# Chapter One

# Introduction

The aim of this book is to describe and explain the beautiful mathematical relationships between matrices, moments, orthogonal polynomials, quadrature rules and the Lanczos and conjugate gradient algorithms. Even though we recall the mathematical basis of the algorithms, this book is computationally oriented. The main goal is to obtain efficient numerical methods to estimate or in some cases to bound quantities like $I[f] = u^T f(A)v$ where $u$ and $v$ are given vectors, $A$ is a symmetric nonsingular matrix and $f$ is a smooth function. The main idea developed in this book is to write $I[f]$ as a Riemann–Stieltjes integral and then to apply Gauss quadrature rules to compute estimates or bounds of the integral. The nodes and weights of these quadrature rules are given by the eigenvalues and eigenvectors of tridiagonal matrices whose nonzero coefficients describe the three-term recurrences satisfied by the orthogonal polynomials associated with the measure of the Riemann–Stieltjes integral. Beautifully, these orthogonal polynomials can be generated by the Lanczos algorithm when $u = v$ or by its variants otherwise. All these topics have a long and rich history starting in the nineteenth century. Our aim is to bring together results and algorithms from different areas. Results about orthogonal polynomials and quadrature rules may not be so well known in the matrix computation community, and conversely the applications in matrix computations that can be done with orthogonal polynomials and quadrature rules may be not too familiar to the community of researchers working on these topics. We will see that it can be very fruitful to mix techniques coming from different areas.

There are many instances in which one would like to compute bilinear forms like $u^T f(A)v$. A first obvious application is the computation of some elements of the matrix $f(A)$ when it is not desired or feasible to compute all of $f(A)$. Computation of quadratic forms $r^T A^{-i} r$ for $i = 1, 2$ is interesting to obtain estimates of error norms when one has an approximate solution $\tilde{x}$ of a linear system $Ax = b$ and $r$ is the residual vector $b - A\tilde{x}$. Bilinear or quadratic forms also arise naturally for the computation of parameters in some numerical methods for solving least squares or total least squares problems and also in Tikhonov regularization for solving ill-posed problems.

The first part of the book provides the necessary mathematical background and explains the theory while the second part describes applications of these results, gives implementation details and studies improvements of some of the algorithms reviewed in the first part. Let us briefly describe the contents of the next chapters.

The second chapter is devoted to orthogonal polynomials, whose history started in the nineteenth century from the study of continued fractions. There are many excellent books on this topic, so we just recall the properties that will be useful in the other chapters. The important point for our purposes is that orthogonal polynomials satisfy three-term recurrences. We are also interested in some properties of the zeros of these polynomials. We give some examples of classical orthogonal polynomials like the Legendre, Chebyshev and Laguerre polynomials. Some of them will be used later in several algorithms and in numerical experiments. We also introduce a less classical topic, matrix orthogonal polynomials, that is, polynomials whose coefficients are square matrices. These polynomials satisfy block three-term recurrences and lead to consideration of block tridiagonal matrices. They will be useful for computing estimates of off-diagonal elements of functions of matrices.

Since tridiagonal matrices will play a prominent role in the algorithms described in this book, chapter 3 recalls properties of these matrices. We consider Cholesky-like factorizations of symmetric tridiagonal matrices and properties of the eigenvalues and eigenvectors. We will see that some elements of the inverse of tridiagonal matrices (particularly the $(1, 1)$ element) come into play for estimating bilinear forms involving the inverse of $A$. Hence, we give expressions of elements of the inverse obtained from Cholesky factorizations and algorithms to cheaply compute elements of the inverse. Finally, we describe the QD algorithm which was introduced by H. Rutishauser to compute eigenvalues of tridiagonal matrices and some of its variants. This algorithm will be used to solve inverse problems, namely reconstruction of symmetric tridiagonal matrices from their spectral properties.

Chapter 4 briefly describes the well-known Lanczos and conjugate gradient (CG) algorithms. The Lanczos algorithm will be used to generate the recurrence coefficients of orthogonal polynomials related to our problem. The conjugate gradient algorithm is closely linked to Gauss quadrature and we will see that quadrature rules can be used to obtain bounds or estimates of norms of the error during CG iterations when solving symmetric positive definite linear systems. We also describe the nonsymmetric Lanczos and the block Lanczos algorithms which will be useful to compute estimates of bilinear forms $u^T f(A)v$ when $u \neq v$. Another topic of interest in this chapter is the Golub–Kahan bidiagonalization algorithms that are useful when solving least squares problems.

Chapter 5 deals with the computation of the tridiagonal matrices containing the coefficients of the three-term recurrences satisfied by orthogonal polynomials. These matrices are called Jacobi matrices. There are many circumstances in which we have to compute the Jacobi matrices either from knowledge of the measure of a Riemann–Stieltjes integral or from the moments related to the measure. It is also important to be able to solve the inverse problem of reconstructing the Jacobi matrices from the nodes and weights of a quadrature formula which defines a discrete measure. We first describe the Stieltjes procedure, which dates back to the nineteenth century. It computes the coefficients from the measure which, in most cases, has to be approximated by a discrete measure. This algorithm can be considered as a predecessor of the Lanczos algorithm although it was not constructed to compute eigenvalues. Unfortunately there are cases for which the Stieltjes algorithm gives poor results due to a sensitivity to roundoff errors. Then we show how the nonzero entries of the Jacobi matrices are related to determinants of Hankel matrices constructed from the moments. These formulas are of little computational interest even though they have been used in some algorithms. More interesting is the modified

Chebyshev algorithm, which uses so-called modified moments to compute the Jacobi matrix. These modified moments are obtained from some known auxiliary orthogonal polynomials. The next section consider several algorithms for solving the problem of constructing the Jacobi matrix from the nodes and weights of a discrete measure. They are the eigenvalues and squares of the first elements of the eigenvectors. Hence, this is in fact an inverse eigenvalue problem of reconstructing a tridiagonal matrix from spectral information. Finally, we describe modification algorithms which compute the Jacobi matrices for measures that are given by a measure for which we know the coefficients of the three-term recurrence multiplied or divided by a polynomial.

The subject of chapter 6 is Gauss quadrature rules to obtain approximations or bounds for Riemann–Stieltjes integrals. The nodes and weights of these rules are related to the orthogonal polynomials associated with the measure and they can be computed using the eigenvalues and eigenvectors of the Jacobi matrix describing the three-term recurrence. With $N$ nodes, the Gauss rule is exact for polynomials of order $2N - 1$. The Jacobi matrix has to be modified if one wants to fix a node at one end or at both ends of the integration interval. This gives respectively the Gauss–Radau and Gauss–Lobatto quadrature rules. We also consider the anti-Gauss quadrature rule devised by D. P. Laurie to obtain a rule whose error is the opposite of the error of the Gauss rule. This is useful to estimate the error of the Gauss quadrature rule. The Gauss–Kronrod quadrature rule uses $2N + 1$ nodes of which $N$ are the Gauss rule nodes to obtain a rule that is exact for polynomials of degree $3N + 1$. It can also be used to estimate errors in Gauss rules. Then we turn to topics that may be less familiar to the reader. The first one is the nonsymmetric Gauss quadrature rule which uses two sets of orthogonal polynomials. The second one is block Gauss quadrature rules to handle the case where the measure is a symmetric matrix. This involves the matrix orthogonal polynomials that were studied in chapter 2.

Chapter 7 is, in a sense, a summary of the previous chapters. It shows how the theoretical results and the techniques presented before allow one to obtain bounds and estimates of bilinear forms $u^T f(A)v$ when $A$ is a symmetric matrix and $f$ a smooth function. First, we consider the case of a quadratic form with $u = v$. To solve this problem we use the Lanczos algorithm which provides a Jacobi matrix. Using the eigenvalues and eigenvectors of this matrix (eventually suitably modified) we can compute the nodes and weights of Gauss quadrature rules. This gives estimates or bounds (if the signs of the derivatives of $f$ are constant over the interval of integration) of the quadratic form. When $u \neq v$ we use either the nonsymmetric Lanczos algorithm or the block Lanczos algorithm. With the former we can in some cases obtain bounds for the bilinear form whereas with the latter we obtain only estimates. However, the block Lanczos algorithm has the advantage of delivering estimates of several elements of $f(A)$ instead of just one for the nonsymmetric Lanczos algorithm.

Chapter 8 briefly describes extensions of the techniques summarized in chapter 7 to the case of a nonsymmetric matrix $A$. The biconjugate gradient and the Arnoldi algorithms have been used to compute estimates of $u^T f(A)v$ or $u^H f(A)v$ in the complex case. Some justifications of this can be obtained through the use of Gauss quadrature in the complex plane [293] or, more interestingly, the Vorobyev moment problem [316].

The first part of the book is ended by chapter 9 which is devoted to solving secular equations. We give some examples of problems for which it is useful to solve such equations. One example is computing the eigenvalues of a matrix $A$ perturbed by a rank-one matrix $cc^T$ where $c$ is a given vector. To compute the eigenvalues $\mu$ we have to solve the equation $1 + c^T(A - \mu I)^{-1}c = 0$. Note that this equation involves a quadratic form. Using the spectral decomposition of $A$, this problem can be reduced to solving a secular equation. We review different numerical techniques to solve such equations. Most of them are based on use of rational interpolants.

The second part of the book describes applications and gives numerical examples of the algorithms and techniques developed in the first nine chapters.

Even though this is not the main topic of the book, chapter 10 gives examples of computation of Gauss quadrature rules. It amounts to computing eigenvalues and the first components of the eigenvectors. We compare the Golub and Welsch algorithm with other implementations of the QR or the QL algorithms. We also show some examples of computation of integrals and describe experiments with modification algorithms where one computes the Jacobi matrix associated with a known measure multiplied or divided by a polynomial.

Chapter 11 is concerned with the computation of bounds for elements of $f(A)$. The functions $f$ we are interested in as examples are $A^{-1}$, $\exp(A)$ and $\sqrt{A}$. We start by giving analytical lower and upper bounds for elements of the inverse. This is obtained by doing "by hand" one or two iterations of the Lanczos algorithm. These results are then extended to any function $f$. We also show how to compute estimates of the trace of the inverse and of the determinant of $A$, a problem which does not exactly fit in the same framework. These algorithms are important for some applications in physics. Several numerical examples are provided to show the efficiency of our techniques for computing bounds and to analyze their accuracy.

Chapter 12 studies the close relationships of the conjugate gradient algorithm with Gauss quadrature. In fact, the square of the $A$-norm of the error at iteration $k$ is the remainder of a $k$-point Gauss quadrature rule for computing $(r^0)^T A^{-1} r^0$ where $r^0$ is the initial residual. Bounds of the $A$-norm of the error can be computed during CG iterations by exploiting this relationship. If one is interested in the $l_2$ norm of the error, it can also be estimated during the CG iterations. This leads to the definition of reliable stopping criteria for the CG algorithm. These estimates have been used when solving finite element problems. One can define a stopping criterion such that the norm of the error with the solution of the continuous problem is at the level one can expect for a given mesh size. Numerous examples of computation of bounds of error norms are provided.

In chapter 13 we consider the least squares fit of some given data by polynomials. The solution to this problem can be expressed using the orthogonal polynomials related to the discrete inner product defined by the data. We are particularly interested in the updating and downdating operations where one adds or deletes data from the sample. This amounts to computing new Jacobi matrices from known ones. We review algorithms using orthogonal transformations to solve these problems, which are also linked to inverse eigenvalue problems. We also consider the problem of computing the backward error of a least squares solution. Use of the exact expression of the backward error is difficult because it amounts to computing the smallest eigenvalue of a rank-one modification of a singular matrix. However one can compute an approximation of the backward error with Gauss quadrature.

Given a matrix $A$ and a right-hand side $c$, the method of Total Least Squares (TLS) looks for the solution of $(A + E)x = c + r$ where $E$ and $r$ are the smallest perturbations in the Frobenius norm such that $c + r$ is in the range of $A + E$. To compute the solution we need the smallest singular value of the matrix $( A \quad c )$. It is given as the solution of a secular equation. In chapter 14, approximations of this solution are obtained by using the Golub–Kahan bidiagonalization algorithm and Gauss quadrature.

Finally, chapter 15 considers the determination of the Tikhonov regularization parameter for discrete ill-posed problems. There are many criteria which have been devised to define good parameters. We mainly study generalized cross-validation (GCV) and the L-curve criteria. The computations of the “optimal” parameters for these methods involve the computation of quadratic forms which can be approximated using Gauss quadrature rules. We describe improvements of algorithms which have been proposed in the literature and we provide numerical experiments to compare the different criteria and the algorithms implementing them.

This book should be useful to researchers in numerical linear algebra and more generally to people interested in matrix computations. It can be of interest too to scientists and engineers solving problems in which computation of bilinear forms arises naturally.