Finding the Roots of a Polynomial with Eigenvalues

Text Reference: Section 6.4, p. 400

The purpose of this set of exercises is to show how we can find the real roots of a polynomial by finding the eigenvalues of a particular matrix. We will find these eigenvalues by the QR method described below.

Let us begin by recalling that a **polynomial of degree** n is a function of the form

$$p(t) = a_0 + a_1 t + \ldots + a_{n-1} t^{n-1} + a_n t^n$$

where a_0, a_1, $\ldots$, a_{n-1}, and a_n are real numbers with $a_n \neq 0$. A **root** of a polynomial is a value of t for which $p(t) = 0$. It is often necessary (especially in calculus-based applications) to find all of the real roots of a given polynomial. In practice this can be a difficult problem even for a polynomial of low degree. For a polynomial of degree 2, every algebra student learns that the roots of $at^2 + bt + c$ can be found by the quadratic formula

$$t = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

If the polynomial is of degree 3 or 4, then there are formulas somewhat resembling the quadratic formula (but much more involved) for finding all the roots of a polynomial. However there is no general formula for finding the roots of a polynomial of degree 5 or higher.

Example: Consider the cubic polynomial $p(t) = t^3 - 2t^2 - 5t + 6$. We can factor this polynomial rather easily to find that its roots are $t = 1$, $t = -2$, and $t = 3$.

If a polynomial cannot easily be factored, we will need to use numerical techniques to find a polynomial's roots. There are problems with this approach as well. Algorithms such as Newton's Method may not converge to a root, or may approach the root very slowly. These methods must also be applied repeatedly to find all of the roots, and usually require a cleverly chosen starting guess for the root we are seeking. However, we will note that there is an algorithm from linear algebra which may be used to find the real roots of a polynomial simultaneously.

Recall that the eigenvalues of a $n \times n$ matrix A are the roots of the characteristic polynomial of A, which is defined as $p(\lambda) = \det(A - \lambda I_n)$ and is a polynomial of degree n. So if we happened to know the eigenvalues of A, we would know the roots of $p(\lambda)$. To find the roots of any polynomial p, then, we would need two things:

1. A way of finding a matrix A whose characteristic polynomial is $p(\lambda)$.
2. A way of finding the eigenvalues of this A which does not depend on finding the roots of $p(\lambda)$, since that is what we are trying to accomplish.

We solve the first problem by defining the **companion matrix** for a polynomial

$$p(t) = t^n + a_{n-1}t^{n-1} + \ldots + a_1 t + a_0$$

Definition: If $p(t)$ is as given above, then the **companion matrix** for p is

$$C_p = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 1 \\ -a_0 & -a_1 & -a_2 & \ldots & -a_{n-1} \end{bmatrix}$$

Example (cont.): The companion matrix for the polynomial $p(t) = t^3 - 2t^2 - 5t + 6$ is

$$C_p = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -6 & 5 & 2 \end{bmatrix}$$

Questions:

1. Find the companion matrices for the following polynomials.

   a) $p(t) = t^2 + 2t - 4$
   b) $p(t) = t^3 - 9t^2 + 12t + 22$
   c) $p(t) = t^4 - 2t^3 - 13t^2 + 14t + 24$

2. Find the characteristic polynomials of the matrices you found in Exercise 1. What do you notice?

3. Show that the characteristic polynomial of a companion matrix for the n^{th} degree polynomial $p(t)$ is $\det(C_p - I_n) = (-1)^n p(\lambda)$ as follows.

   - Show that if C_p is the companion matrix for a quadratic polynomial $p(t) = t^2 + a_1 t + a_0$, then $\det(C_p - I_2) = p(\lambda)$ by direct computation.
   - Use mathematical induction to show that the result holds for $n \geq 2$. Hint: expand the necessary determinant by cofactors down the first column.

Thus we now have a way to create the matrix A whose characteristic polynomial is $p(t)$. We also need a method for finding the eigenvalues of A which does not use the characteristic polynomial. One method which accomplishes this is called the **QR method** because it is based on the QR decomposition of A. Another exercise set which accompanies the text studies this method in depth. To introduce the algorithm, we first establish some properties that underlie this QR method for finding eigenvalues.

Question:

4. Suppose A is a $n \times n$ matrix. Let $A = Q_0 R_0$ be a QR factorization of A and create $A_1 = R_0 Q_0$. Let $A_1 = Q_1 R_1$ be a QR factorization of A_1 and create $A_2 = R_1 Q_1$.

   a) Show that $A = Q_0 A_1 Q_0^T$.
   b) Show that $A = (Q_0 Q_1) A_2 (Q_0 Q_1)^T$
   c) Show that $Q_0 Q_1$ is an orthogonal matrix.
   d) Show that A, A_1, and A_2 all have the same eigenvalues.

The QR method for finding the eigenvalues of an $n \times n$ matrix A extends this process to create a sequence of matrices with the same eigenvalues.

The QR Method:

Step 1: Let $A = Q_0 R_0$ be a QR factorization of A; create $A_1 = R_0 Q_0$.

Step 2: Let $A_1 = Q_1 R_1$ be a QR factorization of A_1; create $A_2 = R_1 Q_1$.

Step 3: Continue this process; Once A_m has been created, let $A_m = Q_m R_m$ be a QR factorization of A_m and create $A_{m+1} = R_m Q_m$.

Step 4: Stop the process when the entries below the main diagonal of A_m are sufficiently small, or stop if it appears that convergence will not happen.

Example (cont.): Let A be the companion matrix for the polynomial in our example; that is,

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -6 & 5 & 2 \end{bmatrix}$$

We find that the QR decomposition of this matrix is

$$Q_0 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 0 & 0 \end{bmatrix} \qquad R_0 = \begin{bmatrix} 6 & -5 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

so

$$A_1 = R_0 Q_0 = \begin{bmatrix} 2 & 6 & -5 \\ 0 & 0 & 1 \\ -1 & 0 & 0 \end{bmatrix}$$

We perform this operation again, getting

$$A_2 = R_1 Q_1 = \begin{bmatrix} 2.236 & 5.367 & -4.472 \\ 0 & 2.683 & -2.236 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} .8944 & .4472 & 0 \\ 0 & 0 & 1 \\ -.4472 & .8944 & 0 \end{bmatrix} = \begin{bmatrix} 4 & -3 & 5.367 \\ 1 & -2 & 2.683 \\ -.4472 & .8944 & 0 \end{bmatrix}$$

This matrix is still far from upper triangular, so we continue the process. Eventually we find that

$$A_{10} = \begin{bmatrix} 3.0296 & -4.9933 & 5.0823 \\ .0297 & -2.0279 & 1.3963 \\ .0001 & .0036 & .9983 \end{bmatrix}$$

so the matrix is converging to an upper triangular matrix, and its diagonal elements are converging to the roots of $p(t)$: $t = 3$, $t = -2$, and $t = 1$.

Question:

5. Approximate the roots of the polynomials in Exercise 1 by applying the QR Method to their companion matrices. Iterate until the entries below the main diagonal are all below 0.1.