MATHEMATICS 23a/E-23a, Fall 2018
Linear Algebra and Real Analysis I
Week 4 (Eigenvectors and Eigenvalues)

Author: Paul Bamberg
R scripts by Paul Bamberg
Last modified: August 11, 2018 by Paul Bamberg

**Reading**

- Hubbard, Section 2.7
- Hubbard, pages 474-475

**Recorded Lectures**

- Lecture 8 (Week 4, Class 1) (watch on October 2 or 3)
- Lecture 9 (Week 4, Class 2) (watch on October 4 or 5)

**Proofs to present in section or to a classmate who has done them.**

- 4.1 Prove that if $\vec{\mathbf{v}}_1, \cdots, \vec{\mathbf{v}}_n$ are eigenvectors of $A : \mathbb{R}^n \to \mathbb{R}^n$ with distinct eigenvalues $\lambda_1 \cdots \lambda_n$, they are linearly independent. Conclude that an $n \times n$ matrix cannot have more than $n$ distinct eigenvalues.
- 4.2
  - For real $n \times n$ matrix $A$, prove that if all the polynomials $p_i(t)$ are simple and have real roots, then there exists a basis for $\mathbb{R}^n$ consisting of eigenvectors of $A$.
  - Prove that if there exists a basis for $\mathbb{R}^n$ consisting of eigenvectors of $A$, then all the polynomials $p_i(t)$ are simple and have real roots.

  Note - Theorem 2.7.9 in Hubbard is more powerful, because it applies to the complex case. The proof is the same. Our proof is restricted to the real case only because we are not doing examples with complex eigenvectors.

**R Scripts**

- 1.4A-EigenvaluesCharacteristic.R
  Topic 1 - Eigenvectors for a 2x2 matrix
  Topic 2 - Not every 2x2 matrix has real eigenvalues
- 1.4B-EigenvectorsAxler.R
  Topic 1 - Finding eigenvectors by row reduction
  Topic 2 - Eigenvectors for a 3 x 3 matrix
- 1.4C-Diagonalization.R
  Topic 1: Basis of real eigenvectors
  Topic 2 - Raising a matrix to a power
  Topic 3 - Wnat if the eigenvalues are complex?
  Topic 4 - What if there is no eigenbasis?
- 1.4X-EigenvectorApplications.R
  Topic 1 - The special case of a symmetric matrix
  Topic 2 - Markov Process (from script 1.1D)
  Topic 3 - Eigenvectors for a reflection
  Topic 4 - Sequences defined by linear recurrences

# 1 Executive Summary

## 1.1 Eigenvalues and eigenvectors

If $A\vec{\mathbf{v}} = \lambda\vec{\mathbf{v}}$, $\vec{\mathbf{v}}$ is called an eigenvector for $A$, and $\lambda$ is the corresponding eigenvalue.

For example, if $A = \begin{bmatrix} -1 & 4 \\ -2 & 5 \end{bmatrix}$, we can check that $\vec{\mathbf{v}} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is an eigenvector of $A$ with eigenvalue 3.

If $A$ is a $2 \times 2$ or $3 \times 3$ matrix, there is a quick, well-known way to find eigenvalues by using determinants.

Rewrite $A\vec{\mathbf{v}} = \lambda\vec{\mathbf{v}}$ as $A\vec{\mathbf{v}} = \lambda I\vec{\mathbf{v}}$, where $I$ is the identity matrix.

Equivalently, $(A - \lambda I)\vec{\mathbf{v}} = \vec{\mathbf{0}}$

Suppose that $\lambda$ is an eigenvalue of $A$. Then the eigenvector $\vec{\mathbf{v}}$ is a nonzero vector in the kernel of the matrix $(A - \lambda I)$.

It follows that the matrix $(A - \lambda I)$ is not invertible. But we have a formula for the inverse of a $2 \times 2$ or $3 \times 3$ matrix, which can fail only if the determinant is zero. Therefore a necessary condition for the existence of an eigenvalue is that $\det(A - \lambda I) = 0$.

The polynomial $\chi_A(\lambda) = \det(A - \lambda I)$ is called the **characteristic polynomial** of matrix $A$. It is easy to compute in the $2 \times 2$ or $3 \times 3$ case, where there is a simple formula for the determinant. For larger matrices $\chi_A(\lambda)$ is hard to compute efficiently, and this approach should be avoided.

Conversely, suppose that $\chi_A(\lambda) = 0$ for some real number $\lambda$. It follows that the columns of the matrix $(A - \lambda I)$ are linearly dependent. If we row reduce the matrix, we will find at least one nonpivotal column, which in turn implies that there is a nonzero vector in the kernel. This vector is an eigenvector.

This was the standard way of finding eigenvectors until 1995, but it has two drawbacks:

- It requires computation of the determinant of a matrix whose entries are polynomials. Efficient algorithms for calculating the determinant of large square matrices use row-reduction techniques, which might require division by a pivotal element that is a polynomial in $\lambda$.

- Once you have found the eigenvalues, finding the corresponding eigenvectors is a nontrivial linear algebra problem.

### 1.2 Finding eigenvalues - a simple example

Let $A = \begin{bmatrix} -1 & 4 \\ -2 & 5 \end{bmatrix}$. Then $A - \lambda I = \begin{bmatrix} -1-\lambda & 4 \\ -2 & 5-\lambda \end{bmatrix}$
and $\chi_A(\lambda) = \det(A - \lambda I) = (-1-\lambda)(5-\lambda) + 8 = \lambda^2 - 4\lambda + 3$.
Setting $\lambda^2 - 4\lambda + 3 = (\lambda - 1)(\lambda - 3) = 0$, we find two eigenvalues, 1 and 3.
Finding the corresponding eigenvectors still requires a bit of algebra.
For $\lambda = 1, A - \lambda I = \begin{bmatrix} -2 & 4 \\ -2 & 4 \end{bmatrix}$.
By inspection we see that $\vec{\mathbf{v}}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ is in the kernel of this matrix.
Check: $A\vec{\mathbf{v}}_1 = \begin{bmatrix} -1 & 4 \\ -2 & 5 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ – eigenvector with eigenvalue 1.
For $\lambda = 3, A - \lambda I = \begin{bmatrix} -4 & 4 \\ -2 & 2 \end{bmatrix}$, and $\vec{\mathbf{v}}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is in the kernel.
Check: $A\vec{\mathbf{v}}_2 = \begin{bmatrix} -1 & 4 \\ -2 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$ – eigenvector with eigenvalue 3.

## 1.3 A better way to find eigenvectors

Given matrix $A$, pick an arbitrary vector $\vec{\mathbf{w}}$. Keep computing $A\vec{\mathbf{w}}$ , $A^2\vec{\mathbf{w}}$, $A^3\vec{\mathbf{w}}$, etc. until you find a vector that is a linear combination of its predecessors. This situation is easily detected by row reduction.
Now you have found a polynomial $p$ of degree $m$ such that $p(A)\vec{\mathbf{w}} = 0$. Furthermore, this is the nonzero polynomial of lowest degree for which $p(A)\vec{\mathbf{w}} = 0$.

Over the complex numbers, this polynomial is guaranteed to have a root $\lambda$ by virtue of the "fundamental theorem of algebra" (Hubbard theorem 1.6.13). Over the real numbers or a finite field, it it will have a root in the field only if you are lucky. Assuming that the root exists, factor it out: $p(t) = (t - \lambda)q(t)$.

Now $p(A)\vec{\mathbf{w}} = (A - \lambda I)q(A)\vec{\mathbf{w}} = 0$.

Thus $q(A)\vec{\mathbf{w}}$ is an eigenvector with eigenvalue $\lambda$.

Again, let $A = \begin{bmatrix} -1 & 4 \\ -2 & 5 \end{bmatrix}$

As the arbitrary vector $\vec{\mathbf{w}}$ choose $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Then $A\vec{\mathbf{w}} = \begin{bmatrix} -1 \\ -2 \end{bmatrix}$ and $A^2\vec{\mathbf{w}} = \begin{bmatrix} -7 \\ -8 \end{bmatrix}$.

We need to express the third of these vectors, $A^2\vec{\mathbf{w}}$, as a linear combination of the first two. This is done by row reducing the matrix
$\begin{bmatrix} 1 & -1 & -7 \\ 0 & -2 & -8 \end{bmatrix}$ to $\begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 4 \end{bmatrix}$ to find that $A^2\vec{\mathbf{w}} = 4A\vec{\mathbf{w}} - 3I\vec{\mathbf{w}}$.

Equivalently, $(A^2 - 4A + 3I)\vec{\mathbf{w}} = 0$.
$p(A) = A^2 - 4A + 3I$ or $p(t) = t^2 - 4t + 3 = (t - 1)(t - 3)$: eigenvalues 1 and 3.

To get the eigenvector for eigenvalue 1, apply the remaining factor of $p(A)$, $A - 3I$, to $\vec{\mathbf{w}}$: $\begin{bmatrix} -4 & 4 \\ -2 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -4 \\ -2 \end{bmatrix}$. Divide by -2 to get $\vec{\mathbf{v}}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$.

To get the eigenvector for eigenvalue 3, apply the remaining factor of $p(A)$, $A - I$, to $\vec{\mathbf{w}}$: $\begin{bmatrix} -2 & 4 \\ -2 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -2 \\ -2 \end{bmatrix}$. Divide by -2 to get $\vec{\mathbf{v}}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

In this case the polynomial $p(t)$ turned out to be the same as the characteristic polynomial, but that is not always the case.

- If we choose $\vec{\mathbf{w}} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, we find $A\vec{\mathbf{w}} = 3\vec{\mathbf{w}}, p(A) = A - 3I, p(t) = t - 3$. We need to start over with a different $\vec{\mathbf{w}}$ to find the other eigenvalue.

- If we choose $A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$, then any vector is an eigenvector with eigenvalue 2. So $p(t) = t - 2$. But the characteristic polynomial is $(t - 2)^2$.

- If we choose $A = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$, the characteristic polynomial is $(t-2)^2$. But now there is only one eigenvector. If we choose $\vec{\mathbf{w}} = \vec{\mathbf{e}}_1$ we find $p(t) = t - 2$ and the eigenvector $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. But if we choose a different $\vec{\mathbf{w}} = \vec{\mathbf{e}}_2$ we find $p(t) = (t - 2)^2$ and we fail to find a second, independent eigenvector.

## 1.4 When is there an eigenbasis?

Choose $\vec{\mathbf{w}}$ successively to equal $\vec{\mathbf{e}}_1\vec{\mathbf{e}}_2, \cdots, \vec{\mathbf{e}}_n$.

In searching for eigenvectors, we find successively polynomials $p_1(t), p_2(t), \cdots, p_n(t)$.

There is a basis of real eigenvectors if and only if each of the polynomials $p_i(t)$ has **simple real roots**, e.g. $p(t) = t(t-2)(t+4)(t-2.3)$. No repeated factors are allowed!

A polynomial like $p(t) = t^2 + 1$, although it has no repeated factors, has no real roots: $p(t) = (t+i)(t-i)$.

If we allow complex roots, then any polynomial can be factored into linear factors (Fundamental Theorem of Algebra, Hubbard page113).

There is a basis of complex eigenvectors if and only if each of the polynomials $p_i(t)$ has **simple roots**, e.g. $p(t) = t(t-i)(t+i)$. No repeated factors are allowed!

Our technique for finding eigenvectors works also for matrices over finite fields, but in that case it is entirely possible for a polynomial to have no linear factors whatever. In that case there are no eigenvectors and no eigenbasis. This is one of the few cases where linear algebra over a finite field is fundamentally different from linear algebra over the real or complex numbers.

## 1.5 Matrix Diagonalization

In the best case we can find a basis of $n$ eigenvectors $\{\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \cdots, \vec{\mathbf{v}}_n\}$ with associated eigenvalues $\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$. Although the eigenvectors must be independent, some of the eigenvalues may repeat.

Create a matrix $P$ whose columns are the eigenvectors. Since the eigenvectors form a basis, they are independent and the matrix $P$ has an inverse $P^{-1}$

The matrix $D = P^{-1}AP$ is a diagonal matrix.

Proof: $D\vec{\mathbf{e}}_k = P^{-1}A(P\vec{\mathbf{e}}_k) = P^{-1}A\vec{\mathbf{v}}_k = P^{-1}\lambda_k\vec{\mathbf{v}}_k = \lambda_k P^{-1}\vec{\mathbf{v}}_k = \lambda_k\vec{\mathbf{e}}_k$.

The matrix $A$ can be expressed as $A = PDP^{-1}$.

Proof: $A\vec{\mathbf{v}}_k = PD(P^{-1}\vec{\mathbf{v}}_k) = PD\vec{\mathbf{e}}_k = P(\lambda_k\vec{\mathbf{e}}_k) = \lambda_k P\vec{\mathbf{e}}_k = \lambda_k\vec{\mathbf{v}}_k$.

A diagonal matrix $D$ is easy to raise to an integer power.

For example, if $D = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$, then $D^k = \begin{bmatrix} \lambda_1^k & 0 & 0 \\ 0 & \lambda_2^k & 0 \\ 0 & 0 & \lambda_3^k \end{bmatrix}$

But now $A = PDP^{-1}$ is also easy to rasie to a power, because $A^k = PD^kP^{-1}$ (will be proved by induction)

The same result extends to $k$th roots of matrices,
where $B = A^{1/k}$ means that $B^k = A$.

## 1.6 Properties of an eigenbasis

- Even if all the eigenvalues are distinct, an eigenbasis is not unique. Any eigenvector in the basis can be multiplied by a nonzero scalar and remain an eigenvector.

- Eigenvectors that correspond to distinct eigenvalues are linearly independent (your proof 4.1)

- If the matrix $A$ is symmetric, eigenvectors that correspond to distinct eigenvalues are orthogonal.

## 1.7 What if there is no eigenbasis?

We consider only the case where $A$ is a $2 \times 2$ matrix. If a real polynomial $p(t)$ does not have two distinct real roots, then it either has a repeated real root or it has a pair of conjugate complex roots.

Case 1: Repeated root: $p(t) = (t - \lambda)^2$.

So $p(A) = (A - \lambda I)^2 = 0$.

Set $N = A - \lambda I$, and $N^2 = 0$. The matrix $N$ is called **nilpotent**.

Now $A = \lambda I + N$, and $A^2 = (\lambda I + N)^2 = \lambda^2 I + 2\lambda N$.

It is easy to prove by induction that $A^k = (\lambda I + N)^k = \lambda^k I + k\lambda^{k-1} N$.

Case 2: Conjugate complex roots:

If a $2 \times 2$ real matrix $A$ has eigenvalues $a \pm ib$, then it can be expressed in the form $A = PCP^{-1}$, where $C$ is the conformal matrix $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ and $P$ is a change of basis matrix. Since a conformal matrix is almost as easy as a diagonal matrix to raise to the $n$th power by virtue of De Moivre's theorem $(r(\cos\theta + i\sin\theta))^n = r^n(\cos n\theta + i \sin n\theta)$, this representation is often useful.

Here is an algorithm for constructing the matrices $C$ and $P$:

Suppose that the eigenvalues of $A$ are $a \pm ib$. Then $A$ has no real eigenvectors, and for any real $\vec{\mathbf{w}}$ we will find the polynomial

$p(t) = (t - a - ib)(t - a + ib) = (t - a)^2 + b^2$

So $p(A) = (A - aI)^2 + b^2 I = 0$ or $(\frac{A-aI}{b})^2 = -I$.

Now we need to construct a new basis, which will not be a basis of eigenvectors but which will still be useful.

Set $\vec{\mathbf{v}}_1 = \vec{\mathbf{e}}_1, \vec{\mathbf{v}}_2 = (\frac{A-aI}{b})\vec{\mathbf{e}}_1$.

Then $(A - aI)\vec{\mathbf{v}}_1 = b\vec{\mathbf{v}}_2$ and $A\vec{\mathbf{v}}_1 = a\vec{\mathbf{v}}_1 + b\vec{\mathbf{v}}_2$ .

Also, $(\frac{A-aI}{b})\vec{\mathbf{v}}_2 = (\frac{A-aI}{b})^2\vec{\mathbf{v}}_1 = -\vec{\mathbf{v}}_1$, so

$(A - aI)\vec{\mathbf{v}}_2 = -b\vec{\mathbf{v}}_1$ and $A\vec{\mathbf{v}}_2 = a\vec{\mathbf{v}}_2 - b\vec{\mathbf{v}}_1$.

With respect to the new basis, the matrix that represents $A$ is the conformal matrix $C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$.

If we define $P$ in the usual way with columns $\vec{\mathbf{v}}_1$ and $\vec{\mathbf{v}}_2$, then $A = PCP^{-1}$, and the matrices $P$ and $C$ are real.

## 1.8 Applications of eigenvectors

- Markov processes

  Suppose that a system can be in one of two or more states and goes through a number of steps, in each of which it may make a transition from one state to another in accordance with specified "transition probabilities."

  For a two-state process, vector $\vec{\mathbf{v}}_n = \begin{bmatrix} p_n \\ q_n \end{bmatrix}$ specifies the probabilities for the system to be in state 1 or state 2 after $n$ steps of the process, where $0 \le p_n, q_n \le 1.$ and $p_n + q_n = 1$ The transition probabilities are spacified by a matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, where all the entries are between 0 and 1 and $a + c = b + d = 1$.

  After a large number of steps, the state of the system is speciifed by $\vec{\mathbf{v}}_n = A^n\vec{\mathbf{v}}_0$.

  The easy way to calculate $A^n$ is by diagonalizing $A$. If there is a "stationary state" $\vec{\mathbf{v}}$ into which the system settles down, it corresponds to an eigenvector with eigenvalue 1, since $\vec{\mathbf{v}}_{n+1} = A\vec{\mathbf{v}}_n$ and $\vec{\mathbf{v}}_{n+1} = \vec{\mathbf{v}}_n = \vec{\mathbf{v}}$.

- Reflections

  If $2 \times 2$ matrix $F$ represents reflection in a line through the origin with direction vector $\vec{\mathbf{v}}$, then $\vec{\mathbf{v}}$ must be an eigenvector with eigenvalue 1 and a vector perpendicular to $\vec{\mathbf{v}}$ must be an eigenvector with eigenvalue -1.

  If $3 \times 3$ matrix $F$ represents reflection in a plane $P$ through the origin with normal vector $\vec{\mathbf{N}}$, then $\vec{\mathbf{N}}$ must be an eigenvector with eigenvalue -1 and there must be a two-dimensional subspace of vectors in $P$, all with eigenvalue +1.

- Linear recurrences and Fibonacci-like sequences.

  In computer science, it is frequently the case that the first two terms of a sequence, $a_0$ and $a_1$, are specified, and subsequent terms are specified by a "linear recurrence" of the form $a_{n+1} = ba_{n-1} + ca_n$. The best-known example is the Fibonacci sequence (Hubbard, pages 220-221) where $a_0 = a_1 = 1$ and $b = c = 1$.

  Then $\begin{bmatrix} a_n \\ a_{n+1} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ b & c \end{bmatrix} \begin{bmatrix} a_{n-1} \\ a_n \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ b & c \end{bmatrix}^n \begin{bmatrix} a_0 \\ a_1 \end{bmatrix}$.

  The easy way to raise matrix $A = \begin{bmatrix} 0 & 1 \\ b & c \end{bmatrix}$ to the $n$th power is to diagonalize it.

- Solving systems of linear differential equations

  This topic, of crucial importance to physics, will be covered after we have done some calculus and infinite series.

## 2 Lecture Outline

1. Using the characteristic polynomial to find eigenvalues and eigenvectors

   If $A\vec{\mathbf{v}} = \lambda\vec{\mathbf{v}}$, $\vec{\mathbf{v}}$ is called an eigenvector for $A$, and $\lambda$ is the corresponding eigenvalue.

   If $A$ is a $2 \times 2$ or $3 \times 3$ matrix, there is a quick, well-known way to find eigenvalues by using determinants.

   Rewrite $A\vec{\mathbf{v}} = \lambda\vec{\mathbf{v}}$ as $A\vec{\mathbf{v}} = \lambda I\vec{\mathbf{v}}$, where $I$ is the identity matrix.

   Equivalently, $(A - \lambda I)\vec{\mathbf{v}} = \vec{\mathbf{0}}$

   Suppose that $\lambda$ is an eigenvalue of $A$. Then the eigenvector $\vec{\mathbf{v}}$ is a nonzero vector in the kernel of the matrix $(A - \lambda I)$.

   It follows that the matrix $(A - \lambda I)$ is not invertible. But we have a formula for the inverse of a $2\times2$ or $3\times3$ matrix, which can fail only if the determinant is zero. Therefore a necessary condition for the existence of an eigenvalue is that $\det(A - \lambda I) = 0$.

   The polynomial $\chi_A(\lambda) = \det(A - \lambda I)$ is called the **characteristic polynomial** of matrix $A$. It is easy to compute in the $2 \times 2$ or $3 \times 3$ case, where there is a simple formula for the determinant. For larger matrices $\chi_A(\lambda)$ is hard to compute efficiently, and this approach should be avoided.

   Conversely, suppose that $\chi_A(\lambda) = 0$ for some real number $\lambda$. It follows that the columns of the matrix $(A - \lambda I)$ are linearly dependent. If we row reduce the matrix, we will find at least one nonpivotal column, which in turn implies that there is a nonzero vector in the kernel. This vector is an eigenvector.

   While considering badminton as a Markov process, we constructed the transition matrix $A = \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix}$. Find its eigenvalues and eigenvectors.

2. Finding eigenvectors – a better approach

   This method is guaranteed to succeed only for the field of complex numbers, but the algorithm is valid for any field, and it finds the eigenvectors whenever they exist.

   Given matrix $A$, pick an arbitrary vector $\vec{\mathbf{w}}$. If you are really lucky, $A\vec{\mathbf{w}}$ is a multiple of $\vec{\mathbf{w}}$ and you have stumbled across an eigenvector. If not, keep computing $A^2\vec{\mathbf{w}}$, $A^3\vec{\mathbf{w}}$, etc. until you find a vector that is a linear combination of its predecessors. This situation is easily detected by row reduction.

   Now you have found a polynomial $p$ of degree $m$ such that $p(A)\vec{\mathbf{w}} = 0$. Furthermore, this is the nonzero polynomial of lowest degree for which $p(A)\vec{\mathbf{w}} = 0$. It is not necessarily the same as the characteristic polynomial.

   Over the complex numbers, this polynomial is guaranteed to have a root $\lambda$ by virtue of the "fundamental theorem of algebra" (Hubbard theorem 1.6.13). Over the real numbers or a finite field, it it will have a root in the field only if you are lucky. Assuming that the root exists, factor it out:

   $p(t) = (t - \lambda)q(t)$

   Now $p(A)\vec{\mathbf{w}} = (A - \lambda I)q(A)\vec{\mathbf{w}} = 0$.

   Thus $q(A)\vec{\mathbf{w}}$ is an eigenvector with eigenvalue $\lambda$.

   Here is a $2 \times 2$ example where the calculation is easy.

   Let $A = \begin{bmatrix} -1 & 4 \\ -2 & 5 \end{bmatrix}$

   As the arbitrary vector $\vec{\mathbf{w}}$ choose $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Compute $A\vec{\mathbf{w}}$ and $A^2\vec{\mathbf{w}}$.

Use row reduction to express the third of these vectors, $A^2\vec{\mathbf{w}}$, as a linear combination of the first two.

$$\begin{bmatrix} 1 & -1 & -7 \\ 0 & -2 & -8 \end{bmatrix}$$

Write the result in the form $p(A)\vec{\mathbf{w}} = 0$.

Factor: $p(t)=$

To get the eigenvector for eigenvalue 1, apply the remaining factor of $p(A), A - 3I$, to $\vec{\mathbf{w}}$.

To get the eigenvector for eigenvalue 3, apply the remaining factor of $p(A), A - I$, to $\vec{\mathbf{w}}$.

Citing your source: This technique was brought to the world's attention by Sheldon Axler's 1995 article "Down with Determinants" (see Hubbard page 224). Unlike most of what is taught in undergraduate math, it should probably be cited when you use it in other courses. An informal comment like "Using Axler's method for finding eigenvectors..." would suffice.

3. Eigenvectors and eigenvalues in a finite field

   Consider the matrix $A = \begin{bmatrix} 3 & 2 \\ 3 & 3 \end{bmatrix}$ with entries from the finite field $\mathbb{Z}_5$.

   (a) Find the eigenvalues of $A$ by solving the characteristic equation $\det(A - \lambda I) = 0$, then find the corresponding eigenvectors. Solving a quadratic equation over $\mathbb{Z}_5$ is easy – in a pinch, just try all five possible roots!

   (b) Find the eigenvalues of $A$ by using the technique of example 2.7.8 of Hubbard. You will get the same equation for the eigenvalues, of course, but it will be more straightforward to find the eigenvectors.

   (c) Write down the matrix $P$ whose columns are the basis of eigenvectors, and show that $P^{-1}AP$ is a diagonal matrix. Why is this reasonable?

4. When is there an eigenbasis?

   Choose $\vec{\mathbf{w}}$ successively to equal $\vec{\mathbf{e}}_1\vec{\mathbf{e}}_2, \cdots, \vec{\mathbf{e}}_n$.

   In searching for eigenvectors, we find successively polynomials $p_1(t), p_2(t), \cdots, p_n(t)$.

   There is a basis of real eigenvectors if and only if each of the polynomials $p_i(t)$ has **simple real roots**, e.g. $p(t) = t(t-2)(t+4)(t-3)$. No repeated factors are allowed!

   A polynomial like $p(t) = t^2 + 1$, although it has no repeated factors, has no real roots: $p(t) = (t+i)(t-i)$.

   If we allow complex roots, then any polynomial can be factored into linear factors (Fundamental Theorem of Algebra, Hubbard page 113).

   There is a basis of complex eigenvectors if and only if each of the polynomials $p_i(t)$ has **simple roots**, e.g. $p(t) = t(t-i)(t+i)$. No repeated factors are allowed!

   Here is a clever way to construct a matrix for which one of the polynomials $p_i(t)$ does not have simple roots and there is no basis of eigenvectors.

   Let $D = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$, $N = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}$. The matrix $N$ is a so-called "nilpotent" matrix: because its kernel is the same as its image, $N^2$ is the zero matrix.

   Show that the matrix $A = D + N$ has the property that if we choose any $\vec{\mathbf{w}}$ that is not in the kernel of $N$, then the polynomial $p(A)$ is $(A - 2I)^2$ and so there is no basis of eigenvectors.

5. The easy case: $n$ distinct real eigenvalues (Proof 4.1)

   If $\vec{\mathbf{v}}_1, \cdots, \vec{\mathbf{v}}_n$ are eigenvectors of $A : \mathbb{R}^n \to \mathbb{R}^n$ with distinct eigenvalues $\lambda_1 \cdots \lambda_n$, they are linearly independent.

   This proof could be done by induction, but there is an equivalent technique, using the "least number principle," that is a little bit easier.

   Suppose, for a contradiction, that the eigenvectors are linearly dependent.

   There exists a first eigenvector (the $j$th one) that is a linear combination of its predecessors:

   $$\vec{\mathbf{v}}_j = a_1\vec{\mathbf{v}}_1 + \cdots + a_{j-1}\vec{\mathbf{v}}_{j-1}.$$

   Multiply both sides by $A - \lambda_j I$. You get zero on the left, and on the right you get a linear combination where all the coefficients are nonzero because $\lambda_j - \lambda_i \neq 0$. This is in contradiction to the assumption that $\vec{\mathbf{v}}_j$ was the first one that is a linear combination of its predecessors.

   Since in $\mathbb{R}^n$ there cannot be more than $n$ linearly independent vectors, there are at most $n$ distinct eigenvalues.

6. Change of basis

   Our "old" basis consists of the standard basis vectors $\vec{\mathbf{e}}_1$ and $\vec{\mathbf{e}}_2$.

   Our "new" basis consists of one eigenvector for each eigenvalue of $A = \begin{bmatrix} -1 & 4 \\ -2 & 5 \end{bmatrix}$, with eigenvalues $\lambda_1 = 1$ and $\lambda_2 = 3$.

   Let's choose $\vec{\mathbf{v}}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and $\vec{\mathbf{v}}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

   It would be all right to multiply either of these vectors by a constant or to reverse their order.

   Write down the change of basis matrix $P$ whose columns express the new basis vectors in term of the old ones.

   Calculate the inverse change of basis matrix $P^{-1}$ whose columns express the old basis vectors in terms of the new ones.

   We are considering a linear transformation that is represented, relative to the standard basis, by the matrix $A$. What diagonal matrix $D$ represents this linear transformation relative to the new basis of eigenvectors?

   Confirm that $A = PDP^{-1}$. We have "diagonalized" the matrix $A$.

$$\begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$$

7. Diagonalization and eigenvectors

Let $P$ be a matrix whose columns are linearly independent eigenvectors of the $n \times n$ matrix $A$. (Such a matrix does not exist for every matrix $A$.) Denote by $\lambda_i$ the eigenvalue corresponding to the eigenvector that is column $i$.

How do we know that the columns of $P$ form a basis for $\mathbb{R}^n$?

Let $D = P^{-1}AP$.

Prove that $D$ is the diagonal matrix

$$D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

.

by applying $D = P^{-1}AP$ to standard basis vectors.

Conversely, let $D$ have the specified form, and prove that $A\vec{\mathbf{v}}_i = \lambda_i\vec{\mathbf{v}}_i$

by applying $A = PDP^{-1}$ to eigenvectors.

8. Eigenvectors for a $3 \times 3$ matrix

For Hubbard Example 2.7.8, the calculation is best subcontracted to R. The matrix is

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

Since we have help with the computation, make the choice $\vec{\mathbf{w}} = \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix}$.

The matrix to row reduce is

$\begin{bmatrix} 2 & -1 & 0 & 3 \\ 3 & -1 & -3 & -9 \\ 5 & 2 & 3 & 6 \end{bmatrix}$, different from the matrix in Hubbard.

The result of row reduction is the same:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 1 & 4 \end{bmatrix}$$

The rest of the work is easily done by hand.

Using the last column, write the polynomial $p(t)$, and factor it.

Find an eigenvector that corresponds to the smallest positive eigenvalue. It is not necessary to use the same $\vec{\mathbf{w}}$; any vector will do, as long as it is not in the subspace spanned by the other eigenvectors. Hubbard uses $\vec{\mathbf{e}}_1$. Use $\vec{\mathbf{e}}_3$ instead.

9. When is there an eigenbasis?

   This is a difficult issue in general. The simple case is where we are lucky and find a polynomial $p$ of degree $n$ that has $n$ distinct roots. In that case we can find $n$ eigenvectors, and it has already been proved that they are linearly independent. They form an eigenbasis. If the roots are real, the eigenvectors are elements of $\mathbb{R}^n$. If the roots are distinct but not all real, the eigenvectors are still a basis of $\mathbb{C}^n$.

   Suppose we try each standard basis vector in turn as $\vec{\mathbf{w}}$. Using $\vec{\mathbf{e}}_i$ leads to a polynomial $p_i$. If every $p_i$ is a polynomial of degree $m_i < n$, the situation is more complicated. Theorem 2.7.9 in Hubbard states the result:

   There exists an eigenbasis of $\mathbb{C}^n$ if and only if all the roots of all the $p_i$ are simple.

   Before doing the difficult proof, look the simplest examples of matrices that do not have $n$ distinct eigenvalues.

   - Let $A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$. In this case every vector in $\mathbb{R}^2$ is an eigenvector with eigenvalue 2. There is only one eigenvalue, but any basis is an eigenbasis.
     If we choose $\vec{\mathbf{w}} = \vec{\mathbf{e}}_1$ and form the matrix whose columns are $\vec{\mathbf{w}}$ and $A\vec{\mathbf{w}}$,
     $\begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}$,

     the matrix is already in echelon form.
     What is $p_1$?

     What eigenvector do we find?

     What eigenvector do we find if we choose $\vec{\mathbf{w}} = \vec{\mathbf{e}}_2$?

     Key point: we found a basis of eigenvectors, even though there was only one eigenvalue, and the polynomial $(t-2)^2$ never showed up.

- Let $A = \begin{bmatrix} 2 & 0 \\ 1 & 2 \end{bmatrix}$. In this case there is only one eigenvalue and there is no eigenbasis.

  What happens if we choose $\vec{\mathbf{w}} = \vec{\mathbf{e}}_2$?

  If we choose $\vec{\mathbf{w}} = \vec{\mathbf{e}}_1$,

  confirm that $\begin{bmatrix} 1 & 2 & 4 \\ 0 & 1 & 4 \end{bmatrix}$

  row reduces to $\begin{bmatrix} 1 & 0 & -4 \\ 0 & 1 & 4 \end{bmatrix}$.

  What is $p_1$?

  What happens when we carry out the procedure that usually gives an eigenvector?

  Key point: There was only one eigenvalue, the polynomial $(t-2)^2$ showed up, and we were unable to find a basis of eigenvectors.

10. An instructive $3 \times 3$ example

The surprising case, and the one that makes the proof difficult, is the one where there exists a basis of eigenvectors but there are fewer than $n$ distinct eigenvalues. A simple example is $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$

Here each standard basis vector is an eigenvector. For the first one the eigenvalue is 1; for the second and third, it is 2.

A less obvious example is

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

When we use row reduction to find the eigenvectors,we obtain the following results:

Using $\vec{\mathbf{w}} = \vec{\mathbf{e}}_1$, we get $p_1(t) = t - 2$ and find an eigenvector

$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ with eigenvalue 2.

Using $\vec{\mathbf{w}} = \vec{\mathbf{e}}_2$, we get $p_2(t) = (t-1)(t-2)$ and find two eigenvectors:

$\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ with eigenvalue 1, $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ with eigenvalue 2.

At this point we have found three linearly independent eigenvectors and we have a basis.

If we use $\vec{\mathbf{w}} = \vec{\mathbf{e}}_3$, we get $p_3(t) = (t-1)(t-2)$ and find two eigenvectors:

$\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ with eigenvalue 1, $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ with eigenvalue 2.

In general, if we use some arbitrary $\vec{\mathbf{w}}$, we will get $p(t) = (t-1)(t-2)$ and we will find the eigenvector with eigenvalue 1 along with some linear combination of the eigenvectors with eigenvalue 2.

Key points about this case:

- The polynomial $p_i(t)$, in order to be simple, must have degree less than $n$.
- We need to use more than one standard basis vector in order to find a basis of eigenvectors.

11. Proof that if all roots are simple there is an eigenbasis

    Assume that whenever we choose $\vec{\mathbf{w}} = \vec{\mathbf{e}}_i$, the polynomial $p_i$ of degree $m_i$ has simple roots. The columns of the matrix that we row reduce are $\vec{\mathbf{e}}_i, A\vec{\mathbf{e}}_i, \cdots A^{m_i}\vec{\mathbf{e}}_i$. The image of this matrix has three properties.

    - It is a subspace $E_i$ of $\mathbb{R}^n$.
    - It includes $m_i$ eigenvectors. Since these correspond to distinct eigenvalues, they are linearly independent, and therefore they span $E_i$.
    - It includes $\vec{\mathbf{e}}_i$.

    Now take the union of all the $E_i$. This union has the following properties:

    - It includes each standard basis vector $\vec{\mathbf{e}}_i$, so it is all of $\mathbb{R}^n$.
    - It is spanned by the union of the sets of eigenvectors. In general there will be more than $n$ vectors in this set. Use them as columns of a matrix. The image of this matrix is all of $\mathbb{R}^n$. We can find a basis for the image consisting of $n$ columns, which are all eigenvectors.

12. Proof that if there is an eigenbasis, each $p_i$ has simple roots.

    There are $k$ distinct eigenvalues, $\lambda_1, \cdots, \lambda_k$. It is entirely possible that $k < n$, since different eigenvectors may have the same eigenvalue.

    Since there is a basis of eigenvectors, we can express each $\vec{\mathbf{e}}_i$ as a linear combination of eigenvectors.

    Define $p_i(t) = \prod(t - \lambda_j)$. The product extends just over the set of eigenvalues that are associated with the eigenvectors needed to express $\vec{\mathbf{e}}_i$ as a linear combination, so there may be fewer than $k$ factors.

    Form $p_i(A) = \prod(A - \lambda_j I)$. The factors can be in any order. If $\vec{\mathbf{w}}$ is any eigenvector whose eigenvalue $\lambda_j$ is included in the product, then $(A - \lambda_j I)\vec{\mathbf{w}} = 0$ and so $p_i(A)\vec{\mathbf{w}} = \vec{\mathbf{0}}$. Since those eigenvectors from a basis for a subspace that includes $\vec{\mathbf{e}}_i$, it follows that $p_i(A)\vec{\mathbf{e}}_i = \vec{\mathbf{0}}$.

    If we form a nonzero polynomial $p_i'(t)$ of lower degree by omitting one factor from the product, then $p_i'(A)\vec{\mathbf{e}}_i \neq \vec{\mathbf{0}}$, since the eigenvectors that correspond to the omitted eigenvalue do not get killed off.

    So $p_i(t)$ is the nonzero polynomial of lowest degree for which $p_i(A)\vec{\mathbf{e}}_i = \vec{\mathbf{0}}$, and by construction it has simple roots.

13. Proof 4.2, first half

    Assume that whenever we choose $\vec{\mathbf{w}} = \vec{\mathbf{e}}_i$, the polynomial $p_i$ of degree $m_i$ has simple real roots. Consider the subspace $E$ that is the image of the matrix whose columns are
    $\vec{\mathbf{e}}_1, A\vec{\mathbf{e}}_1, \cdots A^{m_1}\vec{\mathbf{e}}_1, \vec{\mathbf{e}}_2, A\vec{\mathbf{e}}_2, \cdots A^{m_2}\vec{\mathbf{e}}_2, \cdots, \vec{\mathbf{e}}_n, A\vec{\mathbf{e}}_n, \cdots A^{m_n}\vec{\mathbf{e}}_n$.
    Prove that $E = \mathbb{R}^n$ (easy) and that there exists a basis for $E$ that consists entirely of eigenvectors(harder).

    Theorem 2.7.9 in Hubbard is more powerful, because it applies to the complex case. The proof is the same. Our proof is restricted to the real case only because we are not doing examples with complex eigenvectors.

14. Proof 4.2, second half

    Assume that there is a basis of $\mathbb{R}^n$ consisting of eigenvectors of $n \times n$ matrix $A$, but that $A$ has only $k \leq n$ distinct real eigenvalues. Prove that for any basis vector $\vec{\mathbf{w}} = \vec{\mathbf{e}}_i$, the polynomial $p_i(t)$ has simple roots.

15. Fibonacci numbers by matrices

The usual way to generate the Fibonacci sequence is to set $a_0 = 1, a_1 = 1$, then calculate $a_2 = a_0 + a_1 = 2$, $a_3 = a_1 + a_2 = 3$, etc.

In matrix notation this can be written

$$\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

and more generally

$$\begin{bmatrix} a_n \\ a_{n+1} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}^n \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Use this approach to determine $a_2$ and $a_3$, doing the matrix multiplication first.

$$\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$$

Determine $a_6$ and $a_7$ by using the square of the matrix that was just constructed.

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$$

We have found a slight computational speedup, but it would be nicer to have a general formula for $\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}^n$.

16. Powers of a diagonal matrix.

    For a $2 \times 2$ diagonal matrix, $\begin{bmatrix} c_1 & 0 \\ 0 & c_2 \end{bmatrix}^n = \begin{bmatrix} c_1^n & 0 \\ 0 & c_2^n \end{bmatrix}$.

    Now suppose that we want to compute $A^n$. If there is a basis of eigenvectors, we can construct the matrix $P$, whose columns are eigenvectors, such that

$$P^{-1}AP = \begin{bmatrix} c_1 & 0 \\ 0 & c_2 \end{bmatrix} = D.$$

    Prove by induction that

$$(P^{-1}AP)^n = P^{-1}A^nP.$$

    For the Fibonacci example, where $A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$,

    $D \begin{bmatrix} (1+\sqrt{5})/2 & 0 \\ 0 & (1+\sqrt{5})/2 \end{bmatrix}$ and $P = \begin{bmatrix} 2 & 2 \\ 1+\sqrt{5} & 1-\sqrt{5} \end{bmatrix}$

    You can check that $PD^2P^{-1} = A^2$
    (messy because of the irrational numbers!)
    Example 2.7.1 in the textbook computes $A^n = PD^nP^{-1}$ and gets a formula for the $n$th Fibonacci number that is useful in computer science:

$$a^n = \frac{5+\sqrt{5}}{10}(\frac{1+\sqrt{5}}{2})^n + \frac{5-\sqrt{5}}{10}(\frac{1-\sqrt{5}}{2})^n$$

17. Dealing with complex eigenvalues in the $2 \times 2$ case

If the polynomial $p(t)$ for a matrix $A$ has two distinct complex roots, it is possible to compute $A^n$ by diagonalizing the matrix and calculating $A^n = PD^nP^{-1}$. However, both the eigenvectors and eigenvalues will be complex, and we know that $A^n$ is a real matrix. It seems a shame to have to do complex arithmetic to compute a matrix of real numbers.

Fortunately, there is another type of $2 \times 2$ matrix that is almost as easy to raise to a power as a diagonal matrix – a conformal matrix.

Recall that the matrix $C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} = \begin{bmatrix} r\cos\theta & -r\sin\theta \\ r\sin\theta & r\cos\theta \end{bmatrix}$ represents the complex number $a + bi = re^{i\theta}$.

Its $n$th power is $C^n = \begin{bmatrix} r^n\cos n\theta & -r^n\sin n\theta \\ r^n\sin n\theta & r^n\cos n\theta \end{bmatrix}$.

Given a $2 \times 2$ matrix $A$ with complex eigenvalues $a \pm bi$, we just need to invent a matrix $P$ such that $A = PCP^{-1}$.

Here is an algorithm for constructing the matrices $C$ and $P$:

Suppose that the eigenvalues of $A$ are $a \pm ib$, with $b \neq 0$. Then no real $\vec{\mathbf{w}}$ can be an eigenvector, and for any real $\vec{\mathbf{w}}$ we will find the polynomial

$p(t) = (t - a - ib)(t - a + ib) = (t - a)^2 + b^2$

So $p(A) = (A - aI)^2 + b^2I = 0$ or

$$(\frac{A - aI}{b})^2 = -I.$$

Now we need to construct a new basis, which will not be a basis of eigenvectors but which will still be useful.

Set $\vec{\mathbf{v}}_1 = \vec{\mathbf{e}}_1, \vec{\mathbf{v}}_2 = (\frac{A-aI}{b})\vec{\mathbf{e}}_1$.

Then $(A - aI)\vec{\mathbf{v}}_1 = b\vec{\mathbf{v}}_2$ and $A\vec{\mathbf{v}}_1 = a\vec{\mathbf{v}}_1 + b\vec{\mathbf{v}}_2$ .

Also, $(\frac{A-aI}{b})\vec{\mathbf{v}}_2 = (\frac{A-aI}{b})^2\vec{\mathbf{v}}_1 = -\vec{\mathbf{v}}_1$, so

$(A - aI)\vec{\mathbf{v}}_2 = -b\vec{\mathbf{v}}_1$ and $A\vec{\mathbf{v}}_2 = a\vec{\mathbf{v}}_2 - b\vec{\mathbf{v}}_1$.

With respect to the new basis, the matrix that represents $A$ is the conformal matrix $C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$.

If we define $P$ in the usual way with columns $\vec{\mathbf{v}}_1$ and $\vec{\mathbf{v}}_2$, then $A = PCP^{-1}$, and the matrices $P$ and $C$ are real.

18. Conformal matrices and complex numbers – an example

(a) Show that the polynomial $p(t)$ for the matrix $A = \begin{bmatrix} 7 & -10 \\ 2 & -1 \end{bmatrix}$ has roots $3 \pm 2i$.

(b) Show that $(\frac{A-3I}{2})^2 = -I$.

(c) Choose a new basis with $\vec{\mathbf{v}}_1 = \vec{\mathbf{e}}_1, \vec{\mathbf{v}}_2 = (\frac{A-3I}{2})\vec{\mathbf{e}}_1$.
Use these basis vectors as the columns of matrix $P$.
Confirm that $A = PCP^{-1}$, where $C$ is conformal and $P$ is real.

19. What to do if there is no basis of eigenvectors

    In the $2 \times 2$ case, suppose there is no basis of eigenvectors, just a single eigenvector $\vec{\mathbf{v}}$ with eigenvalue $\lambda$. If we choose $\vec{\mathbf{w}} \neq \vec{\mathbf{v}}$, we will get a polynomial whose roots are not simple: $p(t) = (t - \lambda)^2$.

    So $p(A) = (A - \lambda I)^2 = 0$. In other words, $N = A - \lambda I$ is a nilpotent matrix, with $N^2 = 0$.

    It is easy to raise $A = \lambda I + N$ to the $n$th power.

    Prove by induction that

    $A^k = (\lambda I + N)^k = \lambda^k I + k\lambda^{k-1} N.$

    As an example, let $A = \begin{bmatrix} 5 & -1 \\ 4 & 1 \end{bmatrix}$ and compute $A^4$

20. What about non-integer powers?

    If $A = PDP^{-1}$ or $A = PCP^{-1}$ (basis of real or complex eigenvectors), then it is possible to raise $A$ to a non-integer power! For example, you can compute $B = A^{\frac{1}{2}}$ so that $B^2 = A$.. Be careful, though – there may be many answers or none.

    - In the diagonal case where $A = PDP^{-1}$ the matrix $D$ must have non-negative entries (the eigenvalues cannot be negative.) For a positive eigenvalue, there are two possible square roots.
    - In the conformal case where $A = PCP^{-1}$ the conformal matrix $C$ has two different square roots, one with angle $\frac{1}{2}\theta$ and one with angle $\frac{1}{2}\theta + \pi$.

21. A couple of short but interesting proofs

Suppose that $S$ is a symmetric matrix with $n$ distinct real eigenvalues. Prove that you can construct an orthonormal basis of eigenvectors.

Suppose that matrix $AB$ has a nonzero eigenvalue $\lambda$. Prove that $\lambda$ is also an eigenvalue for $BA$.

# 3 Seminar Topics

Your section instructor will either have emailed a list of topics to prepare or will have posted a signup list of appointments on the Calendar tab of Canvas. Either way, there will be one of the following topics that you should be prepared to present.

Practice your presentation so that it takes about 8 minutes. The text of the presentation will be projected onto a screen so that you need not recopy it. To save time, avoid writing long sentences on the chalkboard. You may use notes, but be discreet about it.

1. Define *eigenvector* and *eigenvalue* for a matrix $A$, and prove that $\lambda$ (which may be a complex number) is an eigenvalue of $A$ if and only if it is a root of the characteristic polynomial $\chi_A(\lambda) = \det(A - \lambda I)$.

2. (Proof 4.1)
   Prove that if $\vec{\mathbf{v}}_1, \cdots, \vec{\mathbf{v}}_n$ are eigenvectors of $A : \mathbb{R}^n \to \mathbb{R}^n$ with distinct eigenvalues $\lambda_1 \cdots \lambda_n$, they are linearly independent. Conclude that an $n \times n$ matrix cannot have more than $n$ distinct eigenvalues.

3. Suppose that for $n \times n$ matrix $A$ there is a basis of $n$ eigenvectors $\{\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \cdots, \vec{\mathbf{v}}_n\}$ with associated eigenvalues $\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$. Although the eigenvectors must be independent, some of the eigenvalues may repeat.

   Create a matrix $P$ whose columns are the eigenvectors. Since the eigenvectors form a basis, they are independent and the matrix $P$ has an inverse $P^{-1}$. Prove the following:

   (a) Matrix $P$ is invertible.

   (b) The matrix $D = P^{-1}AP$ is a diagonal matrix. .

   (c) The matrix $A$ can be expressed as $A = PDP^{-1}$.

   (d) The matrix $A^n$ can be expressed as $A^n = PD^nP^{-1}$ (use induction; do not write $PDP^{-1} \cdots PDP^{-1}$.)

4. For the matrix $A = \begin{bmatrix} 3 & 2 \\ 0 & 3 \end{bmatrix}$, apply our standard technique for finding eigenvectors, first choosing $\vec{\mathbf{w}} = \vec{\mathbf{e}}_1$, then choosing $\vec{\mathbf{w}} = \vec{\mathbf{e}}_2$. Show that $p_2(t)$ does not have simple roots, that you find only one eigenvector, and that $A$ cannot be "diagonalized" (written in the form $A = PDP^{-1}$.)

   Write $A$ in the form $A = \lambda I + N$, where $N$ is called "nilpotent" because its square is the zero matrix. Prove by induction that there is still a simple formula for the powers of $A$, namely $A^n = \lambda^n I + n\lambda^{n-1}N$.

5. (Proof 4.2)

- For real $n \times n$ matrix $A$, prove that if all the polynomials $p_i(t)$ are simple and have real roots, then there exists a basis for $\mathbb{R}^n$ consisting of eigenvectors of $A$.
- Prove that if there exists a basis for $\mathbb{R}^n$ consisting of eigenvectors of $A$, then all the polynomials $p_i(t)$ are simple and have real roots.

# 4 Workshop Problems

1. Some interesting examples with $2 \times 2$ matrices

   (a) Since a polynomial equation with real (or complex) coefficients always has a root (the "fundamental theorem of algebra"), a real matrix is guaranteed to have at least one complex eigenvalue. No such theorem holds for polynomial equations with coefficients in a finite field, so zero eigenvalues is a possibility. This is one of the few results in linear algebra that depends on the underlying field.

   Consider the matrix $A = \begin{bmatrix} 3 & 1 \\ n & 3 \end{bmatrix}$ with entries from the finite field $\mathbb{Z}_5$.

   By considering the characteristic equation, find values of $n$ that lead to 2, 1, or 0 distinct eigenvalues. For the case of 1 eigenvalue, find an eigenvector.

   Hint: After writing the characteristic equation with $n$ isolated on the right side of the equals sign, make a table of the value of $t^2 + 4t + 4$ for each of the five possible eigenvalues. That table lets you determine how many solutions there are for each of the five possible values of $n$. When the characteristic polynomial is the square of a linear factor, there is only one eigenvector and it is easy to construct.

   (b) Extracting square roots by diagonalization.

   The matrix $A = \begin{bmatrix} 2 & 1 \\ 2 & 3 \end{bmatrix}$

   conveniently has two eigenvalues that are perfect squares. Find a basis of eigenvectors and construct a matrix $P$ such that $P^{-1}AP$ is a diagonal matrix.

   Thereby find two independent square roots of $A$, i.e. find matrices $B_1$ and $B_2$ such that $B_1^2 = B_2^2 = A$ , with $B_2 \neq \pm B_1$. Hint: use the negative square root of one of the eigenvalues, the positive square root of the other.

   If you take Physics 15c next year, you may encounter this technique when you study "coupled oscillators."

2. Some proofs. In doing these, you may use the fact that an eigenbasis exists if and only if all the $p_i(t)$ have simple roots.

   (a) Suppose that a $5 \times 5$ matrix has a basis of eigenvectors, but that its only eigenvalues are 1 and 2. Using Hubbard Theorem 2.7.9, show that you must make at least three different choices of $\vec{\mathbf{e}}_i$ in order to find all the eigenvectors.

   (b) An alternative approach to proof 4.1 – use induction.

   Identify a base case (easy). Then show that, if a set of $k-1$ eigenvectors with distinct eigenvalues is linearly independent and you add to the set an eigenvector $\vec{\mathbf{v}}_k$ with an eigenvalue $\lambda_k$ that is different from any of the preceding eigenvalues, the resulting set of $k$ eigenvectors with distinct eigenvalues is linearly independent.

3. Examples where there is no basis of eigenvectors

   (a) The matrix $A = \begin{bmatrix} 1 & -1 \\ 4 & -3 \end{bmatrix}$ has only a single eigenvalue and only one independent eigenvector.

   Find the eigenvalue and eigenvector, show that $A = D + N$ where $D$ is diagonal and $N$ is nilpotent, and use the formula from seminar topic 4 to calculate $A^3$ without ever multiplying $A$ by itself (unless you want to check your answer).

   (b) Find two eigenvectors for the matrix $A = \begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & 1 \\ -2 & 2 & 0 \end{bmatrix}$. and confirm that using each of the three standard basis vectors in turn will not produce a third independent eigenvector.

   Clearly the columns of $A$ are not independent; so 0 is an eigenvalue. This property makes the algebra fairly easy.

4. Problems with $3 \times 3$ matrices, to be solved by writing or editing R scripts

   (a) Sometimes you don't find all the eigenvectors on the first try.

   The matrix $A = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

   has three real, distinct eigenvalues, and there is a basis of eigenvectors. Find what polynomial equation for the eigenvalues arises from each of the following choices, and use it to construct as many eigenvectors as possible.:

   - $\vec{\mathbf{w}} = \vec{\mathbf{e_1}}$.
   - $\vec{\mathbf{w}} = \vec{\mathbf{e_3}}$.
   - $\vec{\mathbf{w}} = \vec{\mathbf{e_1}} + \vec{\mathbf{e_3}}$.

   (b) Use the technique of example 2.7.8 in Hubbard to find the eigenvalues and eigenvectors of the matrix $A = \begin{bmatrix} 3 & 4 & -4 \\ 1 & 3 & -1 \\ 3 & 6 & -4 \end{bmatrix}$

# 5 Homework

1. Consider the sequence of numbers described, in a manner similar to the Fibonacci numbers, by

   $b_3 = 2b_1 + b_2$
   $b_4 = 2b_2 + b_3$

   $b_{n+2} = 2b_n + b_{n+1}$

   (a) Write a matrix $B$ to generate this sequence in the same way that Hubbard generates the Fibonacci numbers.

   (b) By considering the case $b_1 = 1, b_2 = 2$ and the case $b_1 = -1, b_2 = 1$, find the eigenvectors and eigenvalues of $B$.

   (c) Express the vector $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ as a linear combination of the two eigenvectors, and thereby find a formula for $b_n$ if $b_1 = 1, b_2 = 1$.

2. (This is similar to group problem 1b.)

   Consider the matrix $A = \begin{bmatrix} -10 & 9 \\ -18 & 17 \end{bmatrix}$.

   (a) By using a basis of eigenvectors, find a matrix $P$ such that $P^{-1}AP$ is a diagonal matrix.

   (b) Find a cube root of $A$, i.e. find a matrix $B$ such that $B^3 = A$.

3. (a) Prove that if $\vec{\mathbf{v}}_1$ and $\vec{\mathbf{v}}_2$ are eigenvectors of matrix $A$, both with the same eigenvalue $\lambda$, then any linear combination of $\vec{\mathbf{v}}_1$ and $\vec{\mathbf{v}}_2$ is also an eigenvector.

   (b) Suppose that $A$ is a $3 \times 3$ matrix with a basis of eigenvectors but with only two distinct eigenvalues. Prove that for any $\vec{\mathbf{w}}$, the vectors $\vec{\mathbf{w}}$, $A\vec{\mathbf{w}}$, and $A^2\vec{\mathbf{w}}$ are linearly dependent. (This is another way to understand why all the polynomials $p_i(t)$ are simple when $A$ has a basis of eigenvectors but a repeated eigenvalue.)

4. Harvard graduate Ivana Markov, who concentrated in English and mathematics with economics as a secondary field, just cannot decide whether she wants to be a poet or an investment banker, and so her career path is described by the following Markov process:

   - If Ivana works as a poet in year $n$, there is a probability of 0.9 that she will feel poor at the end of the year and take a job as an investment banker for year $n+1$. Otherwise she remains a poet.
   - If Ivana works as an investment banker in year $n$, there is a probability of 0.7 that she will feel overworked and unfulfilled at the end of the year and take a job as a poet for year $n+1$. Otherwise she remains an investment banker.

   Thus, if $\begin{bmatrix} p_n \\ q_n \end{bmatrix}$ describes the probabilities that Ivana works as a poet or a banker respectively in year $n$, the corresponding probabilities for year $n+1$ are given by $\begin{bmatrix} p_{n+1} \\ q_{n+1} \end{bmatrix} = A \begin{bmatrix} p_n \\ q_n \end{bmatrix}$, where $A = \begin{bmatrix} 0.1 & 0.7 \\ 0.9 & 0.3 \end{bmatrix}$

   (a) Find the eigenvalues and eigenvectors of $A$.

   (b) Construct the matrix $P$ whose columns are the eigenvectors, invert it, and thereby express the vector $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ as a linear combination of the eigenvectors.

   (c) Suppose that in year 0 Ivana works as a poet, so that $\begin{bmatrix} p_0 \\ q_0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Find an explicit formula for $\begin{bmatrix} p_n \\ q_n \end{bmatrix}$ and use it to determine $\begin{bmatrix} p_{10} \\ q_{10} \end{bmatrix}$. What happens in the limit of large $n$?

5. In general, the square matrix $A$ that represents a Markov process has the property that all the entries are between 0 and 1 and each column sums to 1. Prove that such a matrix $A$ has an eigenvalue of 1 and that there is a "stationary vector" that is transformed into itself by $A$. You may use the fact, which we have proved so far only for 2×2 and $3 \times 3$ matrices, that if a matrix has a nonzero vector in its kermel, its determinant is zero.

6. (a) Prove by induction (no "$\cdots$" allowed!) that if $F = PCP^{-1}$, then $F^n = PC^nP^{-1}$ for all positive integers $n$.

   (b) Suppose that $2 \times 2$ real matrix $F$ has complex eigenvalues $re^{\pm i\theta}$. Show that, for integer $n$, $F^n$ is a multiple of the identity matrix if and only if $n\theta = m\pi$ for some integer $m$. Hint: write $F = PCP^{-1}$ where $C$ is conformal. This hint also helps with the rest of the problem.

   (c) If $F = \begin{bmatrix} 3 & 7 \\ -1 & -1 \end{bmatrix}$, find the smallest $n$ for which $F^n$ is a multiple of the identity. Check your answer by matrix multiplication.

   (d) If $G = \begin{bmatrix} -2 & -15 \\ 3 & 10 \end{bmatrix}$, use half-angle formulas to find a matrix $A$ for which $A^2 = G$. Check your answer by matrix multiplication.

7. (You may use this problem as a third R script problem, in which case you will automatically get credit for it as an ordinary homework problem)

   Use the technique of example 2.7.8 in Hubbard to find the eigenvalues and eigenvectors of the following two matrices. One has a repeated eigenvalue and will require you to use the technique with two different basis vectors.

   (a) $A = \begin{bmatrix} 3 & 4 & -4 \\ 1 & 3 & -1 \\ 3 & 6 & -4 \end{bmatrix}$

   (b) $B = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 3 & -1 \\ 1 & 2 & 0 \end{bmatrix}$

   Optional problems that require writing or editing R scripts

8. The matrix $A = \begin{bmatrix} 5 & 1 & 1 \\ -1 & 3 & -1 \\ 0 & 0 & 4 \end{bmatrix}$ has only one eigenvalue, 4, and so its characteristic polynomial must be $(t-4)^3$.

   (a) Show that $A$ has a two-dimensional subspace of eigenvectors but that there is no other eigenvector.

   (b) Write $A = D + N$ where $D$ is diagonal and $N$ is nilpotent, and confirm that $N^2$ is the zero matrix.

9. Here is a symmetric matrix, which is guaranteed to have an orthonormal basis of eigenvectors. For once, the numbers have not been rigged to make the eigenvalues be integers.

   $$A = \begin{bmatrix} 4 & -1 & 1 \\ -1 & 3 & 2 \\ 1 & 2 & -3 \end{bmatrix}$$

   Express $A$ in the form $PDP^{-1}$, where $D$ is diagonal and $P$ is an isometry matrix whose columns are orthogonal unit vectors.

   A similar example is in script 1.4X.

10. Problem 7 above.