Iterative Methods for Eigenvalues of Symmetric Matrices as Fixed Point Theorems

A Midterm Report by Jeffrey Springer: In Conjunction with Professor Marty Greenlee

Department of Mathematics: University of Arizona

Introduction and Goals

Thus far in my research I have begun to study the theoretical groundwork for what will become our iterative procedures to determine eigenvalues of Symmetric Matrices. Throughout this midterm report I will describe in detail the different areas of mathematics I have studied throughout the first half of this project, which will become the Mathematical framework used in our research. Currently there is literature discussing many methods for approximating Eigenvalues for a given matrix **A**. The motivation for our research is to approach this from a new viewpoint using fixed point theorems, and hopefully in doing so we will gain some valuable insight into the field of numerical linear algebra. Throughout the last several weeks I have read much literature which discusses concepts such as contraction mappings, metric spaces, the Power method, and the Inverse power method. At first these concepts were a bit difficult to understand as they are from more advanced mathematics courses that I have yet to take in my academic career. As I progressed through the literature I consulted Professor Greenlee many times to better understand the ideas that are fundamental to the project we are undertaking.

In this report I hope to give a description of what I have learned about each of the above topics, and how those topics are applicable to our particular project. After describing the methods I have studied, I will briefly cover how we plan to rewrite the power method so that contraction mapping techniques are useful, and discuss how this can further be applied to the inverse power method. After describing these ideas and methods I will close this midterm summary with comments on where we hope to take our project in the next half of the semester, as well as the possibilities of where this project could go in the future.

Mathematical Methods

Metric Spaces and Cauchy Sequences

The first concept which I studied is that of a metric space, which was somewhat familiar to me, but it was necessary to learn more in order to use the concept for the purpose of contraction mapping theorems. One of the main concepts I studied in detail is the following definition:

Definition 1: *A sequence $\{v_k\}$ is a Cauchy or Fundamental Sequence if for each $\varepsilon > 0$ there exists N such that:*
$d(v_m, v_p) < \varepsilon$ *whenever* $m, p > N$.
This definition is necessary to our use of contraction maps in metric spaces. From this definition followed a theorem which discusses how convergence of a sequence is related to whether or not a sequence is Cauchy.

Theorem 1: *If a sequence $\{v_k\}$ converges, it is fundamental.*

One of the first parts of my research was to read about this theorem and I went through the proof of this theorem which is trivial and follows rather simply from the triangle inequality for metric spaces. The main application of metric spaces to our project is in the fact that we use the concept of a metric space in our discussion of contraction map techniques that will be used in fixed point formulations of the Power Method and Inverse Power Method. The next topic I considered is the idea of a contraction mapping and how this can be applied to iterative methods for eigenvalues of $n \times n$ matrices.

Contraction Mappings: Definition, Theorems, and Applications to Matrices

Now I will give an introduction to transformation of the form:

$$v = Av, \qquad (1)$$

Here we define A to be a transformation of a metric space into itself. We see that the solutions to equation (1) are fixed points which do not vary under the transformation A. This is the idea that we are going to use in order to rewrite the power method as a fixed point theorem, where contraction mapping techniques are clearly applicable. The most popular method for solving these types of equations is the method of successive approximations. In our research we will use successive approximations by defining an initial approximation v_0 and use a successive approximation method that follows the structure $v_1 = Av_0$, $v_2 = Av_1$,....$v_n = Av_{n-1}$ ... Typically for large values of n we will get a good approximation to the fixed point v of A. This rather general method will be applied in rewriting the power method to find a dominant eigenvalue λ_i by successive approximations of a contraction map. Now we need an important definition to discuss how a transformation from a metric space into itself becomes a contraction mapping:

Definition 2: *If there exists a positive constant ρ, independent of u and v in a metric space **X** such that:*

$$d(\mathbf{A}u, \mathbf{A}v) \leq \rho d(u, v) \qquad (2)$$

*Then the transformation **A** is Lipschitz continuous, and if this holds for some fixed $\rho < 1$, **A** is a contraction.*

In order to come to our next theorem we need yet another definition to introduce the concept of a complete metric space.

Definition 3: *A metric space X is said to be complete if every Cauchy sequence of points from X converges to a limit which exists in X.*

This definition is fundamentally important because given a metric space X we shall usually be interested in whether or not Cauchy sequences converge to a limit which is in the space. This is the of the utmost significance in our next theorem.

Theorem 2: *Let **A** be a contraction on a complete metric space X. Then the fixed point equation $v = \mathbf{A}v$ has a unique solution, which can be found by successive approximations from any given starting value.*

Since the proof of this is a bit involved and will largely be used in our development of the power method into a fixed point theorem I include the proof below using general vectors v and u:

Proof of Theorem 2: *To prove uniqueness of the fixed point we suppose that $u = \mathbf{A}u$ and $v = \mathbf{A}v$, here we see that $d(u, v) = d(\mathbf{A}u, \mathbf{A}v)$. Since **A** is a contraction there exists a $\rho < 1$ such that $d(u,v) \leq \rho d(u,v)$. This relationship implies that $d(u, v) = 0$ and therefore by the properties of a metric space $u = v$.*

Now to establish the existence of a fixed point we make an initial approximation v_0, and define the approximations by our previous method $v_n = Av_{n-1}$. First we have to establish that $\{u_k\}$ is a Cauchy sequence. We have that

$d(u_n, u_{n+1}) = d(Au_{n-1}, Au_n) \leq \rho d(u_{n-1}, u_n) \leq \ldots \leq \rho^n d(u_0, u_1)$, *Now if $m > n$ we have that:*

$$
\begin{aligned}
d(u_n, u_m) &\leq d(u_n, u_{n+1}) + \ldots + d(u_{m-1}, u_m) \\
&\leq \rho^n d(u_0, u_1) + \rho^{n+1} d(u_0, u_1) + \ldots + \rho^{m-1} d(u_0, u_1) \\
&= d(u_0, u_1)\rho^n(1+\rho+\rho^2+\ldots+\rho^{m-n-1}) \leq d(u_0, u_1)\, \rho^n/(1-\rho),
\end{aligned}
$$

Notice here that we can replace the finite sum by an infinite series because we have the condition $\rho < 1$ which is necessary for a geometric series to converge. These ideas are the fundamental ones concerning contraction mappings which will allow us to develop our fixed point version of the power method, inverse power method, and eventually the Rayleigh Quotient Method.

The General Power Method

Before discussing how we plan to improve the Power Method I need to discuss the properties of the general power method for approximating eigenvalues of a particular n x n matrix **A.** If our matrix **A** has a largest eigenvalue say λ_1 then for most any vector v, the vectors of the transformations $A^k v$ are on the eigenvector corresponding to our dominant eigenvalue λ_1. Next we assume that A has n linearly independent eigenvectors $x_1, x_2, x_3, \ldots, x_n$ which correspond to eigenvalues $\lambda_1, \lambda_2, \lambda_3, \ldots, \lambda_n$, finally we assume that these eigenvalues are ordered according to magnitude such that $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \ldots \geq |\lambda_n|$.

Now assuming that we are working in the space $R^{n,}$ the vectors $x_1, x_2, x_3, \ldots, x_n$ form a basis for R^n and hence any vector we choose, say v_0, can be expressed as a linear combination of our eigenvectors in the form:

$$v_0 = \gamma_1 x_1 + \gamma_2 x_2 + \ldots + \gamma_n x_n \qquad (3)$$

Now $v_k = Av_{k-1}/\| Av_{k-1} \|$ is the algorithm, for $k \geq 1$.

Because of the ordering of our eigenvalues, i.e., assuming that λ_1 is the largest of the values in magnitude, it follows that,

$\lim_{v\to\infty} (\lambda_i / \lambda_1)^v = 0$, and thus we see that if $\gamma_1 \neq 0$ then $\lim_{k\to\infty} v_k = \gamma_1 x_1.$ (4)

The above is the usual power method, but it is useful to make some comments on the limitations and practicality of using the above algorithm. A problem with this algorithm is that the computation of A^k is incredibly lengthy and thus does not form the basis for a realistic numerical method. To avoid this we can notice that:

$$v_{k+1} = Av_k/\|Av_k\|, (k = 0,1,2,\ldots) \qquad (5)$$

Or with a new scaling factor we have that the recursion now becomes:

$$V_{k+1} = Av_k/\tau_k \; (k = 0,1,2,\ldots) \qquad (6)$$

Now we have a much more realistic method which involves only vector and matrix multiplication. Now we have what is considered the basis of the General Power Method (6) for finding a dominant eigenvalue λ_1 of a given matrix A. Now that I have given the basic groundwork for the Power Method I will proceed to discuss the modifications we plan to make in our reformulation of the Power Method.

Reformulation

Our reformulation of the power method will be such that, as mentioned, we can use the contraction map techniques described earlier. In order to do this, instead of normalizing as in the usual power method (scaling such that the largest component of v_k is unity) we can go ahead and divide by the dominant eigenvalue λ_1 and employ the contraction map techniques to find a solution by iterative procedures in the hyper plane $\langle v,x_1 \rangle = \gamma_1$. Normally we do not know the eigenvalue λ_1, so this makes it difficult to compute the denominator in this algorithm.

This process is not quite applicable at this point but will move from theoretical to practical when using Schwartz or Temple quotients namely $\langle Av_k, v_k \rangle / \langle v_k, v_k \rangle$, to approximate the eigenvector λ_1. After this is undertaken it is easy to extend this process to the shifted Inverse Power Method, and once again use contraction techniques to determine what the algorithm will converge to given an initial guess.

Looking Ahead

In the second part of the semester I plan on finishing the analysis of the modified power method using Schwartz quotient approximations to λ_1. Also I will easily extend my methods to the Inverse power method, and include that reformulation in my final report. After finishing the analysis of the Power Method and Inverse Power Method I will proceed to the rapidly convergent, albeit often unpredictable, Rayleigh quotient algorithm. It is hoped that in the end our modified version of the Rayleigh quotient algorithm will be not only rapidly convergent but also predictable as to what eigenvector and eigenvalue pair it converges to. All of these topics along with some numerical examples to illustrate these methods will comprise my final report on using fixed point formulations to numerically compute eigenvalues of symmetric matrices.

References

(1) C. Beattie and D.W. Fox 1989 *Localization Criteria and Containment for Rayleigh Quotient Iteration*, SIAM J. Matrix Analysis Appl, 80-93

(2) Burden and Faires 1997 *Numerical Analysis, Sixth Edition* 548-554

(3) B.N. Parlett 1980 *The Symmetric Eigenvalue Problem* 59-79

(4) I. Stakgold 1979 *Green's Functions and Boundary Value Problems* 221-259

(5) G.W. Stewart 1973 *Introduction to Matrix Computation* 340-354