

Characteristic Polynomials of Real Symmetric Matrices Generated from Random Graphs

Anagha Deshmane¹
American Institute of Mathematics
360 Portage Avenue
Palo Alto, CA 94306

July 17, 2003

¹E-mail: anaghad@mit.com

Abstract

Eigenvalues are the roots of the characteristic polynomials of an $n \times n$ matrix. In this paper we will use 4×4 real symmetric matrices to represent random graphs, and we will study the eigenvalues of these matrices. We will show that if each non-diagonal entry in a 4×4 matrix is chosen from a uniform distribution on the interval $[0, \alpha]$, where $0 \leq \alpha$, then the eigenvalues of the matrix are rescaled by a factor α . We will also show that if each non-diagonal entry in the matrix is chosen from a uniform distribution on the interval $[\beta, \beta + 1]$, where $0 \leq \beta$, then roughly $\frac{3}{4}$ of the eigenvalues will be concentrated around $-\beta$ and roughly $\frac{1}{4}$ of them will be concentrated around 3β . An investigation into the application of eigenvalues is beyond the scope of this extended essay. However, studying the distribution of these eigenvalues is analogous to studying the roots of a general polynomial with real coefficients. By studying how the distribution of the eigenvalues shifts based on how the inputs in the matrix are altered, we can understand how the roots of a polynomial shift if the coefficients of the polynomial are altered. Thus we expand the problem of the distribution of eigenvalues to encompass the roots of polynomials of the general form

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_0.$$

In this paper we will quantify how much each coefficient can be shifted such that the real roots of the polynomial do not move outside a small window.

Contents

1	Background	2
1.1	Random Graphs, d -Regular Graphs, and Their Matrices	2
1.2	Weighted Matrices	4
1.3	Eigenvalues	4
2	Investigation 1: 4×4 Matrices	7
2.1	Rescaled Intervals	7
2.1.1	Experiments	7
2.1.2	Theory	8
2.2	Shifted Intervals	10
2.2.1	Experiments	10
2.2.2	Theory	11
3	Investigation 2: $n \times n$ Matrices	13
4	Conclusion	16

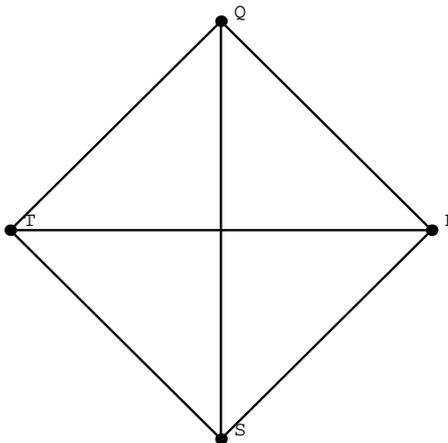
Chapter 1

Background

1.1 Random Graphs, d -Regular Graphs, and Their Matrices

A graph is a collection of vertices (points) and edges (lines which connect the vertices). In a random graph, the vertices are randomly connected to each other. A d -regular graph is a type of random graph in which each vertex is randomly connected to d other vertices. For example, in a 3-regular graph, each vertex is connected to 3 other vertices.

Suppose we make a 3-regular graph with 4 vertices. The only way to construct such a graph is to connect each vertex to every other vertex. Let's label the vertices as Q , R , S , and T . The graph will look like this:



We can represent any graph with n vertices with an $n \times n$ real symmetric matrix (meaning that the entries in the matrix are real numbers, and the entries above the diagonal are reflected below the diagonal). This matrix is called the graph's adjacency matrix. Each column and each row of the adjacency matrix represents a vertex. A 1 is placed in a position if the two vertices that the position represents are connected; a 0 is placed in that position if the two vertices are not connected. The graphs we are investigating do not allow a vertex to be connected to itself, so we place zeros in every position on the diagonal of the adjacency matrix. In a d -regular graph, there are d 1s in each row and d 1s in each column, because there are d connections to each vertex.

In our example of a 3-regular graph with 4 vertices, the graph's adjacency matrix will be a 4×4 real symmetric matrix. The first column and the first row represent the vertices connected to vertex Q , the second column and row represent the vertices connected to vertex R , the third column and row represent the vertices connected to vertex S , and the fourth column and row represent the vertices connected to vertex T . In general, the matrix will look like this:

$$M = \begin{pmatrix} q, q & q, r & q, s & q, t \\ r, q & r, r & r, s & r, t \\ s, q & s, r & s, s & s, t \\ t, q & t, r & t, s & t, t \end{pmatrix}$$

We will represent our graph by a real symmetric matrix with 0s on the diagonal and 1s above and below the diagonal:

$$M = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$$

1.2 Weighted Matrices

Suppose our graph and matrix represents a network of n computers. In this case, it is more beneficial to consider the connection between a pair of vertices as a number other than 1. For example, each entry may represent the speed of a connection between the computers – a number greater or less than 1. Therefore, we can choose each non-diagonal entry in the matrix as a weighted value from a uniform distribution on various intervals.

A uniform distribution on a given interval will yield each subinterval within that interval with equal probability. For example, if we have a uniform distribution on the interval $[0, 1]$, the probability of observing x in a subinterval $[a, b]$ equals $b - a$ for any subinterval within $[0, 1]$. Thus, choosing a number from $[0.1, 0.2]$ has the same probability as choosing from $[0.48, 0.58]$.

Here, we will study the cases in which each such element of the matrix is chosen from uniform distributions on the intervals $[0, \alpha]$, and $[\beta, \beta + 1]$, where α and β are real numbers.

1.3 Eigenvalues

Eigenvalues can be useful in many sections of mathematics and physics. For example, eigenvalues play an important part in the physics of rotating bodies; and if we form a matrix from a random graph, the second eigenvalue reveals information about the graph's connectivity. However, for our purposes, eigenvalues are the solutions to an interesting subset of polynomials.

By definition, λ is an eigenvalue if

$$A\vec{v} = \lambda\vec{v}, \tag{1.1}$$

where A is an $n \times n$, or square, matrix and \vec{v} is a column vector with n components. We call \vec{v} an eigenvector, which is a vector other than the zero vector. Matrix A must be square because the output of the product $A\vec{v}$ is a vector and also has n entries. The output vector is a multiple of \vec{v} , which means that matrix A is a transformation that does *not* change the direction of \vec{v} ; the matrix only changes the length of the vector. Note that not every vector is an eigenvector. Each square matrix has its own eigenvectors, depending on what the eigenvalues of the matrix are.

If we are given a square matrix, we can easily compute its eigenvalues by generating the characteristic polynomial of the matrix. Referring to Equation 1.1, we first collect like terms:

$$\begin{aligned} 0 &= A\vec{v} - \lambda\vec{v} \\ &= (A - \lambda I)\vec{v}. \end{aligned}$$

We must multiply λ by an $n \times n$ identity matrix, I , so that $A - \lambda I$ forms an $n \times n$ matrix. The determinant of this matrix is a degree n polynomial that is equal to zero, because the matrix sends \vec{v} to zero. By solving for λ , we can find the n roots of this characteristic polynomial, which are the eigenvalues of matrix A .

Let M be a 4×4 real symmetric matrix formed from a 3-regular graph:

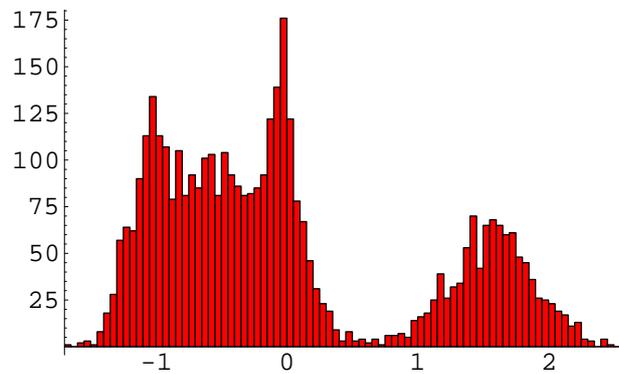
$$M = \begin{pmatrix} 0 & a & b & c \\ a & 0 & d & e \\ b & d & 0 & f \\ c & e & f & 0 \end{pmatrix}.$$

To find the eigenvalues of matrix M , we must find and solve its characteristic polynomial, $P(\lambda)$. This, we have

$$\begin{aligned} P(\lambda) &= \det(A - \lambda I) \\ &= \begin{vmatrix} -\lambda & a & b & c \\ a & -\lambda & d & e \\ b & d & -\lambda & f \\ c & e & f & -\lambda \end{vmatrix} \\ &= \lambda^4 - \lambda^2(a^2 + b^2 + c^2 + d^2 + e^2 + f^2) \\ &\quad - 2\lambda(abd + ace + bcf + def) \\ &\quad + (c^2d^2 - 2bcde + b^2e^2 - 2acdf - 2abef + a^2f^2). \end{aligned} \tag{1.2}$$

The non-diagonal entries of the matrix, a, b, c, d, e , and f , comprise the coefficients of the polynomial $P(\lambda)$. Suppose we take the non-diagonal entries from a uniform distribution on the interval $[0, 1]$. If we look at a sufficiently large number of matrices, what does the projection of the eigenvalues onto the real axis

look like? Below is a histogram of the distribution. We will use this histogram¹ to compare results for cases in which $a, b, c, d, e,$ and f are taken from uniform distributions on different intervals.



¹The range over which the eigenvalues vary is divided into many bins of equal size, each of which represents a subinterval in the range. The histogram depicts how many eigenvalues fall into each bin.

Chapter 2

Investigation 1: 4×4 Matrices

We will begin by studying 4×4 real symmetric matrices, such as that shown below:

$$M = \begin{pmatrix} 0 & a & b & c \\ a & 0 & d & e \\ b & d & 0 & f \\ c & e & f & 0 \end{pmatrix} \quad (2.1)$$

The 4 eigenvalues of this matrix react in distinct ways when the non-diagonal entries $a, b, c, d, e,$ and f are chosen from uniform distributions on the intervals $[0, \alpha]$ and $[\beta, \beta + 1]$, where $0 \leq \alpha$ and $0 \leq \beta$. If the interval is rescaled by a factor α , then the eigenvalues are rescaled by the same factor. If the interval is shifted to $[\beta, \beta + 1]$, then roughly $\frac{3}{4}$ of the eigenvalues will be concentrated around $-\beta$ and roughly $\frac{1}{4}$ will be concentrated around 3β .

2.1 Rescaled Intervals

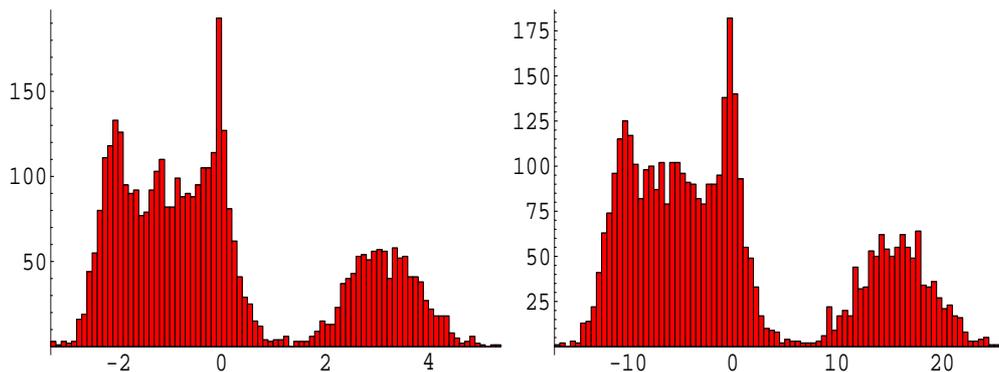
We have already seen the distribution of the eigenvalues if the non-diagonal values are taken from the interval $[0, 1]$. We will now examine what happens when the interval is rescaled by a factor α , where $0 \leq \alpha$, so that the non-diagonal values are chosen from the interval $[0, \alpha]$.

2.1.1 Experiments

Using *Mathematica*, we can average over a finite number of weighted graphs, form their corresponding real symmetric 4×4 matrices that are structured like

matrix M , and observe how the projection of the matrices' eigenvalues onto the real axis is altered when the non-diagonal entries are taken from rescaled intervals.

Consider the cases in which we take each non-diagonal entry in matrix M from a uniform distribution on the intervals $[0, 2]$ and $[0, 10]$. Below are histograms of these matrices.



The two graphs differ only slightly in shape and have roughly the same number of eigenvalues in each bin. However, the overall range of the eigenvalues changes, as do the bin sizes. In the first histogram, the range of the eigenvalues is rescaled by a factor of 2; in the second, by a factor of 10. The same results occur if the interval from which a, b, c, d, e , and f are chosen is rescaled by any real number greater than or equal to zero.

2.1.2 Theory

Assume we rescale the interval of each non-diagonal entry in our 4×4 adjacency matrix by a factor α , so that the entries are chosen from a uniform distribution on $[0, \alpha]$. Then the eigenvalues are rescaled by a factor α .

Recall that each $n \times n$ matrix is associated with a characteristic polynomial of degree n , similar to that in Equation 1.2. The n roots of the polynomial are the eigenvalues of the matrix.

The polynomial for our 4×4 matrix is

$$\begin{aligned}
 P(\lambda) = & \lambda^4 - \lambda^2(a^2 + b^2 + c^2 + d^2 + e^2 + f^2) \\
 & - 2\lambda(abd + ace + bcf + def) \\
 & + (c^2d^2 - 2bcde + b^2e^2 - 2acdf - 2abef + a^2f^2).
 \end{aligned}$$

Note that this polynomial does not have a cubic term. The characteristic polynomial for an $n \times n$ matrix is given by

$$\det(A - \lambda I) = \lambda^n - Tr(A)\lambda^{n-1} + \dots,$$

where $Tr(A)$ is the trace of the matrix (the sum of all its diagonal elements). In an adjacency matrix, each diagonal element is zero, so $Tr(M) = 0$. Therefore there is no λ^{n-1} term in the polynomial.

For the sake of simplicity, we can rewrite the expression by replacing the coefficients such that

$$P(\lambda) = C_0 + C_1\lambda + C_2\lambda^2 + \lambda^4,$$

where C_0, C_1 , and C_2 are polynomials of degree 4, 3, and 2 (respectively) in the coefficients a, b, c, d, e , and f . Now suppose that each of the non-diagonal values in matrix M (a, b, c, d, e , and f) is chosen from a uniform distribution on the interval $[0, \alpha]$. Then writing a as $\alpha a'$, b as $\alpha b'$, c as $\alpha c'$ and so on, we see that

$$C_i(a, b, c, d, e, f) = \alpha^{4-i} C_i(a', b', c', d', e', f').$$

We now have

$$P(\lambda) = C_0\alpha^4 + C_1\alpha^3\lambda + C_2\alpha^2\lambda^2 + \lambda^4.$$

We want to determine the eigenvalues of the matrix, so we must equate $P(\lambda)$ to zero. We can then manipulate the expression to find the roots:

$$\begin{aligned} 0 &= C_0\alpha^4 + C_1\alpha^3\lambda + C_2\alpha^2\lambda^2 + \lambda^4 \\ 0 &= \frac{C_0\alpha^4 + C_1\alpha^3\lambda + C_2\alpha^2\lambda^2 + \lambda^4}{\alpha^4} \\ &= C_0 + C_1\left(\frac{\lambda}{\alpha}\right) + C_2\left(\frac{\lambda}{\alpha}\right)^2 + \left(\frac{\lambda}{\alpha}\right)^4. \end{aligned}$$

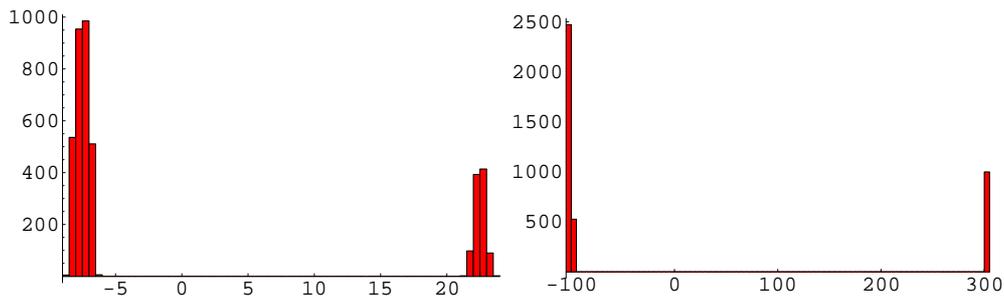
The roots of this expression are $\frac{\lambda}{\alpha}$. The eigenvalues (λ) are therefore rescaled by a factor, α . Thus, when each non-diagonal value of a 4×4 real symmetric matrix is rescaled by a factor α , the eigenvalues are rescaled by the same factor. This result can easily be reproduced for an $n \times n$ matrix.

2.2 Shifted Intervals

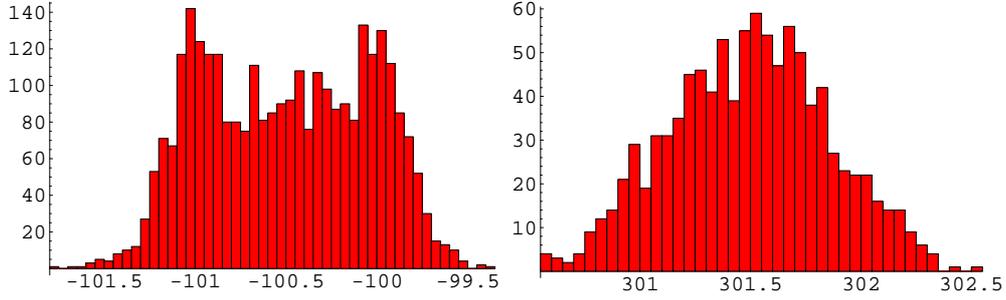
Let us return to matrix M in line 2.1. Now we will choose non-diagonal entries $a, b, c, d, e,$ and f from a uniform distribution that is shifted from its original position on the interval $[0, 1]$ to a new position on the interval $[\beta, \beta + 1]$, where $0 \leq \beta$.

2.2.1 Experiments

We again use *Mathematica* to average over a finite number of weighted graphs, form corresponding real symmetric 4×4 matrices that are structured like matrix M , and observe how the projection of the matrices' eigenvalues onto the real axis is altered when the non-diagonal entries are taken from shifted intervals. If we shift the range of these entries (which also make up the coefficients of the characteristic polynomial of the matrix) to the interval $[7, 8]$ or $[100, 101]$, we will observe the projection of the eigenvalues onto the real axis separating into two masses:



Although it may look as if the sections of the histograms are converging to certain values, a rescaled graph shows otherwise. In fact, the two sections are nearly the same in size and shape as they were when $a, b, c, d, e,$ and f were taken from the interval $[0, 1]$. They have simply moved apart. Below are "zoomed-in" views of the two sections of the histogram of eigenvalues generated from matrices whose non-diagonal values are taken from a uniform distribution on the interval $[100, 101]$.



We see that when $a, b, c, d, e,$ and f are taken from the interval $[100, 101]$, one section of the eigenvalues' distribution is shifted from approximately -1 to approximately -100 , and the other shifts from approximately 3 to approximately 300 . The same shifts in the distribution of the eigenvalues occurs if $a, b, c, d, e,$ and f are taken from β plus some number chosen from a uniform distribution on the interval $[0, 1]$, as β gets large.

2.2.2 Theory

Assume we shift the interval of each non-diagonal entry in our 4×4 adjacency matrix from $[0, 1]$ to $[\beta, \beta + 1]$. Then the eigenvalues are shifted to $-\beta$ and 3β .

Recall the characteristic polynomial from line 1.2. If $a, b, c, d, e,$ and f were exactly equal to β , we would have

$$\begin{aligned}
P(\lambda) &= \lambda^4 - \lambda^2(a^2 + b^2 + c^2 + d^2 + e^2 + f^2) \\
&\quad - 2\lambda(abd + ace + bcf + def) \\
&\quad + (c^2d^2 - 2bcde + b^2e^2 - 2acdf - 2abef + a^2f^2) \\
&= \lambda^4 - 6\beta^2\lambda^2 - (4\beta^3)2\lambda \\
&\quad + (\beta^4 - 2\beta^4 + \beta^4 - 2\beta^4 - 2\beta^4 + \beta^4) \\
&= \lambda^4 - 6\beta^2\lambda^2 - 8\beta^3\lambda - 3\beta^4.
\end{aligned}$$

Again, because we are solving for the eigenvalues, we can set $P(\lambda) = 0$ to get

$$0 = \left(\frac{\lambda}{\beta}\right)^4 - 6\left(\frac{\lambda}{\beta}\right)^2 - 8\left(\frac{\lambda}{\beta}\right) - 3.$$

There are four solutions to this equation:

$$\frac{\lambda}{\beta} = -1, -1, -1, 3.$$

The solutions to this polynomial make sense. If the non-diagonal values in our matrix are chosen on the interval $[\beta, \beta + 1]$, they do not vary much from β , and the eigenvalues remain close to those corresponding to the eigenvalues found if a, b, c, d, e , and f were exactly equal to β . Therefore, we would expect roughly $\frac{3}{4}$ of the eigenvalues to be concentrated around $-\beta$ and roughly $\frac{1}{4}$ of them to be concentrated around 3β .

Furthermore, as β gets very large, a random number chosen from a uniform distribution on the interval $[0, 1]$ becomes small relative to β . Thus in the limit, if a, b, c, d, e , and f are chosen from $\beta + k$, where k is some random number chosen from a uniform distribution on the interval $[0, 1]$, the eigenvalues behave as if the non-diagonal values were chosen from the interval $[\beta, \beta + 1]$.

Chapter 3

Investigation 2: $n \times n$ Matrices

If we have a 4×4 matrix, it is relatively easy to observe how the eigenvalues shift if the non-diagonal entries in the matrix are chosen from altered intervals. However, it is difficult to do so for a larger matrix, such as a 10×10 matrix. If each non-diagonal entry is taken from a uniform distribution on an interval that has been rescaled by a certain factor, then the eigenvalues will be rescaled by that factor. But if we take each non-diagonal element from a uniform distribution on a shifted interval, how much should we expect the eigenvalues to move?

The determinant of a large real matrix yields a large polynomial with real coefficients. For an $n \times n$ matrix, it is of the form

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_0.$$

Assume that at some real simple root, r , there are two points very close to the root, d_1 and d_2 , $d_1 < r < d_2$, with $p(d_1) < 0$ and $p(d_2) > 0$. Assume also that we have a new polynomial $q(x)$ such that

$$q(x) = (a_n + s_n)x^n + (a_{n-1} + s_{n-1})x^{n-1} + \dots + (a_0 + s_0),$$

where $s_n, s_{n-1}, s_{n-2}, \dots, s_0$ are small amounts by which we shift each coefficient of the original polynomial $p(x)$.

Theorem 3.0.1. *If r is a real simple root (its order is 1) of $p(x)$, then if we change the coefficients by small amounts s_i ,*

$$|s_i| < \frac{|p(d_1)|}{(1 + |d_1|^n)(n + 1)}$$

and

$$|s_i| < \frac{|p(d_2)|}{(1 + |d_2|^n)(n + 1)},$$

the new polynomial has a real root near r .

Proof. First consider d_1 . Expanding $q(x)$, we have

$$\begin{aligned} q(x) &= a_n x^n + s_n x^n + a_{n-1} x^{n-1} + s_{n-1} x^{n-1} + \dots + a_0 + s_0 \\ &= p(x) + s_n x^n + s_{n-1} x^{n-1} + \dots + s_0. \end{aligned} \quad (3.1)$$

We want $q(d_1)$ to remain negative, which will be true if

$$|p(d_1)| > |s_n d_1^n + s_{n-1} d_1^{n-1} + \dots + s_0|.$$

If we consider only the last term on the right hand side of the equation, s_0 , we would have

$$|s_0| < |p(d_1)|.$$

If we consider only the $s_1 d_1$ term, we would have

$$|s_1| < \frac{|p(d_1)|}{|d_1|}.$$

In Equation 3.1 we want to ensure that each term in the sum is less than $|p(d_1)|$, so we divide each addition by the highest power of d_1 . However, if $d_1^n \leq 1$, each term will be as large as or greater than $|p(d_1)|$; if we make the denominator of each s_i greater than 1, then each term will be less than $|p(d_1)|$. Therefore we add 1 to the denominator of each term. Also, Equation 3.1 has $n + 1$ terms added to $|p(d_1)|$. To ensure that the sum of all the s terms is less than $|p(d_1)|$, we divide each term by $n + 1$. Thus, $q(d_1)$ remains negative, if each s_i satisfies

$$|s_i| < \frac{|p(d_1)|}{(1 + |d_1|^n)(n + 1)}. \quad (3.2)$$

Now consider d_2 . We want $q(d_2) > 0$. So in Equation 3.1, we want the s_i terms to be less than $|p(d_2)|$ in absolute value. Using the same process we used to find $|s_i|$ above, we find that $q(d_2)$ remains positive if

$$|s_i| < \frac{|p(d_2)|}{(1 + |d_2|^n)(n + 1)}. \quad (3.3)$$

□

Theorem 3.0.2. Assume $p(x)$ has n real simple roots. Then if we change the coefficients by small amounts s_i ,

$$|s_i| < \frac{|p(d_1)|}{(1 + |d_1|^n)(n + 1)}$$

and

$$|s_i| < \frac{|p(d_2)|}{(1 + |d_2|^n)(n + 1)},$$

the new polynomial has n real simple roots, where each root is near a previous one.

Proof. Each root of $q(x)$ is associated with two values for $|s_i|$. If $|s_i|$ is less than the minimum of the $2n$ values for $|s_i|$ generated by Equations 3.2 and 3.3, then $q(d_1)$ will remain negative and $q(d_2)$ will remain positive. In fact, if d_1 and d_2 exist such that $d_1 = r - \Delta r$ and $d_2 = r + \Delta r$, where Δr is less than $\frac{1}{4}$ the difference between two adjacent roots, then the roots of $q(x)$ will be less than Δr away from the roots of $p(x)$, and there will not be more than one root in the interval $[d_1, d_2]$. Thus, shifting the coefficients of a polynomial by a small amount does not have a large effect on the roots. \square

This result explains why choosing the non-diagonal entries in matrix M (a , b , c , d , e , and f) from $\beta + k$, where k is chosen from a uniform distribution on the interval $[0, 1]$, is effectively the same as choosing each non-diagonal entry from a uniform distribution on the interval $[\beta, \beta + 1]$. Shifting the coefficients of the characteristic polynomial from β to β plus some small amount does not significantly change the distribution of the roots (eigenvalues), so the distribution looks very close to the distribution where each non-diagonal element is chosen from the interval $[\beta, \beta + 1]$.

Chapter 4

Conclusion

What began as an investigation of random d -regular graphs and their matrices has evolved into the much broader investigation of a certain class of polynomial. We first observed how rescaling or shifting the entries in a 4×4 real symmetric matrix affected the eigenvalues of the matrix. It should be noted that our results for rescaled intervals hold true for any $n \times n$ matrix. However, our results for shifted intervals are specific to 4×4 matrices.

Because the eigenvalues are the solution to a characteristic polynomial, we found that our observations of the behavior of the eigenvalues under certain conditions are analogous to the behavior of the roots of a polynomial under similar conditions.

If we change the coefficients of a polynomial, it is possible that some real roots will become complex, and that some complex roots will become real. For example, if there was a double root, shifting the coefficients can yield either two distinct real roots or two distinct complex roots. It can be proven that the complex roots will be nearby the original double root, but the proof requires tools derived from Complex Analysis.

However, by expanding our investigation to general polynomials, we can now quantify how much each coefficient in the polynomial can safely be shifted, so that the roots stay within a desired range. This is a handy tool, and can be used in many situations involving similar polynomials.