Math 415 - Lecture 31

Markov matrices and Google

Monday November 9th 2015

Textbook reading: Chapter 5.3

Suggested practice exercises: Chapter 5.3: 8, 9, 12, 14, 10.

Khan Academy video: Finding Eigenvectors and Eigenspaces example

Strang lecture: Lecture 21: Eigenvalues and eigenvectors Lecture 24: Markov Matrices and Fourier Series.

1 Review

1.1 Properties of eigenvectors and eigenvalues

- If $A\mathbf{x} = \lambda\mathbf{x}$ then $\mathbf{x}$ is an **eigenvector** of A with **eigenvalue** λ.All eigenvectors (plus $\mathbf{0}$) with eigenvalue λ form **eigenspace** of λ.
- λ is an eigenvalue of $A \iff \underbrace{\det(A - \lambda I)}_{\text{characteristic polynomial}} = 0$.Why? Because $A\mathbf{x} = \lambda\mathbf{x} \iff (A - \lambda I)\mathbf{x} = \mathbf{0}$.By the way: this means that the eigenspace of λ is just $\mathrm{Nul}(A - \lambda I)$.
- E.g. if $A = \begin{bmatrix} 3 & 2 & 3 \\ 0 & 6 & 10 \\ 0 & 0 & 2 \end{bmatrix}$ then $\det(A - \lambda I) = (3 - \lambda)(6 - \lambda)(2 - \lambda)$.

If $A = \begin{bmatrix} 3 & 2 & 3 \\ 0 & 6 & 10 \\ 0 & 0 & 2 \end{bmatrix}$ then the eigenvalues are $2, 3, 6$ with corresponding eigenvectors

$$\begin{bmatrix} 2 \\ -5/2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2/3 \\ 1 \\ 0 \end{bmatrix}.$$

These three vectors are independent. By the next result, this is always so.

1.2 Independent eigenvectors

Theorem 1. *If $\mathbf{x}_1, \ldots, \mathbf{x}_m$ are eigenvectors of A corresponding to different eigenvalues, then they are independent.*

Proof. Suppose, for contradiction, that $\mathbf{x}_1, \ldots, \mathbf{x}_m$ are dependent. By kicking out some vectors, we may assume that there is (up to multiples) only one linear relation: $c_1\mathbf{x}_1 + \ldots + c_m\mathbf{x}_m = \mathbf{0}$. In other words, the matrix with columns $\mathbf{x}_1, \ldots, \mathbf{x}_m$ has one-dimensional null space. Now multiply this relation with A:

$$A(c_1\mathbf{x}_1 + \ldots + c_m\mathbf{x}_m) = c_1\lambda_1\mathbf{x}_1 + \ldots + c_m\lambda_m\mathbf{x}_m = \mathbf{0}$$

This is a second independent relation! In other words, two independent vectors living in a one-dimensional vector space. Contradiction. $\square$

2 Relations between eigenvalues

2.1 Product of Eigenvalues

If A is $n \times n$ get in principle n eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. How are these eigenvalues related?

Theorem 2. *The product of eigenvalues $\lambda_1\lambda_2 \ldots \lambda_n$ is equal to the determinant of A.*

Proof. The characteristic polynomial $p(\lambda) = \det(A - \lambda I)$ has constant term $\det(A)$. On the other hand $p(\lambda)$ factors, because the roots are the eigenvalues we get $p(\lambda) = (\pm 1)(\lambda - \lambda_1)(\lambda - \lambda_2) \ldots (\lambda - \lambda_n)$, which has constant term $\lambda_1\lambda_2 \ldots \lambda_n$. $\square$

Example 1. Let $A = \begin{bmatrix} \lambda_1 & b \\ 0 & \lambda_2 \end{bmatrix}$. Then the eigenvalues are λ_1, λ_2 and $\det(A) = \lambda_1\lambda_2$.

2.2 Sum of Eigenvalues

What other relations are there between the eigenvalues?

Definition 2. Let $A = \begin{bmatrix} a_{11} & a_{12} & \cdots \\ a_{21} & a_{22} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}$ be $n \times n$. Then the **TRACE** of A is the sum of the diagonal entries: $\mathrm{Tr}(A) = a_{11} + a_{22} + \cdots + a_{nn}$.

Theorem 3. *Let A be $n \times n$. Then the trace of A is the* ***sum*** *of eigenvalues:*

$$\mathrm{Tr}(A) = \lambda_1 + \lambda_2 + \cdots + \lambda_n$$

Example 3. Let $A = \begin{bmatrix} \lambda_1 & b \\ 0 & \lambda_2 \end{bmatrix}$. What are the eigenvalues and what is $\mathrm{Tr}(A)$?

Solution. The eigenvalues are λ_1, λ_2 and $\mathrm{Tr}(A) = \lambda_1 + \lambda_2$.

2.3 The Characteristic Polynomial for 2×2

2×2 matrices are easy.

Theorem 4. *Let* $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. *Then the characteristic polynomial is*

$$p(\lambda) = \lambda^2 - (a+d)\lambda + ad - bc = \lambda^2 - \operatorname{Tr}(A)\lambda + \det(A).$$

Example 4. Let $A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$. What are the eigenvalues and what is the characteristic polynomial?

Solution. $\operatorname{Tr}(A) = 6$, $\det(A) = 8$, so $p(\lambda) = \lambda^2 - 6\lambda + 8$. Also in terms of eigenvalues $\operatorname{Tr}(A) = \lambda_1 + \lambda_2$ and $\det(A) = \lambda_1\lambda_2$. So $\lambda_1 = 2, \lambda_2 = 4$

3 Practice problems

Example 5. Find the eigenvectors and eigenvalues of $A = \begin{bmatrix} 0 & -2 \\ -4 & 2 \end{bmatrix}$.

Example 6. What are the eigenvalues of $A = \begin{bmatrix} 2 & 0 & 0 & 0 \\ -1 & 3 & 0 & 0 \\ -1 & 1 & 3 & 0 \\ 0 & 1 & 2 & 4 \end{bmatrix}$. No calculations!

Example 7. Find the eigenvalues of A as well as a basis for the corresponding eigenspaces, where

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

Solution.

- The characteristic polynomial is:

$$\begin{aligned} \det(A - \lambda I) &= \begin{vmatrix} 2-\lambda & 0 & 0 \\ -1 & 3-\lambda & 1 \\ -1 & 1 & 3-\lambda \end{vmatrix} \\ &= (2-\lambda)\begin{vmatrix} 3-\lambda & 1 \\ 1 & 3-\lambda \end{vmatrix} \\ &= (2-\lambda)[(3-\lambda)^2 - 1] \\ &= (2-\lambda)(\lambda-2)(\lambda-4) \end{aligned}$$

- A has eigenvalues $2, 2, 4$ $\left(A = \begin{bmatrix} 2 & 0 & 0 \\ -1 & 3 & 1 \\ -1 & 1 & 3 \end{bmatrix} \right)$ Since $\lambda = 2$ is a double root, it has **(algebraic) multiplicity** 2.

- $\lambda_1 = 2$:
$$(A - \lambda_1 I)\mathbf{x} = \begin{bmatrix} 0 & 0 & 0 \\ -1 & 1 & 1 \\ -1 & 1 & 1 \end{bmatrix} \mathbf{x} = \mathbf{0}$$
Two independent solutions: $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$ In other words, the eigenspace for $\lambda_1 = 2$ is Span $\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \right\}$.
- $\lambda_2 = 4$: $\left(A = \begin{bmatrix} 2 & 0 & 0 \\ -1 & 3 & 1 \\ -1 & 1 & 3 \end{bmatrix} \right)$
$$(A - \lambda_2 I)\mathbf{x} = \begin{bmatrix} -2 & 0 & 0 \\ -1 & -1 & 1 \\ -1 & 1 & -1 \end{bmatrix} \mathbf{x} = \mathbf{0} \implies \mathbf{x}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$
- In summary, A has eigenvalues 2 and 4:
  - eigenspace for $\lambda_1 = 2$ has basis $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$,
  - eigenspace for $\lambda_2 = 4$ has basis $\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$

4 Markov matrices

Definition 8. An $n \times n$ matrices A is **Markov matrix** if has non negative entries, and the entries in each column add to 1.

Theorem 5. *Let A be a Markov matrix. Then*

(i) 1 is an eigenvalue of A and any other eigenvalue λ satisfies $|\lambda| \leq 1$.

(ii) If A has only positive entries, then any other eigenvalue satisfies $|\lambda| < 1$.

Example 9. Let A be
$$\begin{bmatrix} .9 & .2 \\ .1 & .8 \end{bmatrix}.$$
Is A a Markov matrix?

Theorem 6. *Let A be an $n \times n$-Markov matrix with only positive entries and let $\boldsymbol{v} \in \mathbb{R}^n$. Then*
$$\boldsymbol{v}_\infty := \lim_{k \to \infty} A^k \boldsymbol{v} \text{ exists,}$$
*and $A\boldsymbol{v}_\infty = \boldsymbol{v}_\infty$. In this case $\boldsymbol{v}_\infty$ is often called the **steady state**.*

Proof. If $\mathbf{x}$ is any vector and $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$ is an eigenbasis for a Markov matrix ($A\mathbf{v_1} = 1\mathbf{v_1}$):

$$\mathbf{x} = c_1\mathbf{v_1} + \cdots + c_n\mathbf{v_n},$$

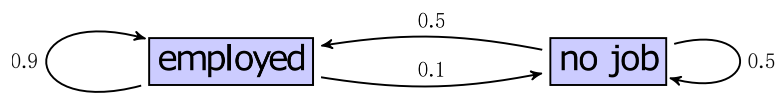
then

$$A^k\mathbf{x} = c_1\lambda_1^k\mathbf{v_1} + \cdots + c_n\lambda_n^k\mathbf{v_n} \to c_1\mathbf{v_1},$$

if the eigenspace of $\lambda = 1$ is 1-dimensional.

□

Example 10. Consider a fixed population of people with or without job. Suppose that each year, 50% of those unemployed find a job while 10% of those employed lose their job.What is the unemployment rate in the long term equilibrium?



Solution. x_t: proportion of population employed at time t (in years) y_t: proportion of population unemployed at time t

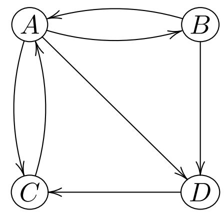
$$\begin{bmatrix} x_{t+1} \\ y_{t+1} \end{bmatrix} = \begin{bmatrix} 0.9x_t + 0.5y_t \\ 0.1x_t + 0.5y_t \end{bmatrix} = \begin{bmatrix} 0.9 & 0.5 \\ 0.1 & 0.5 \end{bmatrix} \begin{bmatrix} x_t \\ y_t \end{bmatrix}$$

The matrix $\begin{bmatrix} 0.9 & 0.5 \\ 0.1 & 0.5 \end{bmatrix}$ is a **Markov matrix**. Its columns add to 1 and it has no negative entries. $\begin{bmatrix} x_\infty \\ y_\infty \end{bmatrix}$ is an equilibrium if $\begin{bmatrix} x_\infty \\ y_\infty \end{bmatrix} = \begin{bmatrix} 0.9 & 0.5 \\ 0.1 & 0.5 \end{bmatrix} \begin{bmatrix} x_\infty \\ y_\infty \end{bmatrix}$.In other words, $\begin{bmatrix} x_\infty \\ y_\infty \end{bmatrix}$ is an eigenvector with eigenvalue 1.Eigenspace of $\lambda = 1$: $\mathrm{Nul}\left(\begin{bmatrix} -0.1 & 0.5 \\ 0.1 & -0.5 \end{bmatrix}\right) = \mathrm{Span}\left\{\begin{bmatrix} 5 \\ 1 \end{bmatrix}\right\}$Since $x_\infty + y_\infty = 1$, we conclude that $\begin{bmatrix} x_\infty \\ y_\infty \end{bmatrix} = \begin{bmatrix} 5/6 \\ 1/6 \end{bmatrix}$. Hence, the unemployment rate in the long term equilibrium is 1/6

5 Page rank (or: the 25000000000 \$ eigenvector)

Google's success is based on an algorithm to rank webpages, the **Page rank**, named after Google founder Larry Page. The idea is to determine how likely it is that a web user randomly gets to a given webpage. The webpages are ranked by these probabilities.
Suppose the internet consisted of the only four webpages A, B, C, D linked as

in the following graph.



Imagine a surfer following these links at random. For the probability $PR_n(A)$ that she is at A (after n steps), we need to think about how she could have reached A. We add:

- the probability that she was at B (at exactly one step before), and left for A,(that's $PR_{n-1}(B) \cdot \frac{1}{2}$)
- the probability that she was at C, and left for A,
- the probability that she was at D, and left for A.

Hence: $PR_n(A) = PR_{n-1}(B) \cdot \frac{1}{2} + PR_{n-1}(C) \cdot \frac{1}{1} + PR_{n-1}(D) \cdot \frac{0}{1}$.

Encode the probabilties at step n in a state vector with four entries. $\begin{bmatrix} PR_n(A) \\ PR_n(B) \\ PR_n(C) \\ PR_n(D) \end{bmatrix} =$

$$\begin{bmatrix} 0 & \frac{1}{2} & 1 & 0 \\ \frac{1}{3} & 0 & 0 & 0 \\ \frac{1}{3} & 0 & 0 & 1 \\ \frac{1}{3} & \frac{1}{2} & 0 & 0 \end{bmatrix} \begin{bmatrix} PR_{n-1}(A) \\ PR_{n-1}(B) \\ PR_{n-1}(C) \\ PR_{n-1}(D) \end{bmatrix}$$

Definition 11. The **PageRank vector** is the long-term equilibrium. It is an eigenvector of the Markov matrix with eigenvalue 1.

Let's call the Markov matrix with the probabilities T:

- $T - 1I = \begin{bmatrix} -1 & \frac{1}{2} & 1 & 0 \\ \frac{1}{3} & -1 & 0 & 0 \\ \frac{1}{3} & 0 & -1 & 1 \\ \frac{1}{3} & \frac{1}{2} & 0 & -1 \end{bmatrix} \overset{\text{RREF}}{\rightsquigarrow} \begin{bmatrix} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & -\frac{2}{3} \\ 0 & 0 & 1 & -\frac{5}{3} \\ 0 & 0 & 0 & 0 \end{bmatrix}$

  $\implies$ eigenspace of $\lambda = 1$ is spanned by $\begin{bmatrix} 2 \\ \frac{2}{3} \\ \frac{5}{3} \\ 1 \end{bmatrix}$.

- Now we need to make the entries add up to 1.

$$\begin{bmatrix} PR(A) \\ PR(B) \\ PR(C) \\ PR(D) \end{bmatrix} = \frac{3}{16} \begin{bmatrix} 2 \\ \frac{2}{3} \\ \frac{5}{3} \\ 1 \end{bmatrix} = \begin{bmatrix} 0.375 \\ 0.125 \\ 0.313 \\ 0.188 \end{bmatrix}.$$

  This is the PageRank vector.

- The corresponding ranking of the webpages is A, C, D, B.

Remark. In practical situations the system might be too large for finding the eigenvalues by row operations.

- Google reports having met 60 trillion webpages. Google's search index is over $100,000,000$ gigabytes. Number of Google's servers is secret: about $2,500,000$ More than 1,000,000,000 websites (i.e. hostnames; about 75% not active)

- Thus we have a gigantic but very sparse matrix.

An alternative to row operations is the **power method** (see Theorem 2):

Power method
If T is an (acyclic and irreducible) Markov matrix, then for any $\mathbf{v}_0$ the vectors $T^n\mathbf{v}_0$ converge to an eigenvector with eigenvalue 1

Here:

$$T = \begin{bmatrix} 0 & \frac{1}{2} & 1 & 0 \\ \frac{1}{3} & 0 & 0 & 0 \\ \frac{1}{3} & 0 & 0 & 1 \\ \frac{1}{3} & \frac{1}{2} & 0 & 0 \end{bmatrix}.$$

Start with an arbitrary state vector, hit it with powers of T.

$$\left(\begin{bmatrix} PR(A) \\ PR(B) \\ PR(C) \\ PR(D) \end{bmatrix} = \begin{bmatrix} 0.375 \\ 0.125 \\ 0.313 \\ 0.188 \end{bmatrix} \right), T \begin{bmatrix} 1/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{bmatrix} = \begin{bmatrix} 0.375 \\ 0.083 \\ 0.333 \\ 0.208 \end{bmatrix}$$

. Note that the ranking of the webpages is already A, C, D, B if we stop here.

$$T \begin{bmatrix} 1/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{bmatrix} = \begin{bmatrix} 0.375 \\ 0.083 \\ 0.333 \\ 0.208 \end{bmatrix}, \qquad T^2 \begin{bmatrix} 1/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{bmatrix} = \begin{bmatrix} 0.375 \\ 0.125 \\ 0.333 \\ 0.167 \end{bmatrix}, \qquad T^3 \begin{bmatrix} 1/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{bmatrix} = \begin{bmatrix} 0.396 \\ 0.125 \\ 0.292 \\ 0.188 \end{bmatrix}$$

Remark.
- If all entries of T are positive (no zero entries!), then the power method is guaranteed to work.

- In the context of PageRank, we can make sure that this is the case by replacing T with

$$(1-p) \cdot \begin{bmatrix} 0 & \frac{1}{2} & 1 & 0 \\ \frac{1}{3} & 0 & 0 & 0 \\ \frac{1}{3} & 0 & 0 & 1 \\ \frac{1}{3} & \frac{1}{2} & 0 & 0 \end{bmatrix} + p \cdot \begin{bmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix}.$$

Just to make sure: still a Markov matrix, now with positive entries Google used to use $p = 0.15$.

6 Practice problems

Problem 12. *Can you see why* 1 *is an eigenvalue for every Markov matrix?*

Problem 13 (just for fun)**.** *The real web contains pages which have no outgoing links. In that case, our random surfer would get "stuck" (the transition matrix is not a Markov matrix). Do you have an idea how to deal with this issue?*