Student: xxx xxx
Math Tutor: William Wu { willywu@stanford.edu }
Date: October 21, 2006

Lesson Plan

(1) Outline:

(a) Discuss the Faustian round table puzzle.
(b) Discuss the homework from last week.
(c) PageRank Part 1: Perron-Frobenius Theorem and Markov Matrices
(d) Mini-lecture: how to sum polynomials, and why it works out.

Last Week's Puzzle

- Week 5: Faustian round table. Hint: Symmetry.

PageRank Part 1: Perron-Frobenius Theorem and Markov Matrices

I hope you found the handout interesting. The results rely on two important and powerful theorems, but their proofs are not given! Therefore, by now you should be frothing at the mouth with voracious hunger for the proofs. Are you dying of hunger for the proofs?

Today we we will prove the first theorem in its entirety. Recall that in the handout, it is only argued that there is an eigenvector with eigenvalue 1. However, it is not explained why that eigenvector should be positive. So this will be our goal today. I should note that the proof of this fact supplied in Strang's textbook is just plain wrong as far as I can tell, so this makes this lecture all the more useful for supplementing and correcting the textbook.

The proof is a bit involved. We will need to prove 5 little things before finally being able to conclude the eigenvector's positivity. There may be an easier way to arrive at this conclusion, but this is the way I was able to find out.

First we need to know a really neat fact from real analysis, which you will learn someday:

> **Brouwer's Fixed Point Theorem**
> Let S be a nonempty, closed, bounded convex set in $\mathbf{R}^n$. Let $f : S \to S$ be a continuous map. Then there exists an $x \in S$ such that $f(x) = x$.

This statement has many terms which are deeper than they appear, and which you are not expected to really understand. For now, just think of it this way: if the set S is "nice", and the function f is continuous (like an unbroken curve), then there must exist at least one point that doesn't move under the map. One example of a map that naturally gives rise to this concept, at least intuitively, is a contraction map. Imagine a rubber oval that gets mapped to a slightly smaller rubber oval, kind of shrinking in on itself. Well, there should exist some point, maybe close to the center of the oval, that doesn't move at all because it is at the black hole of the shrinking process. I may do a fun mini-lecture someday about one way to go about possibly discovering this theorem.

We first prove the following four facts about non-negative matrices, collectively known as the famous Perron-Frobenius Theorem:

> **Perron-Frobenius**
> Let $A > 0$ be an $n \times n$ matrix. Then
>
> (1) There is a positive eigenvector y with positive eigenvector λ_0: $Ay = \lambda_0 y$. (Sidenote: Also, $\sum_j y_j = 1$.)
> (2) The eigenvalue λ_0 is maximal in modulus among all the eigenvalues of A. That is, for any eigenvalue μ of A, $|\mu| \leq \lambda_0$.
> (3) The eigenvalue λ_0 is geometrically simple. That is, any two eigenvectors corresponding to λ_0 are linearly dependent – they all lie on the same line. (Question: What's an example of a matrix with eigenvalues that are not geometrically simple?)
> (4) Any positive eigenvector of A (corresponding to any eigenvalue) is a scalar multiple of y. (Note: We won't actually use this one, but I include it here for completeness.)

(Note that we will be applying this result to the final matrix Q in the handout, which is not only Markov, but also all positive. Although Theorem 1 in the handout is a statement merely about Markov matrices, we will cheat a little bit and add the positivity condition, since the matrix which Google actually uses is both Markov and positive anyways.)

Proof. Let

$$S = \{(x_1, x_2, \ldots, x_n)^T : \sum_i x_i = 1 \text{ and } \forall i x_i \geq 0\}.$$

Note that the vectors in this set are all positive and sum to 1 – this fact will be used several times in the sequel.Also define the map $f : S \to S$ as follows:

$$f(x) = \{\sum_i (Ax)_i\}^{-1} Ax.$$

Here $(Ax)_i$ is the i^{th} coordinate of Ax. If $x \in S$ then, because $A > 0$, the vector Ax is nonzero and the map f is well defined. After taking a course on real analysis, one can easily check that f is continuous and maps S into S. And it so happens that all the other conditions of Brouwer's Fixed Point Theorem are also satisfied (again, from real analysis). Therefore, by applying Brouwer's Theorem, there exists a $y \in S$ such that $f(y) = y$. (Note that from the way S is defined, $y \geq 0$. Thus,

$$\{\sum_i (Ay)_i\}^{-1} Ay = y.$$

If we set $\sum_i (Ay)_i = \lambda_0$, then $\lambda_0 > 0$ and $Ay = \lambda_0 y$. Since $A > 0$, it follows that $y > 0$. To show this explicitly, note that $Ay = \lambda_0 y$ is the same as

$$\sum_{j=1}^{n} a_{ij} y_j = \lambda_0 y_i$$

for every i. Since the $a_{ij} > 0$ and $y_j \geq 0$, then the left hand side above must be nonzero, since we have the condition from S's definition that $\sum_j y_j = 1$. Hence, y_i on the right hand side must be positive, since λ_0 is positive. Hence, (i) is proved.

If we apply (i) to A^T, then we conclude that $A^T z = \lambda_1 z$ for some $\lambda_1 > 0, z > 0$. Now,

$$\lambda_1 y^T z = y^T A^T z = \lambda_0 y^T z,$$

and since $y^T z > 0$, we have $\lambda_0 = \lambda_1$. Thus, $A^T z = \lambda_0 z$. This fact will be used for the rest of the proof. Let $Au = \mu u$ for some eigenvalue μ. Let μ^+ be defined by $\mu^+ = (|u_1|, |u_2|, \ldots, |u_n|)^T$, where $u = (u_1, \ldots, u_n)^T$. Without loss of generality, let u^+ be a probability vector. We have

$$\sum_j a_{ij} |u_j| \geq |\sum_j a_{ij} u_j| = |\mu u_i| = |\mu||u_i|$$

where the first inequality holds by the triangle inequality ($\sum |x_i| \geq |\sum x_i|$). Thus $Au^+ \geq |\mu| u^+$. Premultiply this last inequality by z^T to conclude that $\lambda_0 \geq |\mu|$:

$$\begin{aligned} z^T A u^+ &\geq |\mu| z^T u^+ \\ (A^T z)^T u^+ &\geq |\mu| z^T u^+ \\ \lambda_0 \underbrace{z^T u^+}_{=1} &\geq |\mu| \underbrace{z^T u^+}_{=1} \\ \lambda_0 &\geq |\mu|. \end{aligned}$$

We now prove (iii). First suppose $Av = \lambda_0 v$ for some real, nonzero vector v. We must show that v is a scalar multiple of y. Well, if v and y are linearly independent, then there exists a real number α such that $y - \alpha v$ is a nonnegative, nonzero vector with at least one zero coordinate. (Why? Homework.) Since

$$A(y - \alpha v) = \lambda_0 (y - \alpha v)$$

, $y - \alpha v$ is an eigenvector of A. However, since $A > 0$, any nonnegative eigenvector of A must in fact be positive, and we thus get a contradiction. Thus v is a scalar multiple of y. By considering the real and imaginary parts separately, we can show that any complex eigenvector of A corresponding to λ_0 is a scalar multiple of y.

We omit the proof of (iv) here, since we don't need it today. □

Now we use prove some facts about Markov matrices:

> **Some Facts About Markov Matrices**
>
> (5) There exists an eigenvector v with eigenvalue 1.
> (6) The modulus of every eigenvalue is less than or equal to 1.

Proof. The proof of (i) is given in the handout: "Since the sum of the elements in each column of P is 1, the sum of the elements in each column of the matrix $I - P$ is 0. Thus if we add the first $N - 1$ rows of $I - P$ to the last row, we get a matrix whose last row is 0. This means that the rank of $I - P$ is at most $N - 1$, which means that the equation $(P - I)v = 0$ must have at least one free variable, which means that the equation has a nonzero solution v. Of course such a v is an eigenvector of P with eigenvalue 1."

The proof of (ii) is as follows: Suppose we have a Markov matrix M. Define $A = M^T$. We will show that A satisfies property (ii). Then it follows that the original matrix M also satisfies this property, because the eigenvalues of a matrix and its transpose are the same. (Why? Homework.) Note that every row of A sums to 1, since every column of M summed to 1. Suppose $Ax = \lambda x$. Let k be such that $|x_j| \leq |x_k|$ for all $1 \leq j \leq n$. That is, k is the index of the entry with maximum modulus. Then, equating the k^{th} component of each side of equation yields

$$\sum_{j=1}^{n} a_{kj} x_j = \lambda x_k.$$

Hence

$$\begin{aligned} |\lambda x_k| &= |\lambda||x_k| = |\sum_{j=1}^{n} a_{kj} x_j| \leq \sum_{j=1}^{n} a_{kj}|x_j| \\ &= \sum_{j=1}^{n} a_{kj}|x_k| = |x_k| \end{aligned}$$

where the last equality holds from the row-sum property of A. Hence $|\lambda| \leq 1$. ☐

Finally we are ready to prove the result, by putting all the puzzle pieces together.

(1) From (5), we have "found" an eigenvector v of the Google matrix Q, and it has eigenvalue 1.
(2) From (6), v is an eigenvector with maximal eigenvalue.
(3) From (2) and (3), the eigenvector v is exactly the same as the eigenvector y stated in Equation (1) of the Perron-Frobenius Theorem.
(4) From (1), that eigenvector y is positive. So $v = y$ is positive.

Mini-lecture: Summing Polynomials – Why it works

Recall the following trick from last time:

> **Summing Polynomials**
>
> Assume the sequence $h_0, h_1, h_2, \ldots, h_n, \ldots$
> has a difference table whose 0th diagonal equals $c_0, c_1, c_2, \ldots, c_p, 0, 0, \ldots$
> Then $h_n = c_0\binom{n}{0} + c_1\binom{n}{1} + \ldots + c_p\binom{n}{p}$, and $\sum h_k = c_0\binom{n+1}{1} + c_1\binom{n+1}{2} + \ldots + c_p\binom{n+1}{p+1}$.

In this mini-lecture, we will prove and understand why this formula works.

Outline (following along in Brualdi):

- Zeros for all rows of the difference table beyond p (induction proof)
- Linearity property of differences
- th diagonal of a difference table determines whole table
- Consider the 0th diagonal of "0,0,0,0,1,0,0,..." and relate this to roots of a polynomial
- Apply linearity to get the closed form of polynomials in terms of binomial coefficients
- Finally, use the identity $\sum_{k=0}^{n}\binom{k}{p} = \binom{n+1}{p+1}$. (Why? Discuss this some day. Namely, I actually don't know right now.)

Homework 5

(1) "If v and y are linearly independent, then there exists a real number α such that $y - \alpha v$ is a nonnegative, nonzero vector with at least one zero coordinate. " Justify this statement.
(2) "... the eigenvalues of a matrix and its transpose are the same." Justify this statement.
(3) Work through problems in the Google handout. (It's not hard. The idea is to just get some practice with applying linear algebra practically.)
(4) Puzzle: What row of numbers comes next?

Homework 4 Solutions

(1) Problem 5.4.13 in Strang. (Cross product. See spiral bound notebook.)
   (a) Evaluate and confirm.
   (b) Notice that the derivative of the norm $\|u\|$ with respect to time is equal to 0.
   (c) Three lines of MATLAB:

```
>> syms a b c;
>> A = [ 0 c -b ; -c 0 a ; b -a 0 ];
>> eig(A)

ans =
                      0
  (-a^2-c^2-b^2)^(1/2)
 -(-a^2-c^2-b^2)^(1/2)
```

This is better written as $\pm i\|v\|$, where $v = (a, b, c)$. The matrix operation you were asked to compute corresponds to cross product of two vectors in $\mathbf{R}^3$. The resultant vector is in a direction perpendicular to both vectors. Hence, the eigenvalues are imaginary! We have to "rotate" the input vector to meet up with the output vector.