Matrix perturbations: bounding and computing eigenvalues

Reis da Silva, R.J.

Citation for published version (APA):
This chapter summarizes the contents of this thesis. The initial sections are intended for the reader not familiar with the concepts of eigenvalues and matrices and the connections between the two. The language used is accessible and accuracy and detail were sacrificed in favor of the understanding of the larger issues. The knowledgeable reader may wish to skip to the last section where a more specialized overview of the main results is given.
5.1 The mathematical model

The natural frequency of an object is the frequency at which such object vibrates when hit, touched or disturbed. While some objects might produce sound when vibrating such as the strings in a violin or crystal glasses, others might not. Nevertheless, every object has a natural frequency. Even larger objects such as cars and airplanes, buildings and bridges have natural frequencies. Buildings oscillate when a heavy truck passes down the street or a plane takes off close by. If a structure is not well designed it may vibrate excessively and may even collapse. A famous case was the Tacoma Narrows Bridge in the state of Washington in the United States. The bridge opened on July of 1940 and, not long after, started oscillate excessively. Five months later collapsed. The study of vibrations is, therefore, crucial in civil engineering.

Let us work with a very simple and typical example and see how the study of vibrations relates with the contents of this thesis. Our model is an apartment building with two stories represented in Figure 5.1.

Figure 5.1: Two stories building. The elasticity of the walls is denoted by the $a_j$'s, their displacement by the $\tilde{x}_j$'s and the mass of the ceilings are denoted by the $m_j$'s for $j \in \{1, 2, 3\}$.

To make the model as simple as possible we neglect the mass of the walls and model only their elasticities $a_1$, $a_2$ and $a_3$ which are known quantities. Only
the ceilings have mass which we denote by $m_1, m_2$ and $m_3$ and assume to be
the same. To simplify even more we assume that $m_1 = m_2 = m_3 = 1$. The
horizontal movement of the ceilings (horizontal displacement) is represented by the
functions $\tilde{x}_1, \tilde{x}_2$ and $\tilde{x}_3$. We assume further that the foundations of the building
are strong enough and thus the base of the building is fixed to the ground. Using
the appropriate physical laws governing forces, masses and displacement one can
describe the interactions occurring in the building by a system of equations:

\[
\begin{align*}
(a_1 + a_2)x_1 - a_2x_2 &= \lambda m_1 x_1 \\
-a_2x_1 + (a_2 + a_3)x_2 - a_3x_3 &= \lambda m_2 x_2 \\
-a_3x_2 + a_3x_3 &= \lambda m_3 x_3.
\end{align*}
\] (5.1)

The variables $a_j$ and $m_j$ in the system (5.1) are known but $\lambda$ and the $x_j$s are
unknown. The system turns out to be even slightly more complicated than what
it seems as this is a system of three equations but nine unknowns. In fact, there are
three possible values for $\lambda$ and, for each of them, there is an associated solution
made up of the elements $x_1, x_2$ and $x_3$. The actual solutions, however, depend
always on the specific values of $a_1$, $a_2$ and $a_3$.

The number of equations in the system depends on the underlying problem.
Adding a new floor to the building would result in an additional equation and
an additional variable $a_4$ but in four extra unknowns (one extra $\lambda$ and an extra
element $x_4$ and so to each $\lambda$ now corresponds a combination of $x_1, \ldots, x_4$).

The system we have just described is simple and small. More complicated
structures, and different type of problems may take the number of equations to
reach the billions! A famous one is the system arising from Google’s method
for determining the order of importance of each webpage. It is estimated that
such system of equations contains, currently, more that eleven billion ($11 \times 10^9$
or 11 followed by 9 zeros) equations [52]. (Note that the estimated size of the
same matrix in 2002 was only about three billion [55]). Although such number of
equations is still exceptional, standard problems have now about $10^6$ (1 million)
equations.
5.2 From systems to matrices

With very small systems, say two or three equations such as the one showed earlier, it is still possible to keep control of all elements. One could, in fact, find a solution using just pen and paper. If, however, the number of equations or the number of unknowns increase, the task becomes more cumbersome. The need arose, thus, to develop a way to represent systems of equations that did not require one to write every element of the system every time. This was solved by grouping all common elements together and defining some rules for their manipulation. We can, thus, write the previous system in an equivalent manner as

\[
\begin{bmatrix}
  a_1 + a_2 & -a_2 & 0 \\
  -a_2 & a_2 + a_3 & -a_3 \\
  0 & -a_3 & a_3
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
x_3
\end{bmatrix}
= \lambda
\begin{bmatrix}
m_1 & 0 & 0 \\
0 & m_2 & 0 \\
0 & 0 & m_3
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
x_3
\end{bmatrix}.
\]

(5.2)

The two large square elements are called matrices and are usually represented using capital letters of the alphabet, say \( A \) and \( M \), respectively. The rectangular elements are named vectors and are usually denoted by lower case letters, for instance \( \vec{x} \). Using the new symbols, system (5.1) in the form of equation (5.2) can, indeed, be expressed simply as

\[ A\vec{x} = \lambda M\vec{x}. \]

The size of a matrix, that is, the number of rows or columns is equal to the number of equations in the original system. Note that this can reach a few millions! Using the rules defined for the sum and product of matrices and vectors one can show that the solution of problem (5.2), assuming that \( m_1 = m_2 = m_3 = 1 \), is equivalent to the solution of the equation

\[ A\vec{x} = \lambda \vec{x}. \]

(5.3)

The variable \( \lambda \) is called an eigenvalue (from the German “eigen” meaning proper, thus “proper value”) of \( A \) and \( \vec{x} \) is the eigenvector of \( A \) associated with the eigenvalue \( \lambda \). The number of eigenvalues of a matrix is equal to the number rows and columns of that matrix and, therefore, equal to the number of equations of the
original system. It may happen, however, that some of the eigenvalues are equal. The value of the natural frequency we mentioned in the beginning is obtained from the smallest of those eigenvalues of $A$.

**Perturbation theory**  Perturbation theory for eigenvalues is another component of the study of eigenvalues. In practice, it is not possible to compute the exact eigenvalues of a matrix but just a close approximation. It is, therefore, of interest to understand what happens to $\lambda$ and $\vec{x}$ in the presence of errors for each $a_k$. In other words, what are the values of $\lambda$ and $\vec{x}$ if instead of $a_k$ the actual value is $a_k + e_k$ for some error $e_k$? We may pose the same question in terms of determining the eigenvalues and eigenvectors of a matrix closer to $A$, say $A + E$, for some error matrix $E$, i.e., to solve

$$(A + E)\tilde{x} = \tilde{\lambda}\tilde{x}.$$ 

If, however, the eigenvalues and eigenvectors of $A$ are already known, the investment in computational effort to determine the eigenvalues of $A + E$ may be unnecessary. Perturbation theory attempts to determine the approximate location of the eigenvalues of $A + E$ without explicitly computing them.

**Iterative methods for eigenvalue problems**  Because nowadays the typical matrix is of the order of the millions, computers are needed to determine the eigenvalues and eigenvectors. Iterative methods for eigenvalue problems are computational methods designed to determine a few of the eigenvalues and eigenvectors of a given matrix. These can be a few of the smallest, of the largest or just a few in between. They are called *iterative* because they repeat a certain set of steps over and over. Each step gets us closer to the solution and the method stops once (a very close approximation to) the final solution is found.
5.3 This thesis

Despite the somewhat negative connotation of the word, not every perturbation is a bad perturbation. In fact, while disturbing the matrix entries, many perturbations still preserve useful properties such as the orthonormality of the basis of eigenvectors or the Hermicity of the original matrix.

In the first part of this thesis some of these property preserving perturbations are analyzed. The ultimate goal is to understand how the eigenvalues of the original matrix are influenced by such perturbations. For Hermitian matrices, this resulted in improved Weyl-type bounds for the eigenvalues of Hermitian rank-$k$ perturbations. Their accuracy is dependent on the available eigenvalues and eigenvectors of the original matrix. The same approach is then used to determine bounds for the singular values of perturbations and augmentations of $n \times m$ matrices.

With respect to normal matrices, normality preserving normal perturbations of normal matrices were also considered. For $2 \times 2$ and for rank-one matrices, the analysis is now complete, while for higher rank, all essentially Hermitian normality perturbations are described. The normality preserving augmentation of normal matrices studied by Ikramov and Elsner is revisited and their results are complemented with an analysis on the consequences to the eigenvalues. All augmentations that result in normal matrices with eigenvalues on a quadratic curve in the complex plane are also constructed.

In the second part the Subspace Projected Approximate Matrix (SPAM) method, an iterative method for the Hermitian eigenvalue problem, is investigated from a Numerical Linear Algebra point of view. For certain special choices of the approximation matrix, SPAM turns out to be mathematically equivalent to the Lanczos method. More sophisticated approximations turn SPAM into a boosted version of the Lanczos method. It can also, however, be interpreted as an attempt to enhance a certain instance of the preconditioned Jacobi-Davidson method.