) and (32). Also denote by $rel\delta(L, D)$ the variation vector in (42). In*

addition, let $\lambda \neq 0$ be a simple eigenvalue of B with right eigenvector $\boldsymbol{x}$. And let $\boldsymbol{v}_{LD}^T = (DL^T\boldsymbol{x})^T\mathring{R}$, and $\boldsymbol{w}_{LD}^T = (DL^T\boldsymbol{x})^T H^T diag(\boldsymbol{x})\mathring{L}$. Then relcond$(\lambda; L, D) :=$ $\| \, relgrad_{L,D}(\lambda)\|_2$ is equal to

$$relcond_2(\lambda; L, D) = \lim_{\eta\to 0} \sup\left\{ \frac{|\delta\lambda|}{\eta|\lambda|} \; : \; (\lambda+\delta\lambda) \text{ is an eigenvalue of } \Delta(L+\delta L)(D+\delta D)(L+\delta L)^T, \right.$$
$$\left. ||rel\delta(L,D)||_2 \leq \eta \right\} = \frac{\left[|\boldsymbol{v}_{LD}^T||\boldsymbol{v}_{LD}| + 4|\boldsymbol{w}_{LD}^T||\boldsymbol{w}_{LD}|\right]^{1/2}}{|\lambda|\,|\boldsymbol{x}^T\Delta\boldsymbol{x}|}.$$

Observe that we set to zero 1's and 0's perturbations of L and D .

Now, we express *relcond*$(\lambda; L, D)$ in a more adequate form for computations. To this purpose, we can remove a factor $|\lambda|$ on the right side of (44) by means of,

$$DL^T\boldsymbol{x} = L^{-1}\Delta\boldsymbol{x}\lambda. \tag{46}$$

Equation (46) is derive from the eigenequation $\Delta LDL^T\boldsymbol{x} = \boldsymbol{x}\lambda$. Then, the resultant equations that act replacing (46) in $\boldsymbol{v}_{LD}$ and $\boldsymbol{w}_{LD}$ are,

$$|\mathring{R}^T(DL^T\boldsymbol{x})| = |\lambda||\mathring{R}^T L^{-1}\Delta\boldsymbol{x}| = |\lambda||\mathring{R}L^{-1}\Delta\boldsymbol{x}|, \quad |\mathring{L}^T\text{diag}(\boldsymbol{x})H(DL^T\boldsymbol{x})| = |\lambda||\mathring{L}^T\text{diag}(\boldsymbol{x})HL^{-1}\Delta\boldsymbol{x}|. \tag{47}$$

Note that $\mathring{R}^T = \mathring{R}$ because both matrices are diagonal. Substitute (46) in (44) and cancel $|\lambda|$ $(\neq 0)$ to obtain

$$|\boldsymbol{x}^T\Delta\boldsymbol{x}|\| \, relgrad_{L,D}(\lambda)\|_2 = \left[|L^{-1}\Delta\boldsymbol{x}|^T \left(|\mathring{R}\mathring{R}^T| + 4|H^T\text{diag}(\boldsymbol{x})\mathring{L}(\mathring{L}^T\text{diag}(\boldsymbol{x})H)|\right)|L^{-1}\Delta\boldsymbol{x}|\right]^{\frac{1}{2}}.$$

Once λ is cancelled, the resultant expression may be reduced to the computations of a few systems. So , we obtain $relcond_2(\lambda; L, D)$ for the cost of solving one bidiagonal linear system (L is unit bidiagonal),

$$\boldsymbol{x} = \Delta L\boldsymbol{h} \text{ for } \boldsymbol{h},$$

and for the cost of 3 products: two diagonal-vector products and one bidiagonal-vector product,

$$\boldsymbol{v} = \mathring{R}\boldsymbol{h}, \quad \boldsymbol{w} = \mathring{L}^T diag(\boldsymbol{x})H\boldsymbol{h}$$

we obtain the following expression of the relative condition number for $\Delta T = \Delta LDL^T$.

$$\boxed{relcond_2(\lambda; L, D) := \frac{\left[|\boldsymbol{v}^T||\boldsymbol{v}| + 4|\boldsymbol{w}^T||\boldsymbol{w}|\right]^{\frac{1}{2}}}{|\boldsymbol{x}^T\Delta\boldsymbol{x}|}.} \tag{48}$$

9.4.2 Representation 4 - $\bar{L}, \Omega$ representation of T

The second representation we consider in this section is the factorization $T = \bar{L}\Omega\bar{L}^T$ where $\bar{L} = L|D|^{1/2}$ is lower bidiagonal, $D = |D|^{1/2}\Omega|D|^{1/2}$, $\Omega = \text{diag}(\text{sign}(d_i))$ with $\text{sign}(d_n) = 1$ if $d_n = 0$. A potential advantage of this factorization is the closest to the Cholesky factorization that we can get [12] . Thus,

$$\Delta\bar{L}\Omega\bar{L}^T\boldsymbol{x} = \boldsymbol{x}\lambda. \tag{49}$$

According to [12],there is a related eigenproblem dual to (49):

$$\Omega\bar{L}^T\Delta\bar{L}\boldsymbol{z} = \boldsymbol{z}\lambda \tag{50}$$

obtained by taking a LU transform of (49) by taking one step of the $L\Omega$ algorithm, which preserves the eigenvalues but changes the eigenvectors. This gives, for us, the most elegant (symmetric) form of our problem as

$$\bar{L}\Omega\bar{L}^T\boldsymbol{x} = \Delta\boldsymbol{x}\lambda, \quad \bar{L}^T\Delta\bar{L}\boldsymbol{z} = \Omega\boldsymbol{z}\lambda,$$

with just a single bidiagonal matrix $\bar{L}$". We give, as in previous sections, some relevant definitions,

$$rel\delta\bar{L} = \big(\frac{\delta l_{11}}{l_{11}}, \ldots, \frac{\delta l_{nn}}{l_{nn}}, \frac{\delta l_{21}}{l_{21}}, \ldots, \frac{\delta l_{nn-1}}{l_{nn-1}}\big), \tag{51}$$

$$\mathring{\bar{L}} = \begin{pmatrix} 0 & & & & \\ l_{21} & 0 & & & \\ & l_{32} & 0 & & \\ & & \ddots & \ddots & \\ & & & l_{nn-1}0 \end{pmatrix}, \quad \bar{L} = \begin{pmatrix} l_{11} & & & & \\ l_{21} & l_{22} & & & \\ & l_{32} & l_{33} & & \\ & & \ddots & \ddots & \\ & & & l_{nn-1}l_{nn} \end{pmatrix} \tag{52}$$

and

$$\mathrm{diag}(\bar{L}) = \mathrm{diag}(l_{11}, \ldots, l_{jj}, \ldots, l_{nn}) \tag{53}$$

If in a similar way, as in previous sections, we differentiate λ with respect to the entries $(1,1), \ldots, (n,n)$ and $(2,1), \ldots, (n, n-1)$ of $\bar{L}$, this lead to the *relative condition number* for $\Delta T = \Delta\bar{L}\Omega\bar{L}^T$.

Theorem 9.5. *Let B be a balanced unreduced real tridiagonal matrix that permits a triangular factorization $B = \Delta\bar{L}\Omega\bar{L}^T$, where Δ and Ω are diagonal signature matrices and $\bar{L}$ is a lower bidiagonal matrix. Denote $\bar{L} = \mathrm{diag}(\bar{L}) + \mathring{\bar{L}}$,where $\mathrm{diag}(\bar{L})$ and $\mathring{\bar{L}}$ are matrices defined in (52) and (53), and let $\mathrm{diag}(\boldsymbol{x})$ and H be matrices defined in (32) and (30). In addition, consider the relative perturbation vector $rel\delta\bar{L}$ defined in (51). Let $\lambda \neq 0$ be a simple eigenvalue of B with right eigenvector $\boldsymbol{x}$. And let $\boldsymbol{v}_\Omega^T = (\bar{L}^T\boldsymbol{x})^T\mathrm{diag}(\boldsymbol{x})\mathrm{diag}(\bar{L})$, and $\boldsymbol{w}_\Omega^T = (\bar{L}^T\boldsymbol{x})^T H^T\mathrm{diag}(\boldsymbol{x})\mathring{\bar{L}}$. Then $relcond(\lambda;\bar{L}) := \| relgrad_{\bar{L}}(\lambda)\|_2$ is equal to*

$$\begin{aligned} relcond_2(\lambda;\bar{L}) &= \lim_{\eta\to 0}\sup\left\{\frac{|\delta\lambda|}{\eta|\lambda|} \,:\, (\lambda+\delta\lambda) \text{ is an eigenvalue of } \Delta(\bar{L}+\delta\bar{L})\Omega(\bar{L}+\delta\bar{L})^T,\ \ ||rel\delta\bar{L}||_2 \le \eta\right\} \\ &= \frac{2\,\left[|\boldsymbol{v}_\Omega^T|\,|\boldsymbol{v}_\Omega| + |\boldsymbol{w}_\Omega^T||\boldsymbol{w}_\Omega|\right]^{1/2}}{|\lambda|\,|\boldsymbol{x}^T\Delta\boldsymbol{x}|}\,. \end{aligned}$$

Observe that we set to zero perturbations of the 1's and the 0's of $\bar{L}$.

For computational purposes, it is more convenient to express $relcond(\lambda;\bar{L})$ as follows:

$$relcond_2(\lambda;\bar{L}) := \frac{2(|\boldsymbol{v}^T||\boldsymbol{v}| + |\boldsymbol{w}^T||\boldsymbol{w}|)}{|\boldsymbol{x}^T\Delta\boldsymbol{x}|},$$

where

$$\boldsymbol{x} = \Delta\bar{L}\Omega\boldsymbol{q} \quad \text{for } \boldsymbol{q}.$$

$$\boldsymbol{v} = \mathrm{diag}(\bar{L})\mathrm{diag}(\boldsymbol{x})\boldsymbol{q} \quad \boldsymbol{w} = \mathring{\bar{L}}^T diag(\boldsymbol{x})H\boldsymbol{q}$$

9.5 Equivalence of $relcond_2(\lambda;\mathcal{L},\mathcal{U})$, $relcond_2(\lambda;L,D)$, $relcond_2(\lambda;\bar{L})$

We show in the following lemma that the three condition numbers of the factored forms we have considered before are equivalent .

Lemma 9.2.

$$\frac{1}{\sqrt{2}}\, relcond_2(\lambda;\mathcal{L},\mathcal{U}) \le relcond_2(\lambda;L,D) \le \sqrt{6}\, relcond_2(\lambda;\mathcal{L},\mathcal{U})$$

and

$$\frac{1}{2\sqrt{2}}\, relcond_2(\lambda;\bar{L}) \le relcond_2(\lambda;L,D) \le \frac{\sqrt{6}}{2}\, relcond_2(\lambda;\bar{L}).$$

Proof. For S invertible and diagonal, let $J = SBS^{-1} = S\Delta TS^{-1}$ and consider the triangular factorizations $J = \mathcal{L}\mathcal{U}$ and $B = \Delta T = \Delta LDL^T$. Then we have

$$\mathcal{L}\mathcal{U} = S\Delta LDL^TS^{-1}.$$

Uniqueness of LU factorization guarantees that

$$\mathcal{L} = S\Delta LS^{-1}\Delta \quad \text{and} \quad \mathcal{U} = \Delta SDL^TS^{-1}$$

and

$$\mathring{\mathcal{L}} = S\Delta\mathring{L}S^{-1}\Delta \quad \text{and} \quad \text{diag}(\mathfrak{u}) = \Delta D.$$

with $\mathring{L}$ defined in 43 and $\mathring{\mathcal{L}}, \text{diag}(\mathfrak{u})$ defined in section 9.3. .
Let $\boldsymbol{x}_B$ and $\boldsymbol{x}_J$ be the right eigenvectors of the matrices B and J respectively, corresponding to the simple eigenvalue λ. Since they are related by doing similarity transformations, it is satisfied,

$$\boldsymbol{x}_B = S^{-1}\boldsymbol{x}_J \quad \text{and} \quad \boldsymbol{x}_B^T\Delta = \boldsymbol{y}_B^* = \boldsymbol{y}_J^*S. \tag{54}$$

then, since S is diagonal, it is also satisfied

$$\text{diag}(\boldsymbol{x}_B) = S^{-1}\text{diag}(\boldsymbol{x}_J) \quad \text{and} \quad \text{diag}(\boldsymbol{x}_B^T)\Delta = \text{diag}(\boldsymbol{y}_B^*) = \text{diag}(\boldsymbol{y}_J^*)S. \tag{55}$$

Consequently, $|\boldsymbol{y}_J^*\boldsymbol{x}_J| = |\boldsymbol{y}_B^*\boldsymbol{x}_B| = |\boldsymbol{x}_B^T\Delta\boldsymbol{x}_B|$ and, from (35),

$$|\lambda||\boldsymbol{y}_J^*\boldsymbol{x}_J|\, relcond_2(\lambda;\mathcal{L},\mathcal{U}) =$$
$$= \left[|(\boldsymbol{y}_J^*\mathcal{L})\text{diag}(\mathfrak{u})\text{diag}(\boldsymbol{x}_J)||diag(\boldsymbol{x}_J)\text{diag}(\mathfrak{u})(\boldsymbol{y}_J^*\mathcal{L})^T| + |(\mathcal{U}\boldsymbol{x}_J)^TH^T\text{diag}(\boldsymbol{y}_J)\mathring{\mathcal{L}}||\mathring{\mathcal{L}}^T\text{diag}(\boldsymbol{y}_J)H(\mathcal{U}\boldsymbol{x}_J)|\right]^{1/2},$$

This expression is the expanded version of Theorem 9.3 . It is convenient to focus first at $|(\boldsymbol{y}_J^*\mathcal{L})\text{diag}(\mathfrak{u})\text{diag}(\boldsymbol{x}_J)|$ and notice that the other factor in the first term is its transpose. So, substituting $\boldsymbol{x}_J$, $\boldsymbol{y}_J$, $\text{diag}(\boldsymbol{x}_J)$ and $\text{diag}(\boldsymbol{y}_J)$ from (54) and (55) in the expresion above, we obtain,

$$|\lambda||\boldsymbol{x}_B^T\Delta\boldsymbol{x}_B|\, relcond_2(\lambda;\mathcal{L},\mathcal{U}) =$$
$$= \left[|(DL^T\boldsymbol{x}_B)^T\text{diag}(\boldsymbol{x}_B)||\text{diag}(\boldsymbol{x}_B)(DL^T\boldsymbol{x}_B)| + |(DL^T\boldsymbol{x}_B)^TH^T\text{diag}(\boldsymbol{x}_B)\mathring{L}||\mathring{L}^T\text{diag}(\boldsymbol{x}_B)H(DL^T\boldsymbol{x}_B)|\right]^{1/2} =$$
$$= \left[|(DL^T\boldsymbol{x}_B)^T|(|(\text{diag}^2(\boldsymbol{x}_B)| + |H^T\text{diag}(\boldsymbol{x}_B)\mathring{L}||\mathring{L}^T\text{diag}(\boldsymbol{x}_B)H|)|DL^T\boldsymbol{x}_B)|\right]^{1/2}.$$

We compare the expression above with (44) and (45) in the derivation of $relcond_2(\lambda; L, D)$,which we express as follows:

$$
\begin{aligned}
&|\lambda||\boldsymbol{x}_B^T \Delta \boldsymbol{x}_B|\, relcond_2(\lambda; L, D) = \\
&= \big[|(DL^T\boldsymbol{x}_B)^T \mathrm{diag}(L^T\boldsymbol{x}_B)||\mathrm{diag}(L^T\boldsymbol{x}_B)(DL^T\boldsymbol{x}_B)| + \\
&+ 4|(DL^T\boldsymbol{x}_B)^T H^T \mathrm{diag}(\boldsymbol{x}_B)\mathring{L}||\mathring{L}^T \mathrm{diag}(\boldsymbol{x}_B) H(DL^T\boldsymbol{x}_B)|\big]^{1/2} = \\
&= \left[|(DL^T\boldsymbol{x}_B)^T| \left(|\mathrm{diag}^2(L^T\boldsymbol{x}_B)| + 4|H^T \mathrm{diag}(\boldsymbol{x}_B)\mathring{L}||\mathring{L}^T \mathrm{diag}(\boldsymbol{x_B}) H)|\right) |(DL^T\boldsymbol{x}_B)|\right]^{1/2} .
\end{aligned}
$$

The first difference from the comparison of the two condition numbers above is that $\mathrm{diag}^2(\boldsymbol{x}_B)$ of $relcond_2(\lambda; \mathcal{L}, \mathcal{U})$ is distinct from $\mathrm{diag}^2(L^T\boldsymbol{x}_B)$ of $relcond_2(\lambda; L, D)$. And the second difference found is that $\mathrm{diag}^2(\boldsymbol{x}_B^T\mathring{L})$ (use (45)) is multiplied by one in $relcond_2(\lambda; \mathcal{L}, \mathcal{U})$, whereas in $relcond_2(\lambda; L, D)$ is multiplied by four. We can obtain an equivalence between both condition numbers as follows. First, we transform $\mathrm{diag}^2(L^T\boldsymbol{x}_B)$, so we have,

$$
|\mathrm{diag}^2(\boldsymbol{x}_B^T L)| = |\mathrm{diag}^2(\boldsymbol{x}_B^T(I + \mathring{L}))| = |\mathrm{diag}(\boldsymbol{x}_B^T) + \mathrm{diag}(\boldsymbol{x}_B^T\mathring{L})|^2, \tag{56}
$$

then we include this result in the expression that affect $relcond_2(\lambda; L, D)$,

$$
\begin{aligned}
|\mathrm{diag}(\boldsymbol{x}_B) + \mathrm{diag}(\boldsymbol{x}_B^T\mathring{L})|^2 + 4|\mathrm{diag}(\boldsymbol{x}_B^T\mathring{L})|^2 \leq \left(|\mathrm{diag}(\boldsymbol{x}_B)| + |\mathrm{diag}(\boldsymbol{x}_B^T\mathring{L})|\right)^2 + 4|\mathrm{diag}(\boldsymbol{x}_B^T\mathring{L})|^2 \leq \\
\leq |\mathrm{diag}(\boldsymbol{x}_B)|^2 + |\mathrm{diag}(\boldsymbol{x}_B^T\mathring{L})|^2 + 2|\mathrm{diag}(\boldsymbol{x}_B)||\mathrm{diag}(\boldsymbol{x}_B^T\mathring{L})| + 4|\mathrm{diag}(\boldsymbol{x}_B^T\mathring{L})|^2 \leq \\
\leq |\mathrm{diag}(\boldsymbol{x}_B)|^2 + 5|\mathrm{diag}(\boldsymbol{x}_B^T\mathring{L})|^2 + 2|\mathrm{diag}(\boldsymbol{x}_B)||\mathrm{diag}(\boldsymbol{x}_B^T\mathring{L})| \leq \\
\leq |\mathrm{diag}(\boldsymbol{x}_B)|^2 + 5|\mathrm{diag}(\boldsymbol{x}_B^T\mathring{L})|^2 + |\mathrm{diag}(\boldsymbol{x}_B)|^2 + |\mathrm{diag}(\boldsymbol{x}_B^T\mathring{L})|^2 \leq \\
\leq 6\left(|\mathrm{diag}(\boldsymbol{x}_B)|^2 + |\mathrm{diag}(\boldsymbol{x}_B^T\mathring{L})|^2\right)
\end{aligned}
$$

using the inequality $2ab \leq a^2 + b^2$ for $a, b \in \mathbb{R}$. Therefore,

$$
relcond_2(\lambda; L, D) \leq \sqrt{6}\, relcond_2(\lambda; \mathcal{L}, \mathcal{U}).
$$

Also, since

$$
\begin{aligned}
|\mathrm{diag}(\boldsymbol{x}_B)|^2 + |\mathrm{diag}(\boldsymbol{x}_B^T\mathring{L})|^2 = |\mathrm{diag}(\boldsymbol{x}_B^T(L - \mathring{L}))|^2 + |\mathrm{diag}(\boldsymbol{x}_B^T\mathring{L})|^2 \leq \\
\leq \left(|\mathrm{diag}(\boldsymbol{x}_B^T L)| + |\mathrm{diag}(\boldsymbol{x}_B^T\mathring{L})|\right)^2 + |\mathrm{diag}(\boldsymbol{x}_B^T\mathring{L})|^2 \leq \\
\leq |\mathrm{diag}(\boldsymbol{x}_B^T L)|^2 + |\mathrm{diag}(\boldsymbol{x}_B^T\mathring{L})|^2 + 2|\mathrm{diag}(\boldsymbol{x}_B^T L)||\mathrm{diag}(\boldsymbol{x}_B^T\mathring{L})| + |\mathrm{diag}(\boldsymbol{x}_B^T\mathring{L})|^2 \leq \\
\leq |\mathrm{diag}(\boldsymbol{x}_B^T L)|^2 + 2|\mathrm{diag}(\boldsymbol{x}_B^T L)||\mathrm{diag}(\boldsymbol{x}_B^T\mathring{L})| + 2|\mathrm{diag}(\boldsymbol{x}_B^T\mathring{L})|^2 \leq \\
\leq 2|\mathrm{diag}(\boldsymbol{x}_B^T L)|^2 + 3|\mathrm{diag}(\boldsymbol{x}_B^T\mathring{L})|^2
\end{aligned}
$$

which is less than $2\left(|\mathrm{diag}(\boldsymbol{x}_B^T L)|^2 + 4|\mathrm{diag}(\boldsymbol{x}_B^T\mathring{L})|^2\right)$. So, the equivalence can be expressed as:

$$
\frac{1}{\sqrt{2}}\, relcond_2(\lambda; \mathcal{L}, \mathcal{U}) \leq relcond_2(\lambda; L, D).
$$

The second part of the lemma is derived next. Recall,

$$\bar{L} = L|D|^{1/2} \quad D = |D|^{1/2}\Omega|D|^{1/2} \tag{57}$$

and consider the right eigenvectors, $\boldsymbol{x}$, of $B = \Delta L D L^T = \Delta \bar{L}\Omega\bar{L}^T$. Recall that, acording to Theorem ,

$$\begin{aligned} &|\lambda|\,|\boldsymbol{x}^T\Delta\boldsymbol{x}|\; relcond_2(\lambda;\bar{L}) = \\ &= 2\left[|(\bar{L}^T\boldsymbol{x})^T diag(\boldsymbol{x}) diag(\bar{L})|\,|diag(\bar{L}) diag(\boldsymbol{x})(\bar{L}^T\boldsymbol{x})|+\right. \\ &\left.+\,|(\bar{L}^T\boldsymbol{x})^T H^T diag(\boldsymbol{x})\mathring{\bar{L}}||\mathring{\bar{L}}^T diag(\boldsymbol{x}) H(\bar{L}^T\boldsymbol{x})|\right]^{1/2}. \end{aligned} \tag{58}$$

So , consider

$$\mathrm{diag}(\bar{L}) = |D|^{1/2} \quad \text{and} \quad \mathring{\bar{L}} = \mathring{L}|D|^{1/2} \tag{59}$$

and recall (45), substituting (57) in (58) and applying 59, we have

$$\begin{aligned} &|\lambda|\,|\boldsymbol{x}^T\Delta\boldsymbol{x}|\; relcond(\lambda;\bar{L}) = \\ &= 2\left[|(DL^T\boldsymbol{x})^T \mathrm{diag}(\boldsymbol{x})||\mathrm{diag}(\boldsymbol{x})^T(DL^T\boldsymbol{x})|+\right. \\ &\left.+\,|(DL^T\boldsymbol{x})^T H^T diag(\boldsymbol{x})\mathring{L}||\mathring{L}^T diag(\boldsymbol{x}) H(DL^T\boldsymbol{x})|\right]^{1/2}. \end{aligned} \tag{60}$$

$$\tag{61}$$

Compared with,

$$\begin{aligned} &|\lambda||\boldsymbol{x}^T\Delta\boldsymbol{x}|\; relcond(\lambda;L,D) = \\ &= \left[|(DL^T\boldsymbol{x})^T \mathrm{diag}(L^T\boldsymbol{x})||\mathrm{diag}(L^T\boldsymbol{x})^T(DL^T\boldsymbol{x})|+\right. \\ &\left.+\,4|(DL^T\boldsymbol{x})^T H^T diag(\boldsymbol{x})\mathring{L}||\mathring{L}^T diag(\boldsymbol{x}) H(DL^T\boldsymbol{x})|\right]^{1/2}. \end{aligned} \tag{62}$$

$$\tag{63}$$

to conclude that,

$$|\mathrm{diag}(\boldsymbol{x}^T L)|^2 = |\mathrm{diag}(\boldsymbol{x}^T(I+\mathring{L}))|^2 = |\mathrm{diag}(\boldsymbol{x}) + \mathrm{diag}(\boldsymbol{x}^T\mathring{L})|^2 \tag{64}$$

so, as in the first part of the proof, we have,

$$\begin{gathered} |\mathrm{diag}(\boldsymbol{x}) + \mathrm{diag}(\boldsymbol{x}^T\mathring{L})|^2 + 4|\mathrm{diag}(\boldsymbol{x}^T\mathring{L})|^2 \leq \\ \leq \left(|\mathrm{diag}(\boldsymbol{x})| + |\mathrm{diag}(\boldsymbol{x}^T\mathring{L})|\right)^2 + 4|\mathrm{diag}(\boldsymbol{x}^T\mathring{L})|^2 \\ \leq |\mathrm{diag}(\boldsymbol{x})|^2 + |\mathrm{diag}(\boldsymbol{x}^T\mathring{L})|^2 + 2|\mathrm{diag}(\boldsymbol{x})||\mathrm{diag}(\boldsymbol{x}^T\mathring{L})| + 4|\mathrm{diag}(\boldsymbol{x}^T\mathring{L})|^2 \leq \\ \leq |\mathrm{diag}(\boldsymbol{x})|^2 + 5|\mathrm{diag}(\boldsymbol{x}^T\mathring{L})|^2 + 2|\mathrm{diag}(\boldsymbol{x})||\mathrm{diag}(\boldsymbol{x}^T\mathring{L})| \leq \\ \leq |\mathrm{diag}(\boldsymbol{x})|^2 + 5|\mathrm{diag}(\boldsymbol{x}^T\mathring{L})|^2 + |\mathrm{diag}(\boldsymbol{x})|^2 + |\mathrm{diag}(\boldsymbol{x}^T\mathring{L})|^2 \leq \\ \leq 6\left(|\mathrm{diag}(\boldsymbol{x})|^2 + |\mathrm{diag}(\boldsymbol{x}^T\mathring{L})|^2\right) \end{gathered}$$

and

$$relcond(\lambda;L,D) \leq \frac{\sqrt{6}}{2}\, relcond(\lambda;\bar{L}).$$

We ommit the last proof of equivalence given that the expression we computed in the case of $\mathcal{LU}$ factorization is equally valid for the case of $\bar{L}$ factorization. We invite the reader to do the calculations by himself.

$$\frac{1}{2\sqrt{2}}\, relcond(\lambda;\bar{L}) \leq relcond(\lambda;L,D).$$

Important note: This lemma is based on lemma 6.6 of [12] , in which the equivalence bound for 1-norm were,

$$relcond(\lambda; \mathcal{L}, \mathcal{U}) \leq relcond(\lambda; L, D) \leq 3\, relcond(\lambda; \mathcal{L}, \mathcal{U})$$

and

$$\frac{1}{2}\, relcond(\lambda; \bar{L}) \leq relcond(\lambda; L, D) \leq \frac{3}{2}\, relcond(\lambda; \bar{L}).$$

The difference between the two lemmas is not very notable, since non of them depend on the size of the matrix, n. Althoug it seems 2-norm is more precise.

10 Sensitivity analysis for banded matrices

It is well known [36] that simple eigenvalues of any matrix A are differentiable functions of the entries of A. In particular this is true for banded matrices. This assertion implies that all the partial derivatives of eigenvalues with respect to the matrix entries exist. We will use the 1-norm defining eigenvalue condition numbers of banded matrices, but we will show in this section that the condition number for band matrices can be extended to tridiagonal. There is no specific definition for banded and another for tridiagonal as the title of this section suggests. For simplicity we have used the 1-norm in the definition of eigenvalue condition number for banded matrices , and the 2-norm in the definition for tridiagonal matrices since the 1-norm was already used in [12].
To obtain a condition number using 1-norm, define perturbations of the form :

$$|\delta a_{ij}| \leq \eta |a_{ij}|$$

where a_{ij} are the entries of A, so,

$$|\delta A| \leq \eta |A|, \qquad 0 < \eta \ll 1, \tag{65}$$

and

$$|\, rel\, \delta A| \leq \eta\, (1, 1, \ldots 1)^T. \tag{66}$$

which is the relative perturbation column vector. Then, the condition number with respect to perturbations of the entries is defined as :

$$\begin{aligned} relcond(\lambda; A) &:= \| relgrad_A(\lambda)\|_1 \\ &= \lim_{\eta \to 0} \sup \left\{ \frac{|\delta\lambda|}{\eta|\lambda|} \,:\, (\lambda + \delta\lambda) \text{ is an eigenvalue of } (A + \delta A),\ |\delta A| \leq \eta |A| \right\}. \end{aligned} \tag{67}$$

Moreover, if diagonals of (p,q) banded matrix A defined in (10) are in arrays (p+q+1 arrays). And we identify entries by subindices $k = 1, \ldots, p; \quad j = 1, \ldots, n$ for the lower diagonal arrays and $s = 1, \ldots, q; \quad j = 1, \ldots, n$, for the upper diagonal arrays. We can assert that matrix A is formed by a diagonal array a_{jj}, q upper diagonal arrays $a_{j,j+s}$ and p lower diagonal arrays $a_{j+k,j}$ for $j + s \leq n$ and $j + k \leq n$, which is used to say that no entries fall out of the matrix size. So that one can obtain, for simple eigenvalue λ, the set of partial derivatives $\left\{ \frac{\partial \lambda}{\partial a_{jj}}, \frac{\partial \lambda}{\partial a_{j+k,j}}, \frac{\partial \lambda}{\partial a_{j,j+s}} \right\}$ and obtain the absolute gradient vector over diagonals:

$$\begin{aligned} \text{grad}_A(\lambda) = (&\frac{\partial \lambda}{\partial a_{11}}, \ldots, \frac{\partial \lambda}{\partial a_{nn}}, \frac{\partial \lambda}{\partial a_{21}}, \ldots, \frac{\partial \lambda}{\partial a_{n,n-1}}, \ldots, \frac{\partial \lambda}{\partial a_{p+1,1}}, \ldots, \frac{\partial \lambda}{\partial a_{n,n-p}}, \\ &\frac{\partial \lambda}{\partial a_{12}}, \ldots, \frac{\partial \lambda}{\partial a_{n-1,n}}, \ldots, \frac{\partial \lambda}{\partial a_{1,q+1}}, \ldots, \frac{\partial \lambda}{\partial a_{n-q,n}})^T. \end{aligned} \tag{68}$$

Multiply this vector by infinitesimal absolute changes vector
$(\delta a_{11}, \ldots, \delta a_{nn}, \delta a_{21}, \ldots, \delta a_{n,n-1}, \delta a_{p+1,1}, \ldots, \delta a_{n,n-p}, \delta a_{12}, \ldots, \delta a_{n-1,n}, \ldots, \delta a_{1,q+1}, \ldots, \delta a_{n-q,n})^T =:$
$=: \delta A$, so we have

$$\delta\lambda = \operatorname{grad}_A(\lambda)^T \cdot \delta A + \textit{higher order terms (h.o.t.).} \tag{69}$$

To turn (69) into relative terms, relate $|\delta\lambda/\lambda|$ to $|\delta a_{ij}/a_{ij}|$, with

$$a_{ij} = (a_{11}, \ldots, a_{nn}, a_{12}, \ldots, a_{n-1,n}, \ldots, a_{1,q+1}, \ldots, a_{n-q,n}, \ldots, a_{21}, \ldots, a_{n,n-1}, \ldots, a_{p+1,1}, \ldots, a_{n,n-p})$$

For nonzeros terms, rewrite (69) as

$$\begin{aligned} \frac{\delta\lambda}{\lambda} &= \left(\frac{a_{11}}{\lambda}\frac{\partial\lambda}{\partial a_{11}}, \ldots, \frac{a_{n-q,n}}{\lambda}\frac{\partial\lambda}{\partial a_{n-q,n}}\right) \cdot \left(\frac{\delta a_{11}}{a_{11}}, \ldots, \frac{\delta a_{n-q,n}}{a_{n-q,n}}\right)^T + \textit{h.o.t.} \\ &=: \textit{relgrad}_A(\lambda)^T \cdot \textit{rel}\,\delta A + \textit{h.o.t.}, \end{aligned} \tag{70}$$

"defining the *relative gradient* and the *relative perturbation*. When a parameter vanishes we should omit the corresponding term in the inner product." [12]. Now, obtain the ∞-norm of the perturbation vector,

$$\| \textit{rel}\,\delta A\|_\infty \leq \eta.$$

and, since $|\boldsymbol{u}^T\boldsymbol{v}| \leq \|\boldsymbol{u}\|_1\|\boldsymbol{v}\|_\infty$ (Hölder inequality),

$$\left|\frac{\delta\lambda}{\lambda}\right| \leq \| \textit{relgrad}_A(\lambda)\|_1 \| \textit{rel}\,\delta A\|_\infty + \textit{h.o.t.} \leq \eta \| \textit{relgrad}_A(\lambda)\|_1 + \textit{h.o.t.} \tag{71}$$

So, the relative condition number can be defined as in Theorem 10.1.

Theorem 10.1. *For any unreduced real (p,q) band matrix A of the form of* (10)*, with perturbations of the form of* (65)*, whose relative perturbation vector, $\textit{rel}\,\delta A$, defined in (66), satisfy $\| \textit{rel}\,\delta A\|_\infty \leq \eta$, we can define a relative condition number as in (67) and attain the supremum upper bound for $\|\textit{relgrad}_A(\lambda)\|$, where $\textit{relgrad}_A(\lambda)$ is defined in (70) . So, formally , we can express the relative condition number of A using 1-norms as,*

$$\boxed{\textit{relcond}(\lambda; A) := \| \textit{relgrad}_A(\lambda)\|_1, \quad \lambda \neq 0.}$$

The proof of this theorem is very similar to the proof of theorem 9.1 and we omit it.We bring here some convenient comments of paper [12]. "There is no reason to expect $\textit{relcond}(\lambda; A) > 1$; any value in $[0, +\infty[$ could occur. Should $A \neq O$ be singular then appropriate independent relative changes to the entries will destroy singularity. So we set $\textit{relcond}(0; A) = \infty$. Our other representations will have finite values for $\| \textit{relgrad}(\lambda)\|_1$ and thus may define tiny eigenvalues to high relative accuracy in certain cases, a very desirable property".

10.1 Condition numbers for banded matrices

The case of banded matrices is similar to the tridiagonal case. We begin with the explicit expression of a banded matrix as,

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1,q+1} & & 0 \\ a_{21} & & & & a_{2,q+2} & \\ \vdots & & & & \ddots & \\ a_{p+1,1} & & & \ddots & & a_{n-q,n} \\ & a_{p+2,2} & & & & \vdots \\ & & \ddots & & & \\ 0 & & & a_{n,n-p} & \cdots & a_{nn} \end{bmatrix} \tag{72}$$

We say that $A \in \mathbb{R}^{n\times n}$ has upper bandwith q and lower bandwith p. Now, we infer an explicit expression for *relcond*$(\lambda; A)$ with respect relative perturbations of the entries. Thus, we consider the entries of A as independent variables. This is the challenge of measuring each variation of the eigenvalues with respect each perturbation of each entry of the matrix. This is done in the reference article [12] for the eigenvalue condition numbers of tridiagonal matrices. In addition, we can separate the upper from the lower part, being the upper part an analogue of C_i entries and the lower bandwith an analogue of b_i entries, in the tridiagonal matrix C appearing in equation (11).Thus, with $I = \big(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\big)$,

$$\frac{\partial A}{\partial a_{jj}} = \boldsymbol{e}_j \boldsymbol{e}_j^T,$$

$$\frac{\partial A}{\partial a_{j+k,j}} = \boldsymbol{e}_{j+k} \boldsymbol{e}_j^T \quad \text{with} \;\; k = 1, \ldots, p \;\; \text{for the lower part}$$

and

$$\frac{\partial A}{\partial a_{j,j+s}} = \boldsymbol{e}_j \boldsymbol{e}_{j+s}^T. \quad \text{with} \;\; s = 1, \ldots, q \;\; \text{for the upper part}$$

Consider λ to be a simple nonzero eigenvalue of A and

$$A\boldsymbol{x} = \boldsymbol{x}\lambda, \quad \boldsymbol{y}^* A = \lambda \boldsymbol{y}^*.$$

Then, for $p_{ij} = a_{jj}, a_{j+k,j}, a_{j,j+s}$, with $j = 1, \ldots, n$; $k = 1, \ldots, p$; $s = 1, \ldots, q$, we differentiate $A\boldsymbol{x} = \boldsymbol{x}\lambda$ to get

$$\frac{\partial A}{\partial p_{ij}}\boldsymbol{x} + A\frac{\partial \boldsymbol{x}}{\partial p_{ij}} = \frac{\partial \boldsymbol{x}}{\partial p_{ij}}\lambda + \boldsymbol{x}\frac{\partial \lambda}{\partial p_{ij}}.$$

Multiply by $\boldsymbol{y}^*$ and cancel equal terms to find

$$\frac{\partial \lambda}{\partial p_{ij}}\boldsymbol{y}^*\boldsymbol{x} = \boldsymbol{y}^*\frac{\partial A}{\partial p_{ij}}\boldsymbol{x}, \qquad p_{ij} = a_{j,j}, a_{j+k,j}, a_{j,j+s}, \;\; \text{for} \;\; k = 1, \ldots, p \;\; , \;\; s = 1, \ldots, q \; and \; j = 1, \ldots, n$$

Thus,

$$\frac{\partial \lambda}{\partial a_{jj}} = \frac{\overline{y_j} x_j}{\boldsymbol{y}^*\boldsymbol{x}}, \quad \frac{\partial \lambda}{\partial a_{j+k,j}} = \frac{\overline{y_{j+k}} x_j}{\boldsymbol{y}^*\boldsymbol{x}}, \quad \frac{\partial \lambda}{\partial a_{j,j+s}} = \frac{\overline{y_j} x_{j+s}}{\boldsymbol{y}^*\boldsymbol{x}}$$

and

$$relgrad_A(\lambda) = \frac{1}{\lambda \boldsymbol{y}^*\boldsymbol{x}}\big(a_{11}\overline{y_1}x_1, \ldots, a_{nn}\overline{y_n}x_n, a_{1+k,1}\overline{y_{1+k}}x_1, \ldots, a_{n,n-k}\overline{y_n}x_{n-k}, a_{1,s+1}\overline{y_1}x_{s+1}, \ldots, a_{n-s,n}\overline{y_{n-s}}x_n\big)^T.$$

Finally, we get

$$\| \mathit{relgrad}_A(\lambda)\|_1 = \frac{1}{|\lambda||\boldsymbol{y}^*\boldsymbol{x}|}\left(\sum_{k=0}^{p}\sum_{j=1}^{n-k}|a_{j+k,j}||y_{j+k}||x_j| + \sum_{s=1}^{q}\sum_{j=s+1}^{n}|a_{j-s,j}||y_{j-s}||x_j|\right) = \frac{|\boldsymbol{y}|^T|A||\boldsymbol{x}|}{|\lambda||\boldsymbol{y}^*\boldsymbol{x}|}.$$

We can summarize this result in Theorem 10.2.

Theorem 10.2. *Let $\lambda \neq 0$ be a simple eigenvalue of a (p,q) banded matrix A with left eigenvector $\boldsymbol{y}$ and right eigenvector $\boldsymbol{x}$. Then relcond$(\lambda; A) := \| \mathit{relgrad}_A(\lambda)\|_1$ is equal to*

$$\begin{aligned}\mathit{relcond}(\lambda; A) &= \lim_{\eta\to 0}\sup\left\{\frac{|\delta\lambda|}{\eta|\lambda|} \,:\, (\lambda+\delta\lambda) \text{ is an eigenvalue of } (A+\delta A),\ |\delta A| \leq \eta|A|\right\}\\ &= \frac{|\boldsymbol{y}|^T|A||\boldsymbol{x}|}{|\lambda||\boldsymbol{y}^*\boldsymbol{x}|}.\end{aligned}$$

Note that the expression of the relative condition number $\mathit{relcond}(\lambda; A)$ does not depend on the size of the matrix, neither on p or q in the sense that, the expression for *relcond* is the same in tridiagonals and banded matrices. If $\lambda = 0$, we set $\mathit{relcond}(\lambda; A) = \infty$.

10.1.1 Representation 2 - LU factorization

A (p, q) banded matrix $A \in \mathbb{R}^{n\times n}$ that permits triangular factorization $A = LU$ pass its structure to its factored matrices L and U. This is the main reason for which U has upper bandwidth q, and L has lower bandwidth p (see [14]). We use this inheritance to compute an eigenvalue condition number for matrix A.
Assume that A permits a triangular factorization $A = LU$ and λ is a simple eigenvalue,

$$LU\boldsymbol{x} = \boldsymbol{x}\lambda, \quad \boldsymbol{y}^*LU = \lambda\boldsymbol{y}^*, \qquad \lambda \neq 0.$$

We introduce a notation similar to the one we used in tridiagonal matrices. So, $L = I + \mathring{L}$ with

$$\mathring{L} = \begin{pmatrix} 0 & & & & & \\ l_{21} & 0 & & & & \\ \vdots & & & & & \\ l_{p+1,1} & & \ddots & & & \\ & l_{p+2,2} & & & & \\ & & \ddots & & & \\ & & & l_{n,n-p} & \cdots & 0 \end{pmatrix}$$

and

$$L = \begin{pmatrix} 1 & & & & & & \\ l_{21} & 1 & & & & & \\ \vdots & & & & & & \\ l_{p+1,1} & & & \ddots & & & \\ 0 & l_{p+2,2} & & & & & \\ & & \ddots & & & & \\ & & 0 & l_{n,n-p} & \cdots & 1 \end{pmatrix} \quad \text{and} \quad U = \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1,q+1} & 0 & \\ & u_{22} & & & & \\ & & & & \ddots & 0 \\ & & & \ddots & & u_{n-q,n} \\ & & & & & \vdots \\ & & & & & u_{nn} \end{pmatrix}.$$

The perturbations we consider at level η are given by

$$\begin{aligned} |\delta l_{ij}| &\le \eta |l_{ij}|, \qquad 0 < \eta \ll 1, \\ |\delta u_{ij}| &\le \eta |u_{ij}|, \qquad 0 < \eta \ll 1. \end{aligned}$$

Observe that these perturbations do not change either the 0's or the 1's on the diagonal of L .

Next we derive an explicit expression for the relative condition number for $A = LU$,

$$relcond(\lambda; L, U) := \| \, relgrad_{L,U}(\lambda)\|_1.$$

For $u_{j,j+s}$ we find

$$\frac{\partial \lambda}{\partial u_{j,j+s}} \boldsymbol{y}^* \boldsymbol{x} = \boldsymbol{y}^* L \frac{\partial U}{\partial u_{j,j+s}} \boldsymbol{x} = \boldsymbol{y}^* L \boldsymbol{e}_j \boldsymbol{e}_{j+s}^T \boldsymbol{x} = (\boldsymbol{y}^* L)_j x_{j+s}, \quad s = 0, \ldots, q, \quad j = 1, \ldots, n,$$

and for $l_{j+k,j}$ we find,

$$\frac{\partial \lambda}{\partial l_{j+k,j}} \boldsymbol{y}^* \boldsymbol{x} = \boldsymbol{y}^* \frac{\partial L}{\partial l_{j+k,j}} U \boldsymbol{x} = \boldsymbol{y}^* \boldsymbol{e}_{j+k} \boldsymbol{e}_j^T U \boldsymbol{x} = \overline{y_{j+k}} (U\boldsymbol{x})_j, \quad k = 1, \ldots, p, \quad j = 1, \ldots, n.$$

Then,

$$\begin{aligned} \operatorname{grad}_{L,U}(\lambda) = \big(& \frac{\partial \lambda}{\partial u_{11}}, \ldots, \frac{\partial \lambda}{\partial u_{1,1+q}}, \ldots, \frac{\partial \lambda}{\partial u_{n-1,n-1}}, \frac{\partial \lambda}{\partial u_{n-1,n}}, \frac{\partial \lambda}{\partial u_{n,n}}, \\ & \frac{\partial \lambda}{\partial l_{21}}, \ldots, \frac{\partial \lambda}{\partial l_{p+1,1}}, \ldots, \frac{\partial \lambda}{\partial l_{n-1,n-2}}, \frac{\partial \lambda}{\partial u_{n,n-2}}, \frac{\partial \lambda}{\partial u_{n,n-1}} \big)^T \end{aligned} \tag{73}$$

Observe that U entries in $\operatorname{grad}_{L,U}(\lambda)$ are set in rows and L components by columns.

We finally introduce the parameters $l_{j+k,j}$ and $u_{j,j+s}$ appropriately,

$$\begin{aligned} \lambda(\boldsymbol{y}^* \boldsymbol{x}) \, relgrad_{LU}(\lambda) = & ((\boldsymbol{y}^* L)_1 u_{11} x_1, \ldots, (\boldsymbol{y}^* L)_1 u_{1,1+q} x_{1+q}, \ldots, (\boldsymbol{y}^* L)_{n-1} u_{n-1,n-1} x_{n-1}, (\boldsymbol{y}^* L)_{n-1} u_{n-1,n} x_n, \\ & (\boldsymbol{y}^* L)_n u_{n,n} x_n, (\boldsymbol{y}^* L)_{n-1} u_{n-1,n} x_n, \overline{y_2} l_{21} (U\boldsymbol{x})_1, \ldots, \overline{y_{p+1}} l_{p+1,1} (U\boldsymbol{x})_1, \ldots, \overline{y_{n-1}} l_{n-1,n-2} \\ & (U\boldsymbol{x})_{n-2}, \overline{y_n} l_{n,n-2} (U\boldsymbol{x})_{n-2}, \overline{y_n} l_{n,n-1} (U\boldsymbol{x})_{n-1}))^T. \end{aligned}$$

If we we set the limits of the summatory to be $s = 0, \ldots, \min(q, n-j); \quad k = 0, \ldots, \min(p, n-j)$.
Or this consideration expressed in another form: do not consider the entries that fall out of the size

of the matrix, we finally obtain the following equation,

$$|\lambda||\boldsymbol{y}^*\boldsymbol{x}|\| \, relgrad_{L,U}(\lambda)\|_1 = \sum_{j=1}^{n}\sum_{s=0}^{q}|(\boldsymbol{y}^*L)_j u_{j,j+s}x_{j+s}| + \sum_{j=1}^{n}\sum_{k=1}^{p}|\overline{y_{j+k}}l_{j+k,j}(U\boldsymbol{x})_j| \tag{74}$$
$$= |(\boldsymbol{y}^*L)||U||\boldsymbol{x}| + |\boldsymbol{y}^*||\mathring{L}||(U\boldsymbol{x})|$$

As we did in Theorem 10.2, we can summarize the arguments above in Theorem 10.3.

Theorem 10.3. *Suppose (p,q) banded matrix $A \in \mathbb{R}^{n\times n}$ permits a triangular factorization $A = LU$ with factors as in 18, let $\mathring{L} = L - I$.Let $\lambda \neq 0$ be a simple eigenvalue of A with left eigenvector $\boldsymbol{y}$ and right eigenvector $\boldsymbol{x}$. Then relcond$(\lambda; L, U) := \| \, relgrad_{L,U}(\lambda)\|_1$ is equal to*

$$relcond(\lambda; L, U) = \lim_{\eta\to 0}\sup\left\{\frac{|\delta\lambda|}{\eta|\lambda|} \,:\, (\lambda+\delta\lambda) \text{ is an eigenvalue of } (L+\delta L)(U+\delta U),\right.$$
$$\left. |\delta L| \le \eta|\mathring{L}|, \;\; |\delta U| \le \eta|U|\right\}$$
$$= \frac{|\boldsymbol{y}^*L||U||\boldsymbol{x}| + |\boldsymbol{y}^*||\mathring{L}||U\boldsymbol{x}|}{|\lambda|\,|\boldsymbol{y}^*\boldsymbol{x}|}\,.$$

Observe that the 1's and the 0's perturbations of $\mathcal{L}$ and $\mathcal{U}$ are set to zero.

Note that the factored form has a condition number that does not dependent on the size of the matrix neither p, q. There is a reason by which the resultant expressions for band and tridiagonals are not the same and is that we have used different LU factorizations for band (LU) and tridiagonals ($\mathcal{LU}$). Nevertheless, diag($\mathfrak{u}$) , can be derived from U of the band matrix. It admitt the "fusion" of absolute values, so the expression are rather very similar. Next, we express $relcond(\lambda; \mathcal{L}, \mathcal{U})$ in a form that is more appropiate for computations. In order to extract a factor of $|\lambda|$ on the right in (74), use

$$\boldsymbol{y}^*LU = \lambda\boldsymbol{y}^*, \quad U\boldsymbol{x} = L^{-1}\boldsymbol{x}\lambda, \quad \lambda \neq 0,$$

to find

$$|\boldsymbol{y}^*L| = |\lambda||\boldsymbol{y}^*U^{-1}|, \quad |\boldsymbol{y}^*||\mathring{L}||U\boldsymbol{x}| = |\boldsymbol{y}^*||\mathring{L}||L^{-1}\boldsymbol{x}||\lambda|, \quad |\lambda| \neq 0. \tag{75}$$

Substitute the expressions in (75) into (74) and cancel $|\lambda|$ ($\neq 0$) to find

$$|\boldsymbol{y}^*\boldsymbol{x}|\| \, relgrad_{\mathcal{LU}}(\lambda)\|_1 = |\boldsymbol{y}^*U^{-1}||U||\boldsymbol{x}| + |\boldsymbol{y}^*||\mathring{L}||L^{-1}\boldsymbol{x}|.$$

For the cost of solving the following

$$|\boldsymbol{y}^*U^{-1}||U| = \boldsymbol{w}^*$$
$$\boldsymbol{x} = L\boldsymbol{g} \text{ for } \boldsymbol{g}$$
$$|\boldsymbol{y}^*||\mathring{L}||\boldsymbol{g}| = \boldsymbol{v}$$

we obtain the following expression of the relative condition number for $A = LU$

$$\boxed{relcond(\lambda; L, U) := \frac{|\boldsymbol{w}^T||\boldsymbol{x}| + \boldsymbol{v}}{|\boldsymbol{y}^*\boldsymbol{x}|}.} \tag{76}$$

Therefore, it seems appropriate to set $relcond(0; \mathcal{L}, \mathcal{U}) = 0$.

11 Conclusions and future work

In this undergraduate thesis, we have deduced for first time formulas for the eigenvalue condition numbers of $n \times n$ tridiagonal matrices with respect to perturbations of four different parametrizations or representations of tridiagonal matrices that are important in numerical computations. These representations are: the entries of the matrix; the bidiagonal $\mathcal{L}$ and $\mathcal{U}$ factors of the J-form of tridiagonal matrices; the signed symmetric bidiagonal-diagonal factorization, $\Delta L D L^T$, of the balanced form of tridiagonal matrices; and the double signed bidiagonal factorization $\Delta \bar{L} \Omega \bar{L}^T$ of the balanced form of tridiagonal matrices.

In all the cases we have considered, the perturbations of the parameters preserve the tridiagonal structure of the unperturbed matrix.

The new contribution with respect to previous references available in the scientific literature is that we have measured the sizes of the perturbations of the four different sets of parameters we have analyzed via the 2-norm of the vector of relative variations of the parameters.

It has been shown in very recent ongoing research works [9, 31] that the use of the 2-norm may have relevant advantages for estimating the errors committed by fast modern algorithms for computing the eigenvalues of unsymmetric tridiagonal matrices [1, 11].

Moreover, we have presented numerical procedures for computing the new eigenvalue condition numbers in $O(n)$ flops by solving some very simple linear systems. This makes possible to apply our results in the future for estimating the errors committed by the algorithms presented in [1, 11].

We have also deduced for first time formulas for the eigenvalue condition numbers of $n \times n$ general low-banded matrices with respect to perturbations of the triangular L and U factors of these matrices and we have presented numerical methods for computing these new formulas in $O(n)$ flops.

The most relevant future research work directly related to the original results obtained in this undergraduate thesis would be to combine the new eigenvalue condition numbers we have developed with some efficient algorithms for computing structured eigenvalue backward errors of tridiagonal matrices, which are currently under development [9] [31]. In this way, we would provide a reliable procedure to estimate *a posteriori*, but very efficiently, the errors committed by fast modern algorithms for computing the eigenvalues of unsymmetric tridiagonal matrices [1, 11]. Since these algorithms are not guaranteed to be backward stable, these error estimations would make possible to include fast algorithms for the unsymmetric tridiagonal eigenvalue problem in high quality software professional libraries of Scientific Computing.

References

[1] D. A. Bini, L. Gemignani and F. Tisseur. *The Ehrlich-Aberth method for the non-symmetric tridiagonal eigenvalue problem*. SIAM J. Matrix Anal. Appl., 27(1):153-175 (2005).

[2] K. Braman, R. Byers, and R. Mathias. *The multishift QR algorithm. I. Mantaining well-focused shifts and level 3 performance*. SIAM J. Matrix Anal. Appl., 23:929-947 (2002).

[3] K. Braman, R. Byers, and R. Mathias. *The multishift QR algorithm. II. Aggressive early deflation*. SIAM J. Matrix Anal. Appl., 23:948-973 (2002).

[4] R. C. Craddock et al. *Imaging human connectomes at the macroscale*. Nat. Methods 10:524-536 (2013).

[5] P. A. Clement. *A class of triple-diagonal matrices for test purposes*. SIAM Review, 1:50-52 (1959).

[6] I. S. Dhillon. *A new $O(n^2)$ Algorithm for the Symmetric Tridiagonal Eigenvalue/Eigenvector Problem*. Ph.D. Thesis, Computer Science Division, University of California, Berkeley, May 1997.

[7] I. S. Dhillon and B. N. Parlett. *Multiple representation to compute orthogonal eigenvectors of symmetric tridiagonal matrices*. Linear Algebra Appl., 387:1-28 (2004).

[8] J. Dongarra and F. Sullivan. *The top 10 algorithms*. Comput. Sc. Eng., 2:22-23 (2000).

[9] F. M. Dopico, B. N. Parlett, and C. Ferreira. *The inverse complex eigenvector problem for real tridiagonal matrices*. Invited talk in Special Session "Linear Algebra: Algorithms and Applications" in "First Joint International Meeting RSME-SCM-SEMA-SIMAI-UMI", Bilbao, Spain, 30 June–4 July 2014.

[10] C. Ferreira and B. N. Parlett. *Convergence of LR algorithm for a one-point spectrum tridiagonal matrix*. Numer. Math., 113(3):417-431 (2009).

[11] C. Ferreira and B. N. Parlett. *Real dqds for the nonsymmetric tridiagonal eigenvalue problem*. Submitted.

[12] C. Ferreira, B. N. Parlett, and F. M. Dopico. *Sensitivity of eigenvalues of an unsymmetric tridiagonal matrix*. Numer. Math., 122:527-555 (2012).

[13] R. Granat, B. Kågström, D. Kressner, and M. Shao. *Parallel library software for the multishift QR algorithm with aggressive early deflation*. (To appear in ACM Trans. Math. Software).

[14] G. H. Golub and C. F. Van Loan. *Matrix Computations*. Johns Hopkins University Press, Baltimore, Fourth Edition, 2013.

[15] D. J. Higham and N. J. Higham. *Structured backward error and condition of generalized eigenvalue problems*. SIAM J. Matrix Anal. Appl., 20:493-512 (1998).

[16] N. J. Higham. *Accuracy and Stability of Numerical Algorithms*. Society for Industrial and Applied Mathematics, Philadelphia, Second Edition, 2002.

[17] N. J. Higham, M. Konstantinov, V. Mehrmann, and P. Petkov. *The sensitivity of computational control problems*. IEEE Control Syst. Mag., 24(1):28-43 (2004).

[18] Alston. S. Householder. *The Theory of Matrices in Numerical Analysis* Blaisdell, New York, 1964. xi+257 pp. Reprinted by Dover, New York. ISBN 0-486-61781-5.

[19] I. C. F. Ipsen. *Relative perturbation results for matrix eigenvalues and singular values*. Acta Numer. vol. 7, 1998, pp. 151-201.

[20] M. Karow, D. Kressner, and F. Tisseur. *Structured eigenvalue condition numbers*. SIAM J. Matrix Anal. Appl., 28(4):1052-1068 (2006).

[21] A. J. Laub. *Matrix Analysis for Scientists and Engineers*. Society for Industrial and Applied Mathematics, Philadelphia, 2005.

[22] H. W. J. Lenferink and M. N. Spijker. *On the use of stability regions in the numerical analysis of initial value problems*. Math. Comp., 57(195): 221-237 (1991).

[23] R.-C. Li. *Relative perturbation theory III. More bounds on eigenvalue variations*. Linear Algebra Appl., 266:337-345 (1997).

[24] Z. S. Liu. *On the extended HR algorithm*. Technical Report PAM-564, Center for Pure and Applied Mathematics, University of California, Berkeley, CA, USA, 1992.

[25] L. Merino and E. Santos. *Algebra Lineal con Metodos Elementales*. Thomson, 2006.

[26] C. B. Moler. *Numerical Computing with MATLAB*. Society for Industrial and Applied Mathematics, Philadelphia, 2004.

[27] S. Noschese and L. Pasquini. *Eigenvalue condition numbers: zero-structured versus traditional*. J. Comp. Appl. Math., 185:174-189 (2006).

[28] B. N. Parlett. *Spectral sensitivity of products of bidiagonals*. Linear Algebra Appl., 275-276:417-431 (1998).

[29] B. N. Parlett. *The new qd algorithms*. Acta Numerica, vol. 4, Cambridge University Press, Cambridge, 1995, pp. 459-491.

[30] B. N. Parlett and C. Reinsch. *Balancing a matrix for calculation of eigenvalues and eigenvectors*. Numer. Math., 13:292-304 (1969).

[31] B. N. Parlett, F. M. Dopico, and C. Ferreira. *Structured backward relative error bounds for eigenvalues of tridiagonal matrices*. Invited talk in "Minisymposium on Eigenvalue Computations: Theory and Practice" in "2013 SIAM Annual Meeting", San Diego, California, USA, 8-12 July 2013.

[32] L. Pasquini. *Accurate computation of the zeros of the generalized Bessel polynomials*. Numer. Math., 86:507-538 (2000).

[33] J. Rice. *A theory of condition*. SIAM J. Numer. Anal., 3(2):287-310 (1966).

[34] Y. Saad. *Numerical Methods for Large Eigenvalue Problems*. Society for Industrial and Applied Mathematics, Philadelphia, revised edition, 2011.

[35] J. Slemons. *Toward the Solution of the Eigenproblem: Nonsymmetric Tridiagonal Matrices*. Ph.D Thesis, University of Washington, Seattle, 2008.

[36] G. W. Stewart and J.-G. Sun. *Matrix Perturbation Theory*. Academic Press, New York, 1990.

[37] F. Tisseur and K. Meerbergen. *The quadratic eigenvalue problem*. SIAM Review, 43(2): 235-286 (2001).

[38] L. Trefethen and D. Bau. *Numerical Linear Algebra*. Society for Industrial and Applied Mathematics, Philadelphia, 1997.

[39] J. H. Wilkinson. *The Algebraic Eigenvalue Problem*. Clarendon Press, Oxford, 1965.

[40] "Diffusion MRI". *Wikipedia: The free encyclopedia Wikimedia Foundation,Inc.*, 4 Sep. 2013. Web. 5 Sep. 2013. $< http://en.wikipedia.org/wiki/Diffusion\_MRI >$