Review for Midterm 3

Selected Topics

Eigenvectors and Eigenvalues

Definition

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

1. An **eigenvector** of A is a nonzero vector v in $\mathbf{R}^n$ such that $Av = \lambda v$, for some λ in $\mathbf{R}$. In other words, Av is a multiple of v.
2. An **eigenvalue** of A is a number λ in $\mathbf{R}$ such that the equation $Av = \lambda v$ has a nontrivial solution.

If $Av = \lambda v$ for $v \neq 0$, we say λ is the **eigenvalue for** v, and v is an **eigenvector for** λ.

Definition

Let A be an $n \times n$ matrix and let λ be an eigenvalue of A. The λ-**eigenspace** of A is the set of all eigenvectors of A with eigenvalue λ, plus the zero vector:

$$\begin{aligned}\lambda\text{-eigenspace} &= \{v \text{ in } \mathbf{R}^n \mid Av = \lambda v\} \\ &= \{v \text{ in } \mathbf{R}^n \mid (A - \lambda I)v = 0\} \\ &= \operatorname{Nul}(A - \lambda I).\end{aligned}$$

You find a basis for the λ-eigenspace by finding the parametric vector form for the general solution to $(A - \lambda I)x = 0$ using row reduction.

The Characteristic Polynomial

Definition

Let A be an $n \times n$ matrix. The **characteristic polynomial** of A is

$$f(\lambda) = \det(A - \lambda I).$$

Important Facts:

1. The characteristic polynomial is a polynomial of degree n, of the following form:

$$f(\lambda) = (-1)^n \lambda^n + a_{n-1}\lambda^{n-1} + a_{n-2}\lambda^{n-2} + \cdots + a_1\lambda + a_0.$$

2. The eigenvalues of A are the roots of $f(\lambda)$.
3. The constant term $f(0) = a_0$ is equal to $\det(A)$:

$$f(0) = \det(A - 0I) = \det(A).$$

4. The characteristic polynomial of a 2×2 matrix A is

$$f(\lambda) = \lambda^2 - \mathrm{Tr}(A)\,\lambda + \det(A).$$

Definition

The **algebraic multiplicity** of an eigenvalue λ is its multiplicity as a root of the characteristic polynomial.

Similarity

Definition

Two $n \times n$ matrices A and B are **similar** if there is an invertible $n \times n$ matrix P such that

$$A = PBP^{-1}.$$

Important Facts:

1. Similar matrices have the same characteristic polynomial.
2. It follows that similar matrices have the same eigenvalues.
3. If A is similar to B and B is similar to C, then A is similar to C.

Caveats:

1. Matrices with the same characteristic polynomial need not be similar.
2. Similarity has nothing to do with row equivalence.
3. Similar matrices usually do not have the same eigenvectors.

Let $A = PBP^{-1}$, and let $v_1, v_2, \ldots, v_n$ be the columns of P. These form a basis $\mathcal{B}$ for $\mathbf{R}^n$ because P is invertible. *Key relation:* for any vector x in $\mathbf{R}^n$,

$$[Ax]_{\mathcal{B}} = B[x]_{\mathcal{B}}.$$

This says:

A acts on the usual coordinates of x
in the same way that
B acts on the $\mathcal{B}$-coordinates of x.

Example:

$$A = \frac{1}{4}\begin{pmatrix} 5 & 3 \\ 3 & 5 \end{pmatrix} \qquad B = \begin{pmatrix} 2 & 0 \\ 0 & 1/2 \end{pmatrix} \qquad P = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Then $A = PBP^{-1}$. B acts on the usual coordinates by scaling the first coordinate by 2, and the second by $1/2$:

$$B\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2x_1 \\ x_2/2 \end{pmatrix}.$$

The unit coordinate vectors are eigenvectors: e_1 has eigenvalue 2, and e_2 has eigenvalue $1/2$.

Similarity

Example

$$A = \frac{1}{4}\begin{pmatrix} 5 & 3 \\ 3 & 5 \end{pmatrix} \qquad B = \begin{pmatrix} 2 & 0 \\ 0 & 1/2 \end{pmatrix} \qquad P = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \qquad [Ax]_{\mathcal{B}} = B[x]_{\mathcal{B}}.$$

In this case, $\mathcal{B} = \left\{ \binom{1}{1}, \binom{1}{-1} \right\}$. Let $v_1 = \binom{1}{1}$ and $v_2 = \binom{1}{-1}$.

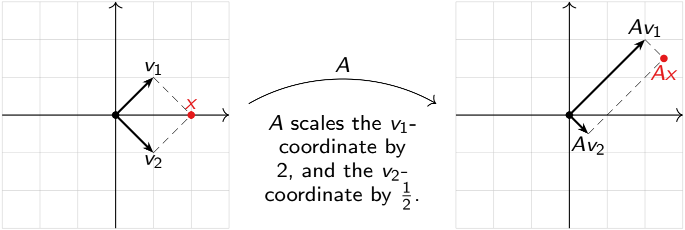
To compute $y = Ax$:

1. Find $[x]_{\mathcal{B}}$.
2. $[y]_{\mathcal{B}} = B[x]_{\mathcal{B}}$.
3. Compute y from $[y]_{\mathcal{B}}$.

Say $x = \binom{2}{0}$.

1. $x = v_1 + v_2$ so $[x]_{\mathcal{B}} = \binom{1}{1}$.
2. $[y]_{\mathcal{B}} = B\binom{1}{1} = \binom{2}{1/2}$.
3. $y = 2v_1 + \frac{1}{2}v_2 = \binom{5/2}{3/2}$.

Picture:



Diagonalization

Definition

An $n \times n$ matrix A is **diagonalizable** if it is similar to a diagonal matrix:

$$A = PDP^{-1} \quad \text{for } D \text{ diagonal.}$$

It is easy to take powers of diagonalizable matrices:

$$A^n = PD^nP^{-1}.$$

The Diagonalization Theorem

An $n \times n$ matrix A is diagonalizable if and only if A has n linearly independent eigenvectors. In this case, $A = PDP^{-1}$ for

$$P = \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{pmatrix} \qquad D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix},$$

where $v_1, v_2, \ldots, v_n$ are linearly independent eigenvectors, and $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the corresponding eigenvalues (in the same order).

Corollary

An $n \times n$ matrix with n distinct eigenvalues is diagonalizable.

Non-Distinct Eigenvalues

Definition

Let A be a square matrix with eigenvalue λ. The **geometric multiplicity** of λ is the dimension of the λ-eigenspace.

Theorem

Let A be an $n \times n$ matrix. Then A is diagonalizable if and only if, for every eigenvalue λ, the algebraic multiplicity of λ is equal to the geometric multiplicity.

(And all eigenvalues are real, unless you want to diagonalize over $\mathbf{C}$.)

Notes:

- The algebraic and geometric multiplicities are both whole numbers ≥ 1, and the algebraic multiplicity is always greater than or equal to the geometric multiplicity. In particular, they're equal if the algebraic multiplicity is 1.
- Equivalently, A is diagonalizable if and only if the sum of the geometric multiplicities of its eigenvalues is n.

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

This has eigenvalues 1 and 2, with algebraic multiplicities 2 and 1, respectively.

The geometric multiplicity of 2 is *automatically* 1.

Let's compute the geometric multiplicity of 1:

$$A - I = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{\text{rref}} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

This has 1 free variable, so the geometric multiplicity of 1 is 1. This is less than the algebraic multiplicity, so the matrix is *not diagonalizable*.

Definition

A square matrix A is **stochastic** if all of its entries are nonnegative, and the sum of the entries of each column is 1. It A is **positive** if all of its entries are positive.

Definition

A *steady state* for a stochastic matrix A is an eigenvector w with eigenvalue 1, such that its entries are positive and sum to 1.

Perron–Frobenius Theorem

If A is a positive stochastic matrix, then it admits a unique steady state vector w, which spans the 1-eigenspace.

Moreover, for any vector v_0 with entries summing to some number c, the iterates $v_1 = Av_0$, $v_2 = Av_1$, ..., $v_n = Av_{n-1}$, ..., approach cw as n gets large.

Think about it in terms of Red Box movies: v_n is the number of movies in each location on day n, and $v_{n+1} = Av_n$. Eventually, the number of movies in each location will be the same every day: $v_n = v_{n+1} = Av_n$. This means v_n is an eigenvector with eigenvalue 1, so it is a multiple of the steady state w: $v_n = cw$. The steady state w tells you the *percentages* of movies that are in each location, so c is the total number of movies. So if you started with $c = 100$ movies on day 0, then you know $v_n = cw = 100w$ for large enough n: the total number of movies doesn't change.

Computing the Steady State

$$A = \begin{pmatrix} .3 & .4 & .5 \\ .3 & .4 & .3 \\ .4 & .2 & .2 \end{pmatrix}$$

This is a positive stochastic matrix. To compute the steady state, first we find *some* eigenvector with eigenvalue 1:

$$A - I = \begin{pmatrix} -.7 & .4 & .5 \\ .3 & -.6 & .3 \\ .4 & .2 & -.8 \end{pmatrix} \xrightarrow{\text{rref}} \begin{pmatrix} 1 & 0 & -7/5 \\ 0 & 1 & -6/5 \\ 0 & 0 & 0 \end{pmatrix}.$$

The parametric vector form is $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = z \begin{pmatrix} 7/5 \\ 6/5 \\ 1 \end{pmatrix}$, so an eigenvector is $\begin{pmatrix} 7 \\ 6 \\ 5 \end{pmatrix}$.

We want the entries of our eigenvector to sum to 1, so we need to divide by the sum of the entries:

$$w = \frac{1}{7+6+5} \begin{pmatrix} 7 \\ 6 \\ 5 \end{pmatrix} = \frac{1}{18} \begin{pmatrix} 7 \\ 6 \\ 5 \end{pmatrix}.$$

This is the steady state. If $v = (6, 22, 8)$ then $A^n v$ approaches $36w = (14, 12, 10)$.

Complex Eigenvectors

Complex eigenvalues and eigenvectors work just like their real counterparts, with the additional fact:

> Both eigenvalues and eigenvectors of real square matrices occur in conjugate pairs.

Example: $A = \begin{pmatrix} \sqrt{3}+1 & -2 \\ 1 & \sqrt{3}-1 \end{pmatrix}$. The characteristic polynomial is

$$f(\lambda) = \lambda^2 - \mathrm{Tr}(A)\,\lambda + \det(A) = \lambda^2 - 2\sqrt{3}\,\lambda + 4.$$

The quadratic formula tells us the eigenvalues are

$$\lambda = \frac{2\sqrt{3} \pm \sqrt{(2\sqrt{3})^2 - 16}}{2} = \sqrt{3} \pm i.$$

Let's compute an eigenvector v with eigenvalue $\lambda = \sqrt{3} - i$.

$$A - \lambda I = \begin{pmatrix} 1+i & -2 \\ \star & \star \end{pmatrix} \rightsquigarrow v = \begin{pmatrix} 2 \\ 1+i \end{pmatrix}.$$

An eigenvector with eigenvalue $\sqrt{3} + i$ is (automatically) $\binom{2}{1-i}$.

Geometric Interpretation of Complex Eigenvalues

Theorem

Let A be a 2×2 matrix with complex (non-real) eigenvalue λ, and let v be an eigenvector. Then

$$A = PCP^{-1}$$

where

$$P = \begin{pmatrix} | & | \\ \operatorname{Re} v & \operatorname{Im} v \\ | & | \end{pmatrix} \quad \text{and} \quad C = \begin{pmatrix} \operatorname{Re}\lambda & \operatorname{Im}\lambda \\ -\operatorname{Im}\lambda & \operatorname{Re}\lambda \end{pmatrix}.$$

The matrix C is a composition of a counterclockwise rotation by $-\arg(\lambda)$, and a scale by a factor of $|\lambda|$.

Example:

$$A = \begin{pmatrix} \sqrt{3}+1 & -2 \\ 1 & \sqrt{3}-1 \end{pmatrix} \qquad \lambda = \sqrt{3} - i \qquad v = \begin{pmatrix} 1-i \\ 1 \end{pmatrix}$$

This gives

$$C = \begin{pmatrix} \operatorname{Re}\lambda & \operatorname{Im}\lambda \\ -\operatorname{Im}\lambda & \operatorname{Re}\lambda \end{pmatrix} = \begin{pmatrix} \sqrt{3} & -1 \\ 1 & \sqrt{3} \end{pmatrix}$$

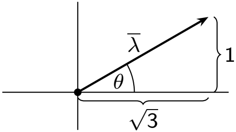
$$P = \begin{pmatrix} \operatorname{Re}(1-i) & \operatorname{Im}(1-i) \\ \operatorname{Re}(1) & \operatorname{Im}(1) \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}$$

$$A = \begin{pmatrix} \sqrt{3}+1 & -2 \\ 1 & \sqrt{3}-1 \end{pmatrix} \quad C = \begin{pmatrix} \sqrt{3} & -1 \\ 1 & \sqrt{3} \end{pmatrix} \quad P = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix} \quad \lambda = \sqrt{3}-i$$

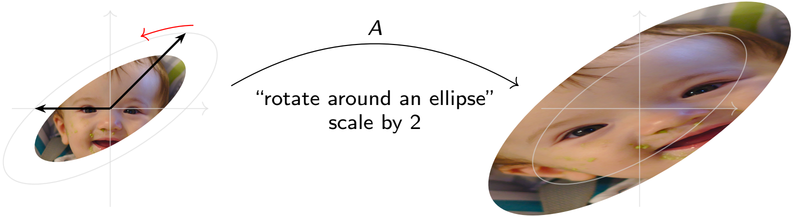
The Theorem says that C scales by a factor of

$$|\lambda| = \sqrt{(\sqrt{3})^2 + (-1)^2} = \sqrt{3+1} = 2.$$

It rotates counterclockwise by the argument of $\overline{\lambda} = \sqrt{3} + i$, which is $\pi/6$:



$$\theta = \tan^{-1}\left(\frac{1}{\sqrt{3}}\right) = \frac{\pi}{6}$$



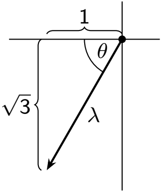
Warning: if $\lambda = a + bi$, you can't just plug $\tan^{-1}(b/a)$ into your calculator and expect to get the argument of λ.

Example: If $\lambda = -1 - \sqrt{3}i$ then

$$\tan^{-1}\left(\frac{-\sqrt{3}}{-1}\right) = \tan^{-1}(\sqrt{3}) = \frac{\pi}{3}.$$

Anyway that's the number your calculator will give you.

You have to *draw a picture*:



$$\theta = \tan^{-1}(\sqrt{3}) = \frac{\pi}{3}$$

$$\text{argument} = \theta + \pi = \frac{4\pi}{3}$$