Problem set 6
The Perron Frobenius theorem.

Math 212a14

Oct 21 2014, Due Oct.28

In a future problem set I want to discuss some criteria which allow us to conclude that that the “ground state” of a self-adjoint operator is non-degenerate. In the physics literature this is known as “the non-degeneracy of the vacuum”.These criteria are extensions to the infinite dimensional setting of theorems about finite matrices which go under the name of Perron-Frobenius. Unfortunately, as far as I can tell, these important theorems are not taught in any course in the math department here at Harvard. So the purpose of this problem set is to familiarize you with the the Perron Frobenius theorem.

This problem set is taken (with some modifications) from a problem set in my Lie algebras course. What I do here is give my version of a proof of the classification of the“extended Dynkin diagrams”. From this classification it is immediate to obtain the classification of the “ordinary Dynkin diagrams” and then with considerable Lie algebra work the classification of all the simple finite dimensional Lie algebras, due to Killing. John Coleman calls Killing’s paper ”The Greatest Mathematical Paper of All Time.”

Contents

1 Graphs.

An **undirected graph** $\Gamma = (N, E)$ consists of a set N (for us finite) and a subset E of the set of subsets of N of cardinality two. We call elements of N "nodes" or "vertices" and the elements of E "edges". If $e = \{i, j\} \in E$ we say that the "edge" e joins the vertices i and j or that "i and j are adjacent". Notice that in this definition our edges are "undirected": $\{i, j\} = \{j, i\}$, and we do not allow self-loops. An example of a graph is the "cycle" $A_\ell^{(1)}$ with $\ell + 1$ vertices, so $N = \{0, 1, 2, \ldots, \ell\}$ with 0 adjacent to ℓ and to 1, with 1 adjacent to 0 and to 2 etc.

The **adjacency matrix** A of a graph Γ is the (symmetric) $0 - 1$ matrix whose rows and columns are indexed by the elements of N and whose i, j-th entry $A_{ij} = 1$ if i is adjacent to j and zero otherwise.

For example, the adjacency matrix of the graph $A_3^{(1)}$ (see the figure below) is

$$\begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}.$$

We can think of A as follows: Let V be the vector space with basis given by the nodes, so we can think of the i-th coordinate of a vector $x \in V$ as assigning the value x_i to the node i. Then $y = Ax$ assigns to i the sum of the values x_j summed over all nodes j adjacent to i.

As I hope to explain in class, the Hückel theory of organic chemistry reduces considerations of stability of organic molecules to that of finding the eigenvalues of the adjacency matrix of a graph.

A **path** of length r is a sequence of nodes $x_{i_1}, x_{i_2}, \ldots, x_{i_r}$ where each node is adjacent to the next. So, for example, the number of paths of length 2 joining i to j is the i, j-th entry in A^2 and similarly, the number of paths of length r joining i to j is the i, j-th entry in A^r. The graph is said to be **connected** if there is a path (of some length) joining every pair of vertices. In terms of the adjacency matrix, this means that for every i and j there is some r such that the i, j entry of A^r is non-zero. In terms of the theory of non-negative matrices (see below) this says that the matrix A is **irreducible**.

Notice that if $\mathbf{1}$ denotes the column vector all of whose entries are 1, then $\mathbf{1}$ is an eigenvector of the adjacency matrix of $A_\ell^{(1)}$, with eigenvalue 2, and all the entries of $\mathbf{1}$ are positive. In view of the Perron-Frobenius theorem to be stated below, this implies that 2 is the maximum eigenvalue of this matrix.

We modify the notion of the adjacency matrix as follows: We start with a connected graph Γ as before, but modify its adjacency matrix by replacing some of the ones that occur by positive integers a_{ij}. If, in this replacement $a_{ij} > 1$, we redraw the graph so that there is an arrow with a_{ij} lines pointing towards the node i. For example, the graph labeled $A_1^{(1)}$ in Table **Aff 1** corresponds to the matrix

$$\begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix}$$

which clearly has $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ as an positive eigenvector with eigenvalue 2.

Similarly, diagram $A_2^{(2)}$ in Table **Aff 2** corresponds to the matrix

$$\begin{pmatrix} 0 & 4 \\ 1 & 0 \end{pmatrix}$$

which has $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ as eigenvector with eigenvalue 2. In the diagrams, the coefficient next to a node gives the coordinates of the eigenvector with eigenvalue 2, and it is immediate to check from the diagram that this is indeed an eigenvector with eigenvalue 2. For example, the 2 next to a node with an arrow pointing toward it in $C_\ell^{(1)}$ satisfies $2 \cdot 2 = 2 \cdot 1 + 2$ etc.

All the graphs so far have zeros along the diagonal. If we relax this condition, and allow for any non-negative integer on the diagonal, then the only new possibilities are those given in Figure 4. Indeed, let us call a matrix **symmetrizable** if $A_{ij} \neq 0 \Rightarrow A_{ji} \neq 0$. The main content of this problem set will be to show that

the lists in the Figures 1-4 exhaust all irreducible matrices with non-negative integer matrices, which are symmetrizable and have maximum eigenvalue 2.

2 Perron-Frobenius.

We say that a real matrix T is **non-negative** (or **positive**) if all the entries of T are non-negative (or positive). We write $T \geq 0$ or $T > 0$. We will use these definitions primarily for square $(n \times n)$ matrices and for column vectors $= (n \times 1)$ matrices. We let

$$Q := \{x \in \mathbf{R}^n : x \geq 0, \quad x \neq 0\}$$

so Q is the non-negative "orthant" excluding the origin. Also let

$$C := \{x \geq 0 : \|x\| = 1\}.$$

So C is the intersection of the orthant with the unit sphere.

A non-negative matrix square T is called **primitive** if there is a k such that all the entries of T^k are positive. It is called **irreducible** if for any i, j there is a $k = k(i,j)$ such that $(T^k)_{ij} > 0$. For example, as mentioned above, the adjacency matrix of a connected graph is irreducible.

If T is irreducible then $I + T$ is primitive.

In this section we will assume that T is non-negative and irreducible.

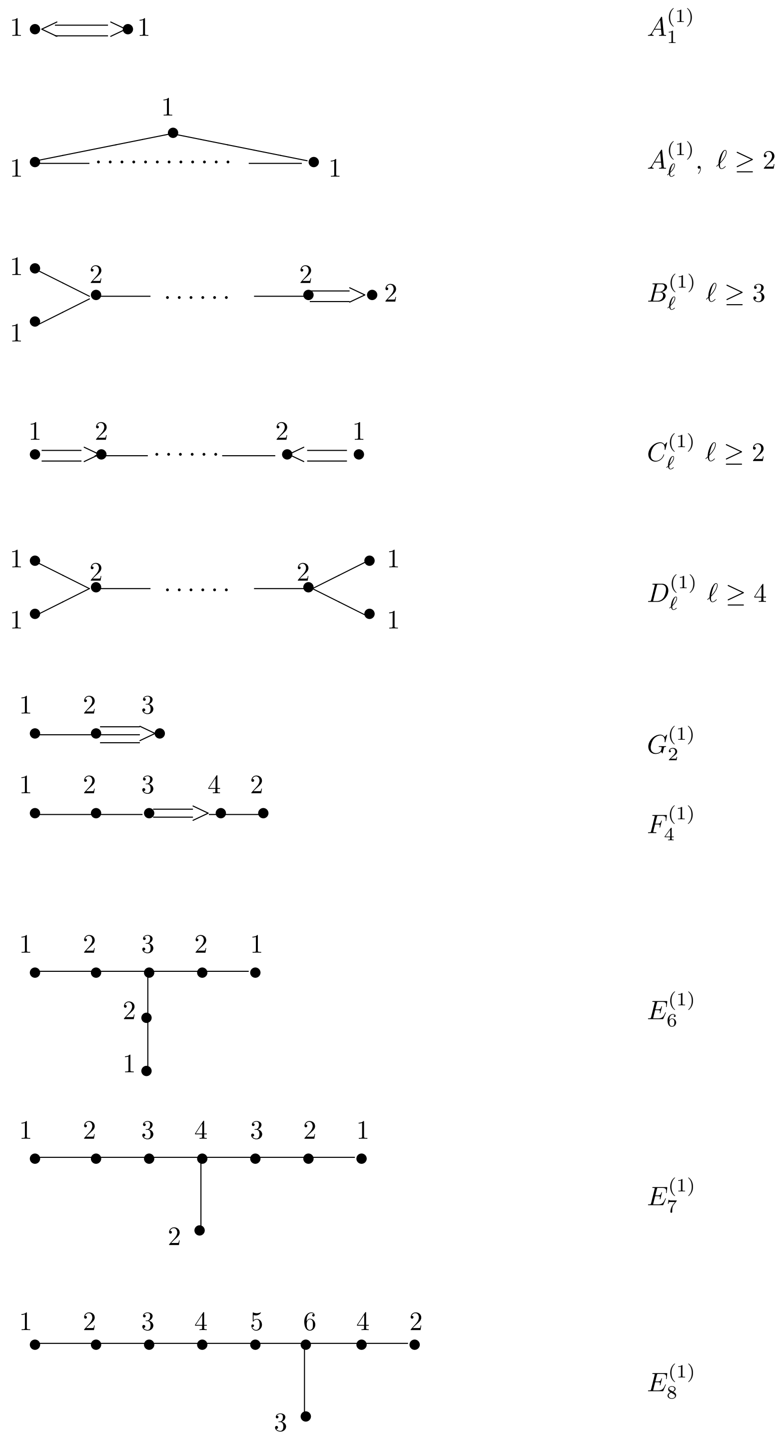
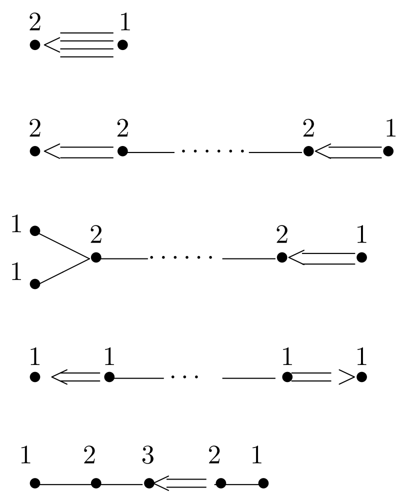


Figure 1: Aff 1.



$A_2^{(2)}$

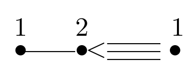
$A_{2\ell}^{(2)}\ \ \ell \geq 2$

$A_{2\ell-1}^{(2)}\ \ \ell \geq 3$

$D_{\ell+1}^{(2)},\ \ \ell \geq 2$

$E_6^{(2)}$

Figure 2: Aff 2



$D_4^{(3)}$

Figure 3: Aff 3

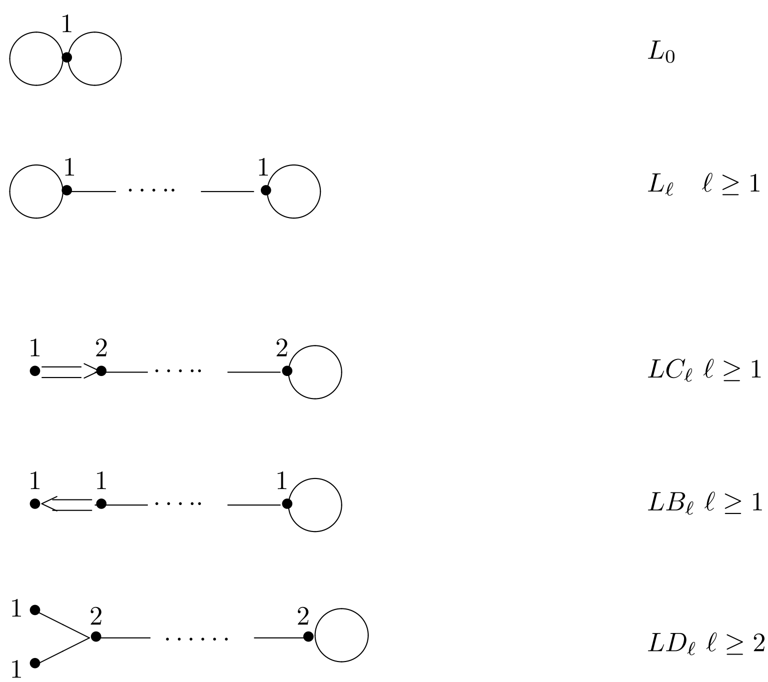


Figure 4: Loops allowed

Theorem 2.1. Perron-Frobenius.

1. *T has a positive (real) eigenvalue $\lambda_{\max}$ such that all other eigenvalues of T satisfy*
$$|\lambda| \leq \lambda_{\max}.$$
2. *Furthermore $\lambda_{\max}$ has algebraic and geometric multiplicity one, and has an eigenvector x with $x > 0$.*
3. *Any non-negative eigenvector is a multiple of x.*
4. *More generally, if $y \geq 0,\ y \neq 0$ is a vector and μ is a number such that*
$$Ty \leq \mu y$$
then
$$y > 0, \quad \text{and} \quad \mu \geq \lambda_{\max}$$
with $\mu = \lambda_{\max}$ if and only if y is a multiple of x.
5. *If $0 \leq S \leq T,\ S \neq T$ then every eigenvalue σ of S satisfies $|\sigma| < \lambda_{\max}$.*
6. *In particular, all the diagonal minors $T_{(i)}$ obtained from T by deleting the i-th row and column have eigenvalues all of which have absolute value $< \lambda_{\max}$.*

I will present a concise proof of this theorem (for those of you who have not seen it in a linear algebra course) later on in this handout. Let me make one clarification as to the last two assertions of the theorem. The matrix $T_{(i)}$ is usually thought of as an $(n-1) \times (n-1)$ matrix obtained by "striking out" the i-th row and column. But we can also consider the matrix T_i obtained from T by replacing the i-th row and column by all zeros. If x is and n-vector which is an eigenvector of T_i, then the $n-1$ vector y obtained from x by omitting the (0) i-th entry of x is then an eigenvector of $T_{(i)}$ with the same eigenvalue. Conversely, if y is an eigenvector of $T_{(i)}$ then inserting 0 at the i-th position will give an n-vector which is an eigenvector of T_i with with the same eigenvalue as that of y.

More generally, suppose that S is obtained from T by replacing a certain number of rows and the corresponding columns by all zeros. Then we may apply item 5) of the theorem to this $n \times n$ matrix, S, or the "compressed version" of S obtained by eliminating all these rows and columns.

We will want to apply this to the following special case. A subgraph Γ' of a graph Γ is the graph obtained by eliminating some nodes, and all edges emanating from these nodes. Thus, if A is the adjacency matrix of Γ and A' is the adjacency matrix of A, then A' is obtained from A by striking out some rows and their corresponding columns. Thus if Γ is irreducible, so that we may apply the Perron Frobenius theorem to A, and if Γ' is a proper subgraph (so we have actually deleted some rows and columns of A to obtain A'), then the maximum eigenvalue of A' is strictly less than the maximum eigenvalue of A. Similarly, if an entry A_{ij} is > 1, the matrix A' obtained from A by decreasing this entry while still keeping it positive will have a strictly smaller maximal eigenvalue.

3 Problems.

Using Perron-Frobenius, these problems will show that that the (generalized) graphs occurring in Figures 1 - 4 are the only irreducible graphs with maximal eigenvector 2.

Here is how to get started. We have already checked that $A_2^{(2)}$ has maximal eigenvector 2. Any graph B which has an entry ≥ 4 must contain $A_2^{(2)}$ as a "subgraph", by the condition $A_{ij} \neq 0 \Rightarrow A_{ji} \neq 0$. The graph $A_2^{(2)}$ can not be a proper subgraph of B, because this will imply that the maximal eigenvalue of B is strictly > 2. This implies that $A_2^{(2)}$ is the only graph whose matrix contains a 4. Similarly, $A_1^{(1)}$ is the only graph for which there is a pair i, j with *both* A_{ij} and $A_{ji} > 1$.

We have verified that $A_1^{\ell}, \ell \geq 2$ has 2 as maximal eigenvalue. This implies that no other graph with maximal eigenvalue 2 can contain a cycle, and that the straight line graph with no double or triple edges or branches has maximal eigenvector < 2 an so is not a candidate.

1. Write out the matrices corresponding to $G_2^{(1)}$ and to $D_4^{(3)}$ and verify that the given vectors are eigenvectors with eigenvalue 2. Conclude that these are the only graphs with a triple edge.

2. Write out the matrices corresponding to $F_4^{(1)}$ and $E_6^{(2)}$ and verify that the given vectors are eigenvectors with eigenvalue 2. Conclude that these are the only graphs with a double edge which does not terminate at an end point of the graph. In other words, any graph with maximal eigenvalue 2 and which has a double edge must have the property that one of the nodes of the double edge is not adjacent to any other node.

3. Write out the matrix corresponding to $D_4^{(1)}$. This is a graph with five nodes, a "star" with the four branches emanating from the center node and connecting to four end nodes. Verify that the given vector is an eigenvector with eigenvalue 2. Conclude that this is the only graph with maximal eigenvalue 2 which has four or more branches emanating from any node (and with no loops).

4. Verify that the maximal eigenvalue of $D_\ell^{(1)}$ is 2, and conclude that these are the only graphs with maximal eigenvalue 2 which has two or more branch points (that is a node with more than two adjacent nodes). Also conclude that a graph with a branch node connecting to two end nodes and no other branch nodes, loops, or double edges has maximal eigenvalue strictly less than 2.

5. Write out the matrix corresponding to $E_6^{(1)}$. Even without writing out the matrix, notice that the given vector is an eigenvector with eigenvalue 2, since the sum of the values at the adjacent nodes is twice the value at any given node.

This graph has three branches (emanating from the central node) and each of the branches has two nodes not counting the central node. Conclude that this is the only graph with maximal eigenvalue 2 with a branching node that has two or more nodes on each branch (not counting the branching node). So any other graph with a branch has at least one branch with only one additional node on it if the graph has maximal eigenvalue 2.

6. The graph $E_7^{(1)}$ has a branch nodes with one additional node on one of the branches and three additional node on the two others. Verify that $E_7^{(1)}$ has maximal eigenvalue 2, and conclude that any graph other than $E_6^{(1)}$ and $E_7^{(1)}$ with a branch node and with maximal eigenvalue 2, has at most one additional node on one of its branches and at most two additional nodes on a second of its branches.

7. Verify that $E_8^{(1)}$ has maximal eigenvalue 2, and conclude that outside of the three $E_\ell^{(1)}$ graphs, $\ell = 6, 7, 8$ any graph with maximal eigenvalue 2, and a branch point, must have two of the branches with only additional node on each, and the other end of the graph must be a double edge or a loop. Conclude that these possibilities are exhausted by $B_\ell^{(1)}, A_{2\ell-1}^{(2)}$ and LD_ℓ.

8. The matrix of L_0 is (2). So this is the only graph with maximal eigenvalue 2 with two (or more) loops emanating from a single node. Draw the graph LD_2 which has three nodes, write out its matrix. Verify that its maximal eigenvalue is 2. Conclude that any other graph with loops and with maximal eigenvalue 2 has its loop(s) emanating from an end node. The possibility of a loop at each end is accounted for by $L_\ell, \ell \geq 1$.

We are thus left to consider graphs with no branches and either double edges at both ends or a double edge at one end and a (single) loop at the other end. In the case of a double edge at both ends, there are there possibilities as to the placement of the arrows. They can both face in, as in $C_\ell^{(1)}$ they can both face out, as in $D_{\ell+1}^{(2)}$or one can face in and the other out as in $A_{2\ell}^{(2)}$. We have thus accounted for all graphs with maximal eigenvalue 2 and with double edges at both ends. If there is a double edge at one end and a loop at the other, there are two possibilities, the arrow can face in as in LC_ℓ or can face out as in LB_ℓ, and both these possibilities occur.

4 Proof of Perron-Frobenius.

Let

$$P := (I + T)^{n-1}$$

Recall that Q denotes the positive orthant minus $\{0\}$ and that C denotes the intersection of the unit sphere with the positive orthant. For any $z \in Q$ let

$$L(z) := \max\{s : sz \leq Tz\} = \min_{1 \leq i \leq n, z_i \neq 0} \frac{(Tz)_i}{z_i}.$$

By definition $L(rz) = L(z)$ for any $r > 0$, so $L(z)$ depends only on the ray through z. If $z \leq y,\ z \neq y$ we have $Pz < Py$. Also $PT = TP$. So if $sz \leq Tz$ then

$$sPz \leq PTz = TPz$$

so

$$L(Pz) \geq L(z).$$

Furthermore, if $L(z)z \neq Tz$ then $L(z)Pz < TPz$. So $L(Pz) > L(z)$ unless z is an eigenvector of T. Consider the image of C under P. It is compact (being the image of a compact set under a continuous map) and all of the elements of $P(C)$ have all their components strictly positive (since P is positive). Hence the function L is continuous on $P(C)$. Thus L achieves a maximum value, $L_{\max}$ on $P(C)$. Since $L(z) \leq L(Pz)$ this is in fact the maximum value of L on all of Q, and since $L(Pz) > L(z)$ unless z is an eigenvector of T, we conclude that $L_{\max}$ is achieved at an eigenvector, call it x of of T and $x > 0$ with $L_{\max}$ the eigenvalue. Since $Tx > 0$ and $Tx = L_{\max}x$ we have $L_{\max} > 0$.

We will now show that this is in fact the maximum eigenvalue in the sense of the theorem. So let y be any eigenvector with eigenvalue λ, and let $|y|$ denote the vector whose components are $|y_j|$, the absolute values of the components of y. We have $|y| \in Q$ and from

$$Ty = \lambda y$$

and the triangle inequality we conclude that

$$|\lambda||y| \leq T|y|.$$

Hence $|\lambda| \leq L(|y|) \leq L_{\max}$. So we may use the notation

$$\lambda_{\max} := L_{\max}$$

since we have proved that

$$|\lambda| \leq \lambda_{\max}.$$

We have proved item 1 in the theorem.

Suppose that $0 \leq S \leq T$. Then $sz \leq Sz$ and $Sz \leq Tz$ implies that $sz \leq Tz$ so $L_S(z) \leq L_T(z)$ for all z and hence

$$L_{\max}(S) \leq L_{\max}(T).$$

We may apply the same argument to $T^\dagger$ to conclude that it also has a positive maximum eigenvalue. Let us call it η. (We shall soon show that $\eta = \lambda_{\max}$.) This means that there is a vector $w > 0$ such that

$$w^\dagger T = \eta w.$$

Recall the $x > 0$ denotes the eigenvector with maximum eigenvalue $\lambda_{\max}$ of T. We have

$$w^\dagger T x = \eta w^\dagger x = \lambda_{\max} w^\dagger x$$

implying that $\eta = \lambda_{\max}$ since $w^\dagger x > 0$.

Now suppose that $y \in Q$ and $Ty \le \mu y$. Then

$$\lambda_{\max} w^\dagger y = w^\dagger T y \le \mu w^\dagger y$$

implying that $\lambda_{\max} \le \mu$, again using the fact that all the components of w are positive and some component of y is positive so $w^\dagger y > 0$. In particular, if $Ty = \mu y$ then then $\mu = \lambda_{\max}$.

Furthermore, if $y \in Q$ and $Ty \le \mu y$ then $\mu \ge 0$ and

$$0 < Py = (I + T)^{n-1} y \le (1 + \mu)^{n-1} y$$

so

$$y > 0.$$

This proves the first two assertions in item 4. If $\mu = \lambda_{\max}$ then $w^\dagger (Ty - \lambda_{\max} y) = 0$ but $Ty - \lambda_{\max} y \le 0$ and so $w^\dagger (Ty - \lambda_{\max} y) = 0$ implies that $Ty = \lambda_{\max} y$. Then the last assertion of item 4) will then follow from item 2) once we prove item 2), since we have shown that y must be an eigenvector with eigenvalue $\lambda_{\max}$.

Suppose that $0 \le S \le T$ and $Sz = \sigma z,\ z \ne 0$. Then

$$T|z| \ge S|z| \ge |\sigma||z|$$

so

$$|\sigma| \le L_{\max}(T) = \lambda_{\max},$$

as we have already seen. But if $|\sigma| = \lambda_{\max}$ then $L(|z|) = L_{\max}(T)$ so $|z| > 0$ and $|z|$ is also an eigenvector of T with the same eigenvalue. But then $(T - S)|z| = 0$ and this is impossible unless $S = T$ since $|z| > 0$. Replacing the i-th row and column of T by zeros give an $S \ge 0$ with $S < T$ since the irreducibility of T precludes all the entries in a row being zero. This proves items 5) and 6).

From linear algebra we know that

$$\frac{d}{d\lambda} \det(\lambda I - T) = \sum_i \det(\lambda I - T_{(i)})$$

and each of the matrices $\lambda_{\max} I - T_{(i)}$ has strictly positive determinant by what we have just proved. This shows that the derivative of the characteristic polynomial of T is not zero at $\lambda_{\max}$, and therefore the algebraic multiplicity and hence the geometric multiplicity of $\lambda_{\max}$ is one. This proves 2) and hence the remaining assertions. QED

5 Markov chains in a nutshell.

A non-negative matrix M is a **stochastic** matrix if each of the row sums equal 1. Then the column vector $\mathbf{1}$ all of whose entries equal 1 is an eigenvector with eigenvalue 1. So if M is irreducible 1 is the maximal eigenvalue since $\mathbf{1}$ has all positive entries.

If M is primitive, then we know from the general theory that

$$M^k \to \begin{pmatrix} \pi_1 & \pi_2 & \cdots & \pi_n \\ \pi_1 & \pi_2 & \cdots & \pi_n \\ \vdots & \vdots & \vdots & \vdots \\ \pi_1 & \pi_2 & \cdots & \pi_n \end{pmatrix}$$

where $\mathbf{p} := (\pi_1, \pi_2, \cdots, \pi_n)$ is the unique vector whose entries sum to one and satisfies $\mathbf{p}M = \mathbf{p}$.

6 One parameter semi-groups of positive matrices.

We are going to be interested (in the infinite dimensional case) in one parameter semi-groups of the form e^{-tA} and will want to know when this is non-negative in for all $t > 0$. If $A = (a)$ is a one-by-one matrix then $e^{-at} > 0$ no matter what a is. Suppose that an $n \times n$ matrix A with $n > 1$ has the property that e^{-tA} is non-negative for all $t > 0$. Then the matrix $I - e^{-tA}$ has non-positive off diagonal matrices. Dividing by $t > 0$ and letting $t \to 0$ shows that the off diagonal entries of A are non-positive.

If B is a matrix with *all* of its entries non-positive, then clearly

$$e^{-tB} = I - tB + \frac{1}{2}t^2B^2 - \cdots$$

has all of its entries non-negative.

9. Show that if the off-diagonal entries of A are non-positive then e^{-tA} has non-negative entries for all $t > 0$.[Hint: Use Lie's formula which says that

$$e^{-t(D+B)} = \lim_{n \to \infty} \left(e^{-\frac{t}{n}D} \cdot e^{-\frac{t}{n}D} \right)^n.$$

We will prove a usable infinite dimensional version of this later which is known as the Trotter product formula. At the moment, take Lie's formula for granted.]