Math 1600B Lecture 23, Section 2, 5 Mar 2014

Announcements:

Read Section 4.2 for next class. Work through recommended homework questions.

Midterms available for pick-up starting Thursday. (If you have a Wednesday tutorial, your TA will be available Thursday or Friday.) Solutions will be posted Thursday.

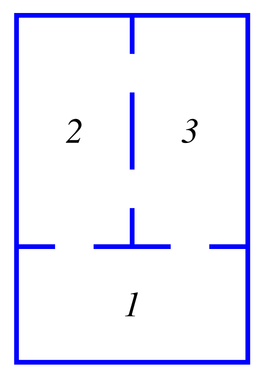
Drop date: Friday, March 7.

Tutorials: Quiz this week covers Sections 3.5 and 3.6, focusing on 3.6.

Help Centers: Monday-Friday 2:30-6:30 in MC 106.

New Material: Section 3.7: Markov chains (cont)

Example 3.65: A Markov chain can have more than two states. A rat is in a maze with three rooms, and always chooses to go through one of the doors with equal probability. Draw the state diagram, determine the transition matrix P and describe how to find a steady-state vector.



Solution: Draw **state diagram** on board.

From this, we find the **transition matrix**

$$P = \begin{bmatrix} 0 & 1/3 & 1/3 \\ 1/2 & 0 & 2/3 \\ 1/2 & 2/3 & 0 \end{bmatrix}$$

The P_{ij} entry is the probability of going from room j to room i. Note that the columns are **probability vectors** (non-negative entries that sum to 1) and so P is a **stochastic matrix** (square, with columns being probability vectors).

A **steady state vector** is a vector $\vec{x}$ such that $P\vec{x} = \vec{x}$. That is, $\vec{x} - P\vec{x} = \vec{0}$, or $(I - P)\vec{x} = \vec{0}$. To see if there is a non-trivial steady state vector for this Markov chain, we solve the homogeneous system with coefficient matrix $I - P$:

$$\left[\begin{array}{ccc|c} 1 & -1/3 & -1/3 & 0 \\ -1/2 & 1 & -2/3 & 0 \\ -1/2 & -2/3 & 1 & 0 \end{array}\right]$$

In RREF:

$$\left[\begin{array}{ccc|c} 1 & 0 & -2/3 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

So $x_3 = t$, $x_2 = t$ and $x_1 = \frac{2}{3}t$. If we want a probability vector, then we want $t + t + \frac{2}{3}t = 1$, so $t = 3/8$, so we get $\begin{bmatrix} 2/8 \\ 3/8 \\ 3/8 \end{bmatrix}$.

Theorem: Every Markov chain has a non-trivial steady state vector.

This appears in the book as Theorem 4.30 in Section 4.6.

Proof: Let P be the transition matrix. We want to find a non-trivial solution to $(I - P)\vec{x} = \vec{0}$. By the fundamental theorem of invertible matrices and the fact that $\operatorname{rank}(I - P) = \operatorname{rank}((I - P)^T)$, this is equivalent to $(I - P)^T\vec{x} = \vec{0}$ having a non-trivial solution. That is, finding a non-trivial $\vec{x}$ such that

$$P^T\vec{x} = \vec{x} \quad (\text{since } I^T = I).$$

But since P is a stochastic matrix, we always have

$$P^T \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

So therefore $P\vec{x} = \vec{x}$ also has a (different) non-trivial solution. $\square$

We'll probably study Markov chains again in Section 4.6.

Section 4.1: Eigenvalues and eigenvectors

We saw when studying Markov chains that it was important to find solutions to the system $A\vec{x} = \vec{x}$, where A is a square matrix. We did this by solving $(I - A)\vec{x} = \vec{0}$.

More generally, a central problem in linear algebra is to find $\vec{x}$ such that $A\vec{x}$ is a *scalar multiple* of $\vec{x}$.

Definition: Let A be an $n \times n$ matrix. A scalar λ (lambda) is called an **eigenvalue** of A if there is a nonzero vector $\vec{x}$ such that $A\vec{x} = \lambda\vec{x}$. Such a vector $\vec{x}$ is called an **eigenvector** of A corresponding to λ.

We showed that $\lambda = 1$ is an eigenvalue of every stochastic matrix P.

Example A: Since

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

we see that 2 is an eigenvalue of $\begin{bmatrix} 1 & 2 \\ 2 & -2 \end{bmatrix}$ with eigenvector $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$.

Example 4.2: Show that 5 is an eigenvalue of $A = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}$ and determine all eigenvectors corresponding to this eigenvalue.

Solution: We are looking for nonzero solutions to $A\vec{x} = 5\vec{x}$. This is the same as $(A - 5I)\vec{x} = \vec{0}$, so we compute the coefficient matrix:

$$A - 5I = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix} - \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix} = \begin{bmatrix} -4 & 2 \\ 4 & -2 \end{bmatrix}$$

The columns are linearly dependent, so the null space of $A - 5I$ is nonzero. So $A\vec{x} = 5\vec{x}$ has a nontrivial solution, which is what it means for 5 to be an eigenvalue.

To find the eigenvectors, we compute the null space of $A - 5I$:

$$[\, A - 5I \mid \vec{0}\,] = \left[\begin{array}{cc|c} -4 & 2 & 0 \\ 4 & -2 & 0 \end{array}\right] \to \left[\begin{array}{cc|c} 1 & -1/2 & 0 \\ 0 & 0 & 0 \end{array}\right]$$

The solutions are of the form $\begin{bmatrix} t/2 \\ t \end{bmatrix} = t\begin{bmatrix} 1/2 \\ 1 \end{bmatrix}$. So the eigenvectors for the eigenvalue 5 are the *nonzero* multiples of $\begin{bmatrix} 1/2 \\ 1 \end{bmatrix}$.

Definition: Let A be an $n \times n$ matrix and let λ be an eigenvalue of A. The collection of all eigenvectors corresponding to λ, together with the zero vector, is a subspace called the **eigenspace** of λ and is denoted E_λ. In other words,

$$E_\lambda = \text{null}(A - \lambda I).$$

In the above Example, $E_5 = \text{span}\left\{ \begin{bmatrix} 1/2 \\ 1 \end{bmatrix} \right\}$.

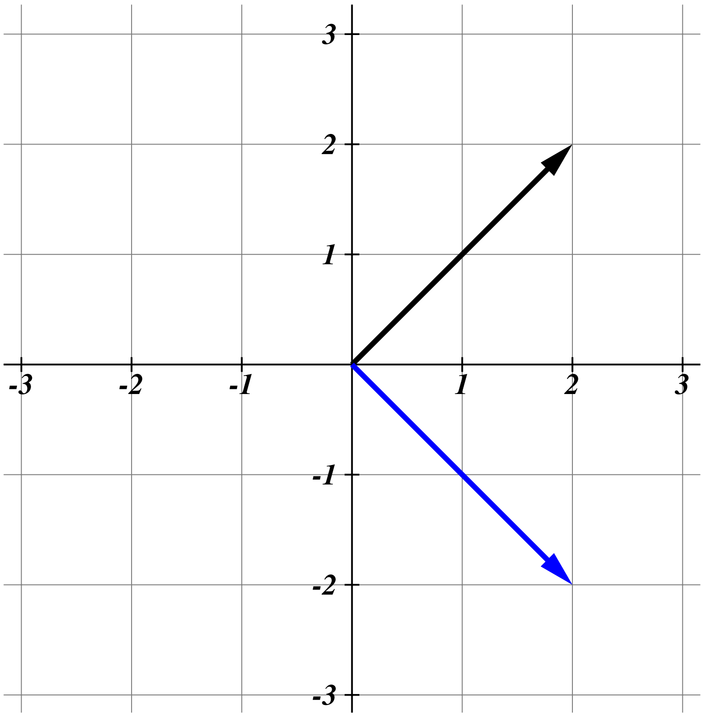
Example: Give an eigenvalue of the matrix $A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ and compute its eigenspace.

Since $A\vec{x} = 2\vec{x}$ for every $\vec{x}$, 2 is an eigenvalue, and is the only eigenvalue. In this case, $E_2 = \mathbb{R}^2$.

Example: If 0 is an eigenvalue of A, what is another name for E_0?

E_0 is the null space of $A - 0I = A$. That is, $E_0 = \text{null}(A)$.

An applet illustrating the transformation $T_A : \mathbb{R}^2 \to \mathbb{R}^2$, for A the 2×2 matrix shown. The black vector is the input $\vec{x}$, and the blue vector is the output $T_A(\vec{x}) = A\vec{x}$.



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

$$\vec{x} = \begin{bmatrix} 2.0 \\ 2.0 \end{bmatrix} \quad A\vec{x} = \begin{bmatrix} 2.0 \\ -2.0 \end{bmatrix}$$

Reflection in x-axis.
Reflection in y-axis.
Projection onto x-axis.
Rotation by 90° ccw.
Rotate and scale.
Example A from above.
A rank 1 example.
Custom:

(Click to move input vector. Hit 't' to toggle modes. Click on a phrase to the right to change the matrix. Enter four numbers, separated by spaces, for a custom matrix.)

Applet: See also this java applet. (Instructions.) If that doesn't work, here is another applet.

Read Example 4.3 in the text for a 3×3 example.

Finding eigenvalues

Given a specific number λ, we now know how to check whether λ is an eigenvalue: we check whether $A - \lambda I$ has a nontrivial null space. And we can find the eigenvectors by finding the null space.

We also have a geometric way to find **all** eigenvalues λ, at least in the 2×2 case. Is there an algebraic way to check all λ at once?

By the fundamental theorem of invertible matrices, $A - \lambda I$ has a nontrivial null space if and only if it is not invertible. For 2×2 matrices, we can check invertibility using the determinant!

Example: Find all eigenvalues of $A = \begin{bmatrix} 1 & 2 \\ 2 & -2 \end{bmatrix}$.

Solution: We need to find all λ such that $\det(A - \lambda I) = 0$.

$$\begin{aligned} \det(A - \lambda I) &= \det \begin{bmatrix} 1-\lambda & 2 \\ 2 & -2-\lambda \end{bmatrix} \\ &= (1-\lambda)(-2-\lambda) - 4 = \lambda^2 + \lambda - 6, \end{aligned}$$

so we need to solve the quadratic equation $\lambda^2 + \lambda - 6 = 0$. This can be factored as $(\lambda - 2)(\lambda + 3) = 0$, and so $\lambda = 2$ or $\lambda = -3$, the same as we saw above and with the applet.

We proceed to find the eigenvectors for these eigenvalues, by solving $(A - 2I)\vec{x} = \vec{0}$ and $(A + 3I)\vec{x} = \vec{0}$. On board, if time.

Appendix D provides review of polynomials and their solutions.

See also Example 4.5 in text.

So now we know how to handle the 2×2 case. To handle larger matrices, we need to learn about their determinants, which is Section 4.2.

We won't discuss eigenvectors and eigenvalues for matrices over $\mathbb{Z}_m$. We will discuss complex numbers $\mathbb{C}$ in a later lecture.