Eigenbases

Let A be a 3×3 matrix. Suppose that there are vectors $\vec{u}$, $\vec{v}$, $\vec{w}$ with $A\vec{u} = 3\vec{u}$, $A\vec{v} = 5\vec{v}$, $A\vec{w} = 7\vec{w}$. Recall that $\vec{u}$, $\vec{v}$, $\vec{w}$ must be linearly independent, so they are a basis.

Wake up question: What is the matrix of A in the coordinates of the basis $\vec{u}$, $\vec{v}$, $\vec{w}$?

Let A be a 3×3 matrix. Suppose that there are vectors $\vec{u}$, $\vec{v}$, $\vec{w}$ with $A\vec{u} = 3\vec{u}$, $A\vec{v} = 5\vec{v}$, $A\vec{w} = 7\vec{w}$. Recall that $\vec{u}$, $\vec{v}$, $\vec{w}$ must be linearly independent, so they are a basis.

Wake up question: What is the matrix of A in the coordinates of the basis $\vec{u}$, $\vec{v}$, $\vec{w}$?

$$\begin{bmatrix} 3 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 7 \end{bmatrix}.$$

Let A be a 3×3 matrix. Suppose that there are vectors $\vec{u}$, $\vec{v}$, $\vec{w}$ with $A\vec{u} = 3\vec{u}$, $A\vec{v} = 5\vec{v}$, $A\vec{w} = 7\vec{w}$. Recall that $\vec{u}$, $\vec{v}$, $\vec{w}$ must be linearly independent, so they are a basis.

Wake up question: What is the matrix of A in the coordinates of the basis $\vec{u}$, $\vec{v}$, $\vec{w}$?

$$\begin{bmatrix} 3 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 7 \end{bmatrix}.$$

If S is the matrix with columns $\vec{u}$, $\vec{v}$, $\vec{w}$, then we have

$$A = S \left[\begin{smallmatrix} 3 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 7 \end{smallmatrix}\right] S^{-1}.$$

Let A be a 3×3 matrix. Suppose that there are vectors $\vec{u}$, $\vec{v}$, $\vec{w}$ with $A\vec{u} = 3\vec{u}$, $A\vec{v} = 5\vec{v}$, $A\vec{w} = 7\vec{w}$. Recall that $\vec{u}$, $\vec{v}$, $\vec{w}$ must be linearly independent, so they are a basis.

Wake up question: What is the matrix of A in the coordinates of the basis $\vec{u}$, $\vec{v}$, $\vec{w}$?

$$\begin{bmatrix} 3 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 7 \end{bmatrix}.$$

If S is the matrix with columns $\vec{u}$, $\vec{v}$, $\vec{w}$, then we have

$$A = S \left[\begin{smallmatrix} 3 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 7 \end{smallmatrix}\right] S^{-1}.$$

We call $\vec{u}$, $\vec{v}$, $\vec{w}$ an ***eigenbasis*** of A. The general definition is that, if $A : V \to V$ is a linear map, then $\vec{v}_1$, $\vec{v}_2$, $\ldots$, $\vec{v}_n$ is an ***eigenbasis of*** A if it is a basis of V and the $\vec{v}_i$ are eigenvectors of A. If A has an eigenbasis, we say that A is ***diagonalizable***.

The general definition is that, if $A : V \to V$ is a linear map, then $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ is an ***eigenbasis of*** A if it is a basis of V and the $\vec{v}_i$ are eigenvectors of A. If A has an eigenbasis, we say that A is ***diagonalizable***.

Remark: If A has n distinct eigenvalues, then it must have an eigenbasis, because the eigenvectors must be linearly independent. We'll come back to this.

The general definition is that, if $A : V \to V$ is a linear map, then $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ is an ***eigenbasis of*** A if it is a basis of V and the $\vec{v}_i$ are eigenvectors of A. If A has an eigenbasis, we say that A is ***diagonalizable***.

If $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ is an eigenbasis of A, with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ then the coordinates of A in the basis $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ is

$$\begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}.$$

In regular coordinates, we have

$$A = \begin{bmatrix} | & | & | & & | \\ | & | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \vec{v}_3 & \cdots & \vec{v}_n \\ | & | & | & & | \\ | & | & | & & | \end{bmatrix} \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} \begin{bmatrix} | & | & | & & | \\ | & | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \vec{v}_3 & \cdots & \vec{v}_n \\ | & | & | & & | \\ | & | & | & & | \end{bmatrix}^{-1}.$$

Exercise: Let A be a 2×2 matrix which has an eigenbasis with eigenvalues 3 and 5. Show that $(A - 3\mathrm{Id})(A - 5\mathrm{Id}) = 0$.

Exercise: Let A be a 2×2 matrix which has an eigenbasis with eigenvalues 3 and 5. Show that $(A - 3\mathrm{Id})(A - 5\mathrm{Id}) = 0$.

Proof: Let $\vec{u}$, $\vec{v}$ be the eigenbasis. Let $\vec{z}$ be any vector, and write $\vec{z} = a\vec{u} + b\vec{v}$. Then **Proof:** Let $\vec{u}$, $\vec{v}$ be the eigenbasis. Let $\vec{z}$ be any vector, and write $\vec{z} = a\vec{u} + b\vec{v}$. Then

$$(A - 3\mathrm{Id})(A - 5\mathrm{Id})\vec{z} = (A - 3\mathrm{Id})(A - 5\mathrm{Id})(a\vec{u} + b\vec{v}) =$$

$$(A - 3\mathrm{Id})(a(3-5)\vec{u} + b(5-5)\vec{v}) = (A - 3\mathrm{Id})a(3-5)\vec{u} = a(3-3)(3-5)\vec{u} = 0.$$

Exercise: Let A be a 2×2 matrix which has an eigenbasis with eigenvalues 3 and 5. Show that $(A - 3\mathrm{Id})(A - 5\mathrm{Id}) = 0$.

Proof: Let $\vec{u}$, $\vec{v}$ be the eigenbasis. Let $\vec{z}$ be any vector, and write $\vec{z} = a\vec{u} + b\vec{v}$. Then

$$(A - 3\mathrm{Id})(A - 5\mathrm{Id})\vec{z} = (A - 3\mathrm{Id})(A - 5\mathrm{Id})(a\vec{u} + b\vec{v}) =$$

$$(A-3\mathrm{Id})(a(3-5)\vec{u}+b(5-5)\vec{v}) = (A-3\mathrm{Id})a(3-5)\vec{u} = a(3-3)(3-5)\vec{u} = 0.$$

Question: Suppose that A were, instead, a 100×100 matrix with an eigenbasis made up of 50 eigenvectors with eigenvalue 3 and 50 eigenvectors with eigenvalue 5. Would we still have $(A - 3\mathrm{Id})(A - 5\mathrm{Id}) = 0$?

Exercise: Let A be a 2×2 matrix which has an eigenbasis with eigenvalues 3 and 5. Show that $(A - 3\mathrm{Id})(A - 5\mathrm{Id}) = 0$.

Proof: Let $\vec{u}$, $\vec{v}$ be the eigenbasis. Let $\vec{z}$ be any vector, and write $\vec{z} = a\vec{u} + b\vec{v}$. Then

$$(A - 3\mathrm{Id})(A - 5\mathrm{Id})\vec{z} = (A - 3\mathrm{Id})(A - 5\mathrm{Id})(a\vec{u} + b\vec{v}) =$$

$$(A-3\mathrm{Id})(a(3-5)\vec{u}+b(5-5)\vec{v}) = (A-3\mathrm{Id})a(3-5)\vec{u} = a(3-3)(3-5)\vec{u} = 0.$$

Question: Suppose that A were, instead, a 100×100 matrix with an eigenbasis made up of 50 eigenvectors with eigenvalue 3 and 50 eigenvectors with eigenvalue 5. Would we still have $(A - 3\mathrm{Id})(A - 5\mathrm{Id}) = 0$?

Yes! Just need to check that $(A - 3\mathrm{Id})(A - 5\mathrm{Id})\vec{v} = 0$ for each basis vector $\vec{v}$. If $\vec{v}$ is a 3-eigenvector, then $(A - 3\mathrm{Id})(A - 5\mathrm{Id})\vec{v} = (3 - 3)(3 - 5)\vec{v} = \vec{0}$ and, if $\vec{v}$ is a 5-eigenvector, then $(A - 3\mathrm{Id})(A - 5\mathrm{Id})\vec{v} = (5 - 3)(5 - 5)\vec{v} = \vec{0}$.

In general, if A has an eigenbasis with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_r$, then $\prod(A - \lambda_j \mathrm{Id}) = 0$ where we just include each eigenvalue once.

The formula $\prod(A - \lambda_j \mathrm{Id})$ should remind us of the characteristic polynomial which we met last week. Recall that the ***characteristic polynomial*** is

$$\chi_A(x) = \det(x\mathrm{Id} - A).$$

The roots of the characteristic polynomial are the eigenvalues of A.

The formula $\prod(A - \lambda_j \mathrm{Id})$ should remind us of the characteristic polynomial which we met last week. Recall that the ***characteristic polynomial*** is

$$\chi_A(x) = \det(x\mathrm{Id} - A).$$

The roots of the characteristic polynomial are the eigenvalues of A.

Exercise: What is the characteristic polynomial of

$$\begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix}?$$

The formula $\prod(A - \lambda_j \mathrm{Id})$ should remind us of the characteristic polynomial which we met last week. Recall that the ***characteristic polynomial*** is

$$\chi_A(x) = \det(x\mathrm{Id} - A).$$

The roots of the characteristic polynomial are the eigenvalues of A.

Exercise: What is the characteristic polynomial of

$$\begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix}?$$

$$\det \begin{bmatrix} x-3 & 0 & 0 \\ 0 & x-3 & 0 \\ 0 & 0 & x-5 \end{bmatrix} = (x-3)(x-3)(x-5).$$

The formula $\prod(A - \lambda_j \mathrm{Id})$ should remind us of the characteristic polynomial which we met last week. Recall that the ***characteristic polynomial*** is

$$\chi_A(x) = \det(x\mathrm{Id} - A).$$

The roots of the characteristic polynomial are the eigenvalues of A.

Exercise: What is the characteristic polynomial of

$$\begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix}?$$

$$\det \begin{bmatrix} x-3 & 0 & 0 \\ 0 & x-3 & 0 \\ 0 & 0 & x-5 \end{bmatrix} = (x-3)(x-3)(x-5).$$

In general, if A has an eigenbasis with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, then $\chi_A(x) = \prod_{j=1}^{n}(x - \lambda_j)$, with multiple eigenvalues used multiple times.

In short, if A is diagonalizable, then $\chi_A(x) = \prod_{j=1}^{n}(x - \lambda_j)$ with multiple eigenvalues used multiple times. In this case, we have $\prod(A - \lambda_i \mathrm{Id}) = 0$ even just using each eigenvalue once (and, of course, also if we use them more than once).

Minimal polynomials

We now want to think more generally about what we can say about matrices obeying polynomials. For example, let

$$J = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

Geometrically, J is a 90° rotation. It has no nonzero real eigenvectors.

We now want to think more generally about what we can say about matrices obeying polynomials. For example, let

$$J = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

Geometrically, J is a 90° rotation. It has no nonzero real eigenvectors.

We have $J^2 = -\mathrm{Id}$. So J obeys the polynomial $x^2 + 1 = 0$.

We now want to think more generally about what we can say about matrices obeying polynomials. For example, let

$$J = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

Geometrically, J is a 90° rotation. It has no nonzero real eigenvectors.

We have $J^2 = -\mathrm{Id}$. So J obeys the polynomial $x^2 + 1 = 0$.

We also have $J^3 = -J$ and $J^4 = \mathrm{Id}$, so J also obeys the polynomials $x^3 + x$ and $x^4 - 1$.

We now want to think more generally about what we can say about matrices obeying polynomials. For example, let

$$J = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

Geometrically, J is a $90°$ rotation. It has no nonzero real eigenvectors.

We have $J^2 = -\mathrm{Id}$. So J obeys the polynomial $x^2 + 1 = 0$.

We also have $J^3 = -J$ and $J^4 = \mathrm{Id}$, so J also obeys the polynomials $x^3 + x$ and $x^4 - 1$. Notice that $x^3 + x = x(x^2 + 1)$ and $x^4 - 1 = (x^2 - 1)(x^2 + 1)$.

We now want to think more generally about what we can say about matrices obeying polynomials. For example, let

$$J = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

Geometrically, J is a 90° rotation. It has no nonzero real eigenvectors.

We have $J^2 = -\mathrm{Id}$. So J obeys the polynomial $x^2 + 1 = 0$.

We also have $J^3 = -J$ and $J^4 = \mathrm{Id}$, so J also obeys the polynomials $x^3 + x$ and $x^4 - 1$. Notice that $x^3 + x = x(x^2 + 1)$ and $x^4 - 1 = (x^2 - 1)(x^2 + 1)$.

This is part of a general pattern: Let $A : V \to V$ be a linear operator and let $f(x)$ be 0. If $f(A) = 0$ and $f(x)$ divides $g(x)$, then $g(A) = 0$.

Indeed, let $g(x) = h(x)f(x)$. Then $g(A) = h(A)f(A) = h(A)0 = 0$.

The polynomial $x^2 + 1$ is what is called the ***minimal polynomial*** of A.

Theorem/Definition: Let V be a finite dimensional vector space and let $A : V \to V$ be a linear map. Then there is a nonzero polynomial $m(x)$ such that $m(A) = 0$ and, if $f(x)$ is any other polynomial with $f(x) = 0$, then $m(x)$ divides $f(x)$. We can describe $m(x)$ as the polynomial of minimal degree with $m(A) = 0$; the polynomial $m(x)$ is called the ***minimal polynomial*** of A.

We need to prove:

(1) There is a nonzero polynomial $f(x)$ with $f(A) = 0$.

(2) If $m(x)$ is the polynomial of minimal degree with $m(A) = 0$, and $f(x)$ is any other polynomial with $f(A) = 0$, then $m(x)$ divides $f(x)$.

We need to prove:

(1) There is a nonzero polynomial $f(x)$ with $f(A) = 0$.

(2) If $m(x)$ is the polynomial of minimal degree with $m(A) = 0$, and $f(x)$ is any other polynomial with $f(A) = 0$, then $m(x)$ divides $f(x)$.

(1) Think of A as an $n \times n$ matrix. Then the powers A^0, A^1, A^2, $\ldots$, A^{n^2} are $n^2 + 1$ matrices of size $n \times n$, we can think of these as $n^2 + 1$ vectors in an n^2 dimensional space, so there is a linear relationship

$$f_{n^2} A^{n^2} + f_{n^2-1} A^{n^2-1} + \cdots + f_2 A^2 + f_1 A + f_0 \mathrm{Id} = 0.$$

We need to prove:

(1) There is a nonzero polynomial $f(x)$ with $f(A) = 0$.

(2) If $m(x)$ is a polynomial of minimal degree with $m(A) = 0$, and $f(x)$ is any other polynomial with $f(A) = 0$, then $m(x)$ divides $f(x)$.

(2) Preliminary claim: For any polynomials $f(x)$ and $m(x)$, $m(x) \neq 0$, we can write $f(x) = q(x)m(x) + r(x)$ with $\deg r(x) < \deg m(x)$.

Preliminary claim: For any polynomials $f(x)$ and $m(x)$, $m(x) \neq 0$, we can write $f(x) = q(x)m(x) + r(x)$ with $\deg r(x) < \deg m(x)$.

Proof: Subtract off multiples of $m(x)$ from $f(x)$ to write $f(x) = b(x)m(x) + r(x)$ with $\deg r(x) < d$.

For example, let $m(x) = x^2 + x + 1$ and let $f(x) = x^4 + 2x^3 + 4x^2 + 8x + 9$. Then

$$f(x) - x^2m(x) = x^3 + 4x^2 + 8x + 9$$

$$f(x) - x^2m(x) - xm(x) = 3x^2 + 7x + 9$$

$$f(x) - x^2m(x) - xm(x) - 3m(x) = 4x + 6$$

Preliminary claim: For any polynomials $f(x)$ and $m(x)$, $m(x) \neq 0$, we can write $f(x) = q(x)m(x) + r(x)$ with $\deg r(x) < \deg m(x)$.

Proof: Subtract off multiples of $m(x)$ from $f(x)$ to write $f(x) = b(x)m(x) + r(x)$ with $\deg r(x) < d$.

For example, let $m(x) = x^2 + x + 1$ and let $f(x) = x^4 + 2x^3 + 4x^2 + 8x + 9$. Then

$$f(x) - x^2m(x) = x^3 + 3x^2 + 8x + 9$$

$$f(x) - x^2m(x) - xm(x) = 2x^2 + 7x + 9$$

$$f(x) - x^2m(x) - xm(x) - 2m(x) = 5x + 7$$

$f(x) - (x^2 + x + 2)m(x) = 5x + 7$ so $f(x) = (x^2 + x + 3)m(x) + 5x + 7$.

We need to prove:

(1) There is a nonzero polynomial $f(x)$ with $f(A) = 0$.

(2) If $m(x)$ is a polynomial of minimal degree with $m(A) = 0$, and $f(x)$ is any other polynomial with $f(A) = 0$, then $m(x)$ divides $f(x)$.

(2) For any polynomials $f(x)$ and $m(x)$, $m(x) \neq 0$, we can write $f(x) = q(x)m(x) + r(x)$ with $\deg r(x) < \deg m(x)$.

We need to prove:

(1) There is a nonzero polynomial $f(x)$ with $f(A) = 0$.

(2) If $m(x)$ is a polynomial of minimal degree with $m(A) = 0$, and $f(x)$ is any other polynomial with $f(A) = 0$, then $m(x)$ divides $f(x)$.

(2) For any polynomials $f(x)$ and $m(x)$, $m(x) \neq 0$, we can write $f(x) = q(x)m(x) + r(x)$ with $\deg r(x) < \deg m(x)$.

If $m(A) = 0$ and $f(A) = 0$, then we have
$0 = q(A)m(A) + r(A) = q(A) \cdot 0 + r(A) = r(A)$. So, if $m(A) = 0$ and $f(A) = 0$, then $r(A) = 0$ as well. This would make $r(x)$ a lower degree polynomial than $m(x)$ with $r(A) = 0$, contradicting our choice of m ...

We need to prove:

(1) There is a nonzero polynomial $f(x)$ with $f(A) = 0$.

(2) If $m(x)$ is a polynomial of minimal degree with $m(A) = 0$, and $f(x)$ is any other polynomial with $f(A) = 0$, then $m(x)$ divides $f(x)$.

(2) For any polynomials $f(x)$ and $m(x)$, $m(x) \neq 0$, we can write $f(x) = q(x)m(x) + r(x)$ with $\deg r(x) < \deg m(x)$.

If $m(A) = 0$ and $f(A) = 0$, then we have
$0 = q(A)m(A) + r(A) = q(A) \cdot 0 + r(A) = r(A)$. So, if $m(A) = 0$ and $f(A) = 0$, then $r(A) = 0$ as well. This would make $r(x)$ a lower degree polynomial than $m(x)$ with $r(A) = 0$, contradicting our choice of m ... unless $r(x) = 0$!

We need to prove:

(1) There is a nonzero polynomial $f(x)$ with $f(A) = 0$.

(2) If $m(x)$ is a polynomial of minimal degree with $m(A) = 0$, and $f(x)$ is any other polynomial with $f(A) = 0$, then $m(x)$ divides $f(x)$.

(2) For any polynomials $f(x)$ and $m(x)$, $m(x) \neq 0$, we can write $f(x) = q(x)m(x) + r(x)$ with $\deg r(x) < \deg m(x)$.

If $m(A) = 0$ and $f(A) = 0$, then we have $0 = q(A)m(A) + r(A) = q(A) \cdot 0 + r(A) = r(A)$. So, if $m(A) = 0$ and $f(A) = 0$, then $r(A) = 0$ as well. This would make $r(x)$ a lower degree polynomial than $m(x)$ with $r(A) = 0$, contradicting our choice of m ... unless $r(x) = 0$!

So $r(x) = 0$ and $m(x)$ divides $f(x)$. **QED**

One last note: Why do we say that $m(x)$ is ***the*** minimal polynomial? Could we have more than one?

One last note: Why do we say that $m(x)$ is ***the*** minimal polynomial? Could we have more than one?

Let $m_1(x)$ and $m_2(x)$ be two polynomials of minimal degree with $m_1(A) = m_2(A) = 0$. Then $m_1(x)$ divides $m_2(x)$ and $m_2(x)$ divides $m_1(x)$, so m_1 and m_2 are the same up to a scalar multiple.

One last note: Why do we say that $m(x)$ is ***the*** minimal polynomial? Could we have more than one?

Let $m_1(x)$ and $m_2(x)$ be two polynomials of minimal degree with $m_1(A) = m_2(A) = 0$. Then $m_1(x)$ divides $m_2(x)$ and $m_2(x)$ divides $m_1(x)$, so m_1 and m_2 are the same up to a scalar multiple.

So all minimal polynomials are the same up to a scalar multiple; we will usually adopt the normalization of taking the highest degree term of $m(x)$ to have leading degree 1 and call $m(x)$ ***the*** minimal polynomial of A.