Math REU Apprentice Program Alternate Proof of the Spectral Theorem

Julian Hartman

September 17, 2009

Abstract

This paper is a proof of the Spectral Theorem; a theorem that was proven in the apprentice class but not in this manner. Because the Spectral Theorem is such a fundamental theorem in linear algebra, it is important to explore it further through an alternate proof.

Contents

1	Intro	1
2	Lemmas	2
3	The Spectral Theorem	3
4	Exploration	4
5	Examples	4

1 Intro

Specifically, this proof differs from the one done in class by the use of a matrix's characteristic polynomial- a concept that we will introduce in the paper and prove several lemmas about. The statement of the theorem is: If A is an $n \times n$ symmetric real matrix, then \exists an orthanormal basis of real eigenvectors for A. First we will prove a series of lemmas. These will bring us to a point where we can prove the Spectral Theorem true through simple induction. The theorem is particularly useful, because if we know that an orthonormal basis of eigenvectors exists, it becomes much easier to find a basis of eigenvectors for a given matrix. Finding a basis of eigenvectors is very important because it makes calculations using the matrix easier.

Since you know that this basis of eigenvectors exists, it becomes (relatively) easy to find it. To do so, one must find an eigenvector (by finding a zero of the characteristic polynomial). Then, try to find an eigenvector which is orthogonal to the first. If this is not possible, try another eigenvector. If it is, find a third vector that is orthogonal to both, and continue until you have a complete basis of eigenvectors. Then normalize them by dividing each vector by its length, and at this point the collection of eigenvectors forms an orthonormal basis. After the theorem is proven, an example of this process will be shown.

Definition 1. Let A be a linear map from V to V. Let W be a subspace of V. W is A-invariant if for all w in W, Aw is in W.

Definition 2. The characteristic polynomial of a square matrix A is defined as f(t) = det(tI - A)

Definition 3. An eigenvector of a matrix A is a nonzero vector v such that $Av = \lambda v$ where λ is some constant. λ is known as an eigenvalue of the matrix A.

2 Lemmas

Lemma 1. The determinant of a matrix is 0 iff the kernel is nontrivial.

Proof. Let A be an arbitrary $n \times n$ matrix.

Suppose that its kernel is nontrivial. This implies that A maps some vector to zero other than the zero vector. Therefore A is not one-to-one, and thus not invertible. Because A is a singular matrix, its determinant is 0.

Suppose that the determinant of the matrix is 0. Then the matrix A is not invertible. This means that the linear map is not injective, and therefore the kernel is nontrivial.

Lemma 2. The zeros of the characteristic polynomial of a matrix are the matrix's eigenvalues.

Proof. Suppose $f(\lambda) = 0$. This means that $det(\lambda I - A) = 0$. The determinant of a matrix is 0 iff the kernel is nontrivial. So, the fact that f(t) has a zero means that the kernel of $(\lambda I - A)$ is nontrivial. That is, for some nonzero v, $(\lambda I - A)v = 0$. This implies that $Av = \lambda v$, which is the definition of an eigenvalue and eigenvector. Thus, when $f(\lambda) = 0$, λ is an eigenvalue.

Suppose λ is an eigenvalue with eigenvector v. Then, $Av = \lambda v$. This implies that $\lambda v - Av = 0$ and then that $(\lambda I - A)v = 0$. This implies that the kernel of the matrix $(\lambda I - A)$ is nontrivial. Since the kernel of this matrix is nontrivial, $det(\lambda I - A) = 0$. Since f(t), the characteristic polynomial, is defined as f(t) = det(tI - A), $f(\lambda) = 0$.

Lemma 3. If A is a real symmetric square matrix then all the zeros of its characteristic polynomial are real.

Proof. Suppose A is a real symmetric square matrix. Suppose A is a real symmetric square matrix. Then, by Lemma 2, the complex eigenvalues of A are the zeros of its characteristic polynomial. By the Fundamental Theorem of Algebra, this polynomial has a complex root, and hence A has a complex eigenvalue. Therefore the eigenvector for that eigenvalue must exist. Let v be this eigenvector, \overline{v} be its complex conjugate (by which I mean the vector whose coordinates with respect to the standard basis are the complex conjugates of those of v), and A be a real symmetric square matrix operating on the vectors. Suppose λ is the eigenvalue of v.

 $\langle v, Av \rangle = v^T \overline{Av} = v^T A^T \overline{v}$, because A is symmetric. $v^T A^T \overline{v} = (Av)^T \overline{v}$. Because v is an eigenvector, $Av = \lambda v$ and thus $(\lambda v)^T \overline{v} = \lambda (v)^T \overline{v} = v^T \overline{Av} = v^T \overline{\lambda v} = \overline{\lambda} v^T \overline{v}$. Thus, $\lambda = \overline{\lambda} \implies \lambda$ is real. Then, by Lemma 2, Lemma 3 is true.

Definition 4. W^{\perp} is defined to be the collection of all vectors w^{\perp} such that for all w in W, $< w, w^{\perp} >= 0$.

Lemma 4. If A is symmetric, and W is an A-invariant subspace, then W^{\perp} is also an A-invariant subspace.

Proof. Let w^{\perp} be a vector in W^{\perp} , and w be a vector in W.

$$\langle w, Aw^{\perp} \rangle = w^T Aw^{\perp} = w^T A^T w^{\perp}$$
, because A is symmetric.
 $w^T A^T = (Aw)^T \implies w^T A^T w^{\perp} = (Aw)^T w^{\perp} = \langle Aw, w^{\perp} \rangle = 0$

Because W is A-invariant, and as we found that $\langle w, Aw^{\perp} \rangle = \langle Aw, w^{\perp} \rangle = 0, w \perp Aw^{\perp}.$ $\implies w^{\perp} \in W^{\perp} \implies W^{\perp}$ is A-invariant.

Lemma 5. v is an eigenvector of A iff the subspace V spanned by v is A-invariant.

Proof. Let v be an eigenvector for A. Thus, $Av = \lambda v$. λV is just $C \times v$, where C is a constant; therefore it spans the same space as $v \implies Av$ has the same span as $V \implies V$ is A invariant.

Suppose V is A invariant. $\implies Av$ has the same span as V. Therefore, $Av = C \times v$, where C is a constant. Let $C = \lambda$. Then, $Av = \lambda v$, and v therefore is an eigenvector of A.

3 The Spectral Theorem

Theorem 1. If A is an $n \times n$ symmetric real matrix, then \exists an orthanormal basis of real eigenvectors for A.

Proof. We will prove this theorem through induction. Suppose A is a $k \times k$ symmetric matrix. Let k = 1 Then, a basis for A consists of only one eigenvector. Because it only includes one vector, the basis is orthonormal. Thus, if k = 1, there exists an orthonormal basis of real eigenvectors for A.

Now suppose this theorem is true for some arbitrary k. Take a $k + 1 \times k + 1$ matrix. Chose an arbitrary eigenvector for the matrix. It covers a one dimensional subspace that is A invariant. Call this subspace W. By Lemma 4, W^{\perp} is a k dimensional subspaces that is A invariant. There is an orthogonal basis of eigenvectors for A for this k dimensional subspace, and the arbitrary eigenvector previously chosen is perpendicular to this subspace. This yields an orthogonal basis of eigenvectors in k + 1 dimensions. By dividing all the vectors in the orthogonal basis by their norm, we find an orthonormal basis of eigenvectors. Thus, by induction, the theorem is true.

4 Exploration

It is important to note that for the spectral theorem to be true, the matrix must be symmetric. It is simple to find a square real matrix that has no real eigenvalues or eigenvectors. Let $A = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$, and let $v = \begin{pmatrix} a \\ b \end{pmatrix}$. $Av = \begin{pmatrix} a+b \\ b-a \end{pmatrix}$. Suppose that v is an eigenvector for A. Then, $Av = \begin{pmatrix} a+b \\ b-a \end{pmatrix} = \begin{pmatrix} \lambda a \\ \lambda b \end{pmatrix} = \lambda v$. Consequently, $a + b = \lambda a$ and $b - a = \lambda b$. From this, we can say that $b = (\lambda - 1)a$, and substitute for b in the other equation. $(\lambda - 1)a - a = \lambda(\lambda - 1)a \implies (\lambda - 2)a = (\lambda^2 - \lambda)a$. By dividing a out, we find that $(\lambda - 2) = (\lambda^2 - \lambda)$, or $0 = \lambda^2 - 2\lambda + 2$. By the quadratic formula, $\lambda = (2 \pm \sqrt{4 - 4 \times 1 \times 2})/2 = (2 \pm \sqrt{4 - 8})/2 = (2 \pm \sqrt{-4})/2$. The values for λ are obviously complex. Since the eigenvalues for A are all complex, so are the eigenvectors. For this linear map, there is no orthonormal basis of real eigenvectors. Even a small difference from the conditions of the theorem, just one asymmetrically placed negative sign, invalidates the whole theorem.

5 Examples

Let's go through an example of finding a matrix's orthonormal basis, which we know must exist because of the Spectral Theorem.

Let A be the 3×3 matrix $\begin{bmatrix} 1 & 2 & 2 \\ 2 & 4 & 1 \\ 2 & 4 & 1 \end{bmatrix}$. The first step to finding an orthonormal basis of eigenvectors for A is to find the eigenvectors of A, which entails finding the zeros of the characteristic polynomial of A. By definition two, we know that the characteristic polynomial of a square matrix A is defined as f(t) = det(tI - A). In this case,

$$f(t) = det\left(\begin{bmatrix} t-1 & -2 & -2 \\ -2 & t-1 & -4 \\ -2 & -4 & t-1 \end{bmatrix}\right)$$

= $(t-1)^3 + (-2)(-4)(-2) + (-2)(-2)(-4) - (t-1)(-4)(-4) - (-2)(-2)(t-1) - (-2)(t-1)(-2) =$

$$(t^3 - 3t^2 + 3t - 1) - 32 - 16t + 16 - 8t + 8 = t^3 - 3t^2 - 21t - 10.$$

The solutions of this cubic equation are the eigenvalues of the matrix A. There are three solutions for this particular equation and they are: $t = \lambda = -3$, $t = \lambda' = 6.4641$, and $t = \lambda'' = -.4641$. From these eigenvalues, we can start figuring out the eigenvectors of A using definition 3. Let $v = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ be an eigenvector of A with the eigenvalue of -3. This means that $\begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & 4 \\ 2 & 4 & 1 \end{bmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -3x \\ -3y \\ -3z \end{pmatrix} = \begin{pmatrix} x+2y+2z \\ 2x+y+2z \\ 2x+4y+z \end{pmatrix}$.

This means that:

$$x + 2y + 2z = -3x$$

$$2x + y + 4z = -3y$$

$$2x + 4y + z = -3z$$

If we let x = 0, we find that -2y = 2z, y + 4z = -3y, and 4y + z = -3z. Thus, y = -z. We can now say that a normal eigenvector of the A with eigenvalue -3 is $< 0, -\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2} >$.

Now, for the eigenvalue λ' :

$$x + 2y + 2z = 6.4641x$$

$$2x + y + 4z = 6.4641y$$

$$2x + 4y + z = 6.4641z$$

By subtracting the second two equations, we can clearly see that -3y + 3z = 6.4641y - 6.4641z, and therefore $9.4641y = 9.4641z \implies y = z$. Substituting one variable for the other in the first equation, we find that x + 4y = 6.4641x, and that $y = z = \frac{5.4641x}{4}$. Therefore, and eigenvector for the eigenvalue of 6.4641 is < 0.732, 1, 1, >.

To determine whether this eigenvector is normal to the first one we found, we take their interproduct. $< 0, -\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2} > < 0.732, 1, 1, >= 0 - \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2} = 0$. The two vectors are indeed orthogonal. Now, to normalize the second eigenvector, we divide it by its length and get < 0.4597, 0.6279, 0.6280 >.

The third eigenvalue:

$$x + 2y + 2z = -.4641x$$

$$2x + y + 4z = -.4641y$$

$$2x + 4y + z = -.4641z$$

It is pretty clear that in this case, y = z as well. Using the first equation, we find that $y = z = \frac{-1.4641x}{4}$. A corresponding eigenvector is $\langle -2.732, 1, 1 \rangle$.

It is obviously orthogonal to the original eigenvector:

$$<0, -\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}><-2.732, 1, 1, >=0 - \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2} = 0$$

It is also orthogonal to the second eigenvector:

$$< 0.732, 1, 1 > < -2.732, 1, 1 > = -2 + 1 + 1$$

When we normalize the third eigenvector, we get < -0.888, 0.325, 0.325 >.

These three vectors $(< 0, -\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2} >, < 0.4597, 0.6279, 0.6280 >, and < -0.888, 0.325, 0.325 >)$ are all orthogonal to each other, and therefore linearly independent. Since they span the all three dimensions that the linear transformation A does, they form a basis for R^3 . Since they are normalized eigenvectors, they form the orthonormal basis of eigenvectors for A that the Spectral Theorem predicated to exist.

Acknowledgments

I would like to thank my advisor, Mike Miller, for putting up with and correcting my stream of drafts.