CPSC 531F: SUPPLEMENTAL NOTES AND HOMEWORK

JOEL FRIEDMAN

Last revised: Monday 12th April, 2021, at 13:42

Contents

Research supported in part by an NSERC grant.

Disclaimer: The material may sketchy and/or contain errors, which I will elaborate upon and/or correct in class. For those not in CPSC 531F: use this material at your own risk...

1. Introduction

The main goal of this course is to use spectral information, i.e., information on the eigenvalues and eigenvectors of a matrix, to infer information about a "large" or "complex" system. Typical applications are:

(1) You are given a graph on a large number of vertices; at times the graph is "sparse" in the sense that each vertex has a "small" number of neighbours. You want to infer something about certain "global" properties of the graph than are not easy to infer, neither by a feasible computation nor by "local" considerations. However, you do know something about the largest few eigenvalues and corresponding eigenvalues of the adjacency matrix, and you'd like to use this "spectral information" to get some information about these global properties.
(2) You are given a (discrete) Markov chain on a large number of states. Same type of global questions. One typical global property is the "mixing time" of the chain.
(3) In some cases, you want to build a graph or network for the sake of running an algorithm, and you get to choose whichever graph or network you like.
(4) There are many special cases and related problems to the above (e.g., regular graphs, weighted graphs, directed graphs, reversible Markov chains).
(5) You are given a question that has no a priori connection to any particular matrix; at times you can use the spectral decomposition of a cleverly chosen matrix to answer the question. A typical example concerns certain questions about Boolean functions, where the eigenvalues/eigenvectors of the Boolean n-cube (often called "Fourier analysis" of the n-cube) is a powerful tool. Another situation is when you can think of a large data set as a matrix, where you'd like to compress the information using a low rank approximation via its SVD (singular-value decomposition).

1.1. References.

For Spring 2021, the main references are:

(1) *Matrix Analysis*, by Horn and Johnson; the first edition [HJ85] is currently available online at the UBC Library; the second edition [HJ13], which is significantly expanded in some sections, is currently available in print; there is also a corrected printing of the first edition [HJ90]. When we refer to this textbook, we will make sure to note any difference in section numbers between [HJ85, HJ13].
(2) *Markov Chains and Mixing Times*, by Levin and Peres, with contributions from Wilmer et al., [LP17], especially Chapter 4 on Markov chain mixing; at present there is a version of this available at https://pages.uoregon.edu/dlevin/MARKOV/markovmixing.pdf, and this is the version to which I will refer.
(3) "Expander Graphs and Their Applications," by Hoory, Linial, and Wigderson [HLW06], (this article received the 2008 AMS Conant Prize); at present there is a version of this available at https://www.cs.huji.ac.il/~nati/PAPERS/expander_survey.pdf.

(4) Other references?

Here are some other notable references for applications:

(1) "Graph spectra in Computer Science," by Cvetković and Simić, https://www.sciencedirect.com/science/article/pii/S0024379510006117?via%3Dihub, "Selected Topics in Applications of Graph Spectra," Editors Cvetkovic and Gutman, 2009, http://www.mi.sanu.ac.rs/novi_sajt/research/projects/ZbR14-22.pdf and many other articles and books of Cvetković and co-authors, including
(2) Others?
(3)

1.2. **Calculus Motivation and Main Theorems.** In this subsection we introduce way symmetric matrices arise in calculus, namely to determine when a function $f\colon \mathbb{R}^2 \to \mathbb{R}$ has a local minimum, local maximum, or a saddle. When then state the fundamental theorem about symmetric matrices, as well as Sylvester's "Law of Inertia" to show how these theorems generalize to function $f\colon \mathbb{R}^n \to \mathbb{R}$. matrices.

In Spring 2021 we use the textbook *Matrix Analysis* by Horn and Johnson for results on linear algebra and matrices. You should look over Sections 0.0–0.6 to review the ideas we need.

1.2.1. *Local Maxima, Minima, Saddles, Etc.* Taylor's theorem says that if $f\colon \mathbb{R} \to \mathbb{R}$ is twice differentiable, then for any $x_0 \in \mathbb{R}$ we have

$$f(x) = f(x_0) + (x - x_0)f'(x_0) + \frac{(x-x_0)^2}{2} f''(x_0) + o(x - x_0)^2.$$

It follows that if $f'(x_0) = 0$, then f has a local min at x_0 if $f''(x_0) > 0$, and a local min there if $f''(x_0) < 0$. The analog for $f\colon \mathbb{R}^n \to \mathbb{R}$ and $\mathbf{x}_0 \in \mathbb{R}^n$ is that

$$f(\mathbf{x}) = f(\mathbf{x}_0) + (\mathbf{x}-\mathbf{x}_0)\cdot\big((\nabla f)(\mathbf{x}_0)\big) + \frac{1}{2}(\mathbf{x}-\mathbf{x}_0)^{\mathrm{T}}\big((D^2 f)(\mathbf{x}_0)\big)(\mathbf{x}-\mathbf{x}_0) + o|\mathbf{x}-\mathbf{x}_0|^2,$$

where $D^2 f$, sometimes called the *Hessian of* f is the matrix of second partial derivatives: for $n = 2$ this is

$$D^2 f \stackrel{\text{def}}{=} \begin{bmatrix} f_{x_1x_1} & f_{x_1x_2} \\ f_{x_2x_1} & f_{x_2x_2} \end{bmatrix}$$

The fact that this matrix is symmetric, i.e., that $f_{x_2x_1} = f_{x_1x_2}$ and hence $D^2 f$ equals its transpose, gives a remarkable way to understand the local behaviour of f when $\nabla f(\mathbf{x}_0) = 0$. For simplicity, we state these results in the case where $\mathbf{x}_0 = \mathbf{0}$ and $f(\mathbf{x}_0) = 0$.

Theorem 1.1. *Consider a* 2×2 *real symmetric matrix, i.e., a matrix of the form*

$$\begin{bmatrix} a & b \\ b & c \end{bmatrix}.$$

If $ac - b^2 \neq 0$*, then the function*

$$f(\mathbf{x}) = f(x_1, x_2) = \mathbf{x}^{\mathrm{T}} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \mathbf{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = ax_1^2 + 2bx_1x_2 + cx_2^2$$

has (1) a local minimum at $\mathbf{x} = \mathbf{0}$ *if* $ac - b^2 > 0$ *and* $a > 0$ *(which implies* $c > 0$*), (2) a local maximum at* $\mathbf{x} = \mathbf{0}$ *if* $ac - b^2 > 0$ *and* $a < 0$ *(which implies* $c < 0$*), (3) a* saddle *at* $\mathbf{x} = \mathbf{0}$ *if* $ac - b^2 < 0$.

The case where $ac - b^2 = 0$, i.e., the determinant of the above 2×2 matrix is zero, i.e., the above matrix is not invertible, is a degenerate case (where higher derivatives may determine the local behaviour of f).

A stronger theorem says that regardless of the value of $ac - b^2$, there is an "orthonormal change of basis," which here means a rotation of coordinates in $\mathbb{R}^2$, which turns the above matrix into a diagonal 2×2 matrix, i.e., of the form

$$\begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix}.$$

We now describe that generalize the above theorems from two variables to any number of variables.

1.2.2. *The Main Theorem on Symmetric Matrices.* This subsection can be skipped for now. However, I encourage the reader to have a look and see how much of it makes sense now—some students may have already seen this theorem in a second term course on linear algebra. After the next few sections this should make complete sense.

The main goal of the next few sections is to give several proofs of the following fundamental theorem on symmetric matrices, and to explain its applications to graphs, Markov chains, and a host of applications that arise in the SVD (singular-value decomposition). In the sections that follow we will review what these terms mean.

This comes from the theorem below. At this point we will use the notation in *Matrix Analysis* by Horn and Johnson (either edition): e.g., $M_{m,n}(\mathbb{R})$ denotes the space of $m \times n$ matrices with real entries and $M_n(\mathbb{R}) = M_{n,n}(\mathbb{R})$; the notation $\mathbb{R}^{m \times n}$ is also very common in the literature. In this notation we generally assume that $m, n \in \mathbb{Z} = \{1, 2, \ldots\}$, although at times it is convenient to all $m, n = 0$[1]

Theorem 1.2. *Let $A \in M_n(\mathbb{R}) = M_{n,n}(\mathbb{R})$ (also commonly written $A \in \mathbb{R}^{n \times n}$) with $A^{\mathrm{T}} = A$, i.e., A is an $n \times n$, real symmetric matrix A (hence $n \in \mathbb{N} = \{1, 2, \ldots\}$). Then the following hold (these are all equivalent):*

1. *A can be written as a real, diagonal matrix after an orthonormal change of basis; this basis is unique (up to $\pm$) when the diagonal entries are distinct.*
2. *There are real numbers $\lambda_1(A) \geq \cdots \geq \lambda_n(A)$ and orthonormal vectors $v_1, \ldots, v_n$ such that $Av_i = \lambda_i v_i$; the sequence $\lambda_1, \ldots, \lambda_n$ is uniquely determined.*
3. *A can be written as MDM^{-1}, where M is an orthogonal matrix (i.e., $M^{-1} = M^{\mathrm{T}}$, or equivalently the rows of M are orthonormal, or equivalently, the columns of M are orthonormal), where D is a diagonal matrix with real entries:*

$$D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}.$$

[1] One of my first college teachers, George Mackey, used to say that in every classroom there is always an expert on the empty set. At times it can be useful to be an expert on the empty set (i.e., the empty union, the empty sieve, etc.). We leave it to the reader to verify that $\mathbb{R}^0 = \{0\}$ and to interpret the meaning of $M_{m,n}(\mathbb{R})$ in case $m = 0$ and/or $n = 0$.

If we order $\lambda_1, \dots, \lambda_n$ as $\lambda_1 \geq \cdots \geq \lambda_n$, then the sequence $\lambda_1, \dots, \lambda_n$ is uniquely determined, and u_i is uniquely determined (up to $\pm$) if λ_i occurs "with multiplicity one."

Here is a sample application; it is extremely important to understand this in an intuitive way.

Corollary 1.3. *Let $Q\colon \mathbb{R}^n \to \mathbb{R}$ be a quadratic form, i.e., a function of the form $Q(\mathbf{x}) = \mathbf{x}^{\mathrm{T}} B \mathbf{x}$ for a matrix $B \in M_n(\mathbb{R})$. Let $A = (B + B^{\mathrm{T}})/2$, so that A is a (and the only) symmetric matrix with $Q(\mathbf{x}) = \mathbf{x}^{\mathrm{T}} B \mathbf{x} = \mathbf{x}^{\mathrm{T}} A \mathbf{x}$. Let $\lambda_1, \dots, \lambda_n$ be the eigenvalues of A, and $\mathbf{u}_1, \dots, \mathbf{u}_n$ a corresponding orthonormal basis of eigenvectors (corresponding means $A\mathbf{u}_i = \lambda_i \mathbf{u}_i$ for all i), then*

$$Q(c_1 \mathbf{u}_1 + \cdots + c_n \mathbf{u}_n) = \lambda_1 c_1^2 + \lambda_2 c_2^2 + \cdots + \lambda_n c_n^2.$$

Intuitively this corollary says that there is a change of coordinates with respect to which Q is of the form

$$\lambda_1 c_1^2 + \lambda_2 c_2^2 + \cdots + \lambda_n c_n^2.$$

One can also make this precise: if we define

$$\tilde{Q}(c_1, \dots, c_n) \stackrel{\text{def}}{=} Q(c_1 \mathbf{u}_1 + \cdots + c_n \mathbf{u}_n),$$

Then $\tilde{Q}$ is also a quadratic form, and represents Q in the "coordinate system $\mathbf{u}_1, \dots, \mathbf{u}_n$," and this "coordinate system" is comprised of orthonormal vectors (this orthonormality is extremely important in applications). The above corollary then states that

$$\tilde{Q}(c_1, \dots, c_n) = \lambda_1 c_1^2 + \lambda_2 c_2^2 + \cdots + \lambda_n c_n^2,$$

which makes it easy to see if $\tilde{Q}$ has a local maximum or minimum (or something "in between") at $\mathbf{0}$; but we easily see that $\tilde{Q}$ has a local maximum or minimum iff Q does.

1.2.3. *The LDU-Factorization and Sylvester's "Law of Inertia".* Another set of results, which are less important to these notes, allows one to quickly check how many of the eigenvalues of a symmetric matrix of moderate size are positive, negative, and zero. These results also imply that if a twice continuously differentiable function $f\colon \mathbb{R}^3 \to \mathbb{R}$ has a critical point at $\mathbf{x}_0 \in \mathbb{R}^3$, i.e., $\nabla f(\mathbf{x}_0) = 0$, then f has a local maximum (respectively minimum) if the three matrices

$$\left[f_{x_1 x_1}\right], \quad \begin{bmatrix} f_{x_1 x_1} & f_{x_1 x_2} \\ f_{x_2 x_1} & f x_2 x_2 \end{bmatrix}, \quad \begin{bmatrix} f_{x_1 x_1} & f_{x_1 x_2} & f_{x_1 x_3} \\ f_{x_2 x_1} & f_{x_2 x_2} & f_{x_2 x_3} \\ f_{x_3 x_1} & f_{x_3 x_2} & f_{x_3 x_3} \end{bmatrix}, \tag{1}$$

each have negative (respectively, positive) determinant at $\mathbf{x}_0$. This is the analog of the two variable criterion.

Let us briefly describe the main theorems.

The basic result is called Sylvester's "Law of Inertia," which says that if A is a symmetric matrix, and for some invertible matrix, S, we set $B = SAS^{\mathrm{T}}$, then (B is also symmetric) and A, B have the same number of positive eigenvalues, negative eignevalue, and zero eigenvalues.

Definition 1.4. We say that a symmetric $A \in M_n(\mathbb{R})$, is *positive definite* if $f(\mathbf{x}) = \mathbf{x}^{\mathrm{T}} A \mathbf{x}$ satisfies $f(\mathbf{x}) \geq 0$ with equality iff $\mathbf{x} = 0$. One similarly defines *negative definite*.

Hence a twice continuously differentiable function $f\colon \mathbb{R}^n \to \mathbb{R}$ with a critical point at $\mathbf{x}_0 \in \mathbb{R}$, i.e., $\nabla f(\mathbf{x}_0) = 0$, has a local maximum at $\mathbf{x}_0$ iff $D^2 f(\mathbf{x}_0)$ is negative definite, and a local minimum iff $D^2 f(\mathbf{x}_0)$ is positive definite. Also A is positive definite iff $-A$ is negative definite.

This gives a quick way to check if A is positive definite, at least when A's dimesion is of small enough size to allow for a reasonable quick and accurate Gaussian elimination. Here are the details.

Gaussian elimination on an invertible square matrix $A \in M_n(\mathbb{R})$ essentially writes A as a product $PLDU$, where P is a permutation matrix (indicating when we need to exchange rows to get a non-zero pivot element for the elimination), L is lower triangular with 1's on the diagonal, D is diagonal, and U is upper triangular with 1's on the diagonal. If A is symmetric and positive definite, then by induction on n one seems that one doesn't need the permutation matrix, and that $U = L^{\mathrm{T}}$; hence Guassian elimination essentially factors A as $A = LDL^{\mathrm{T}}$. Sylverster's law of inertia then implies that D must have positive entries; conversely, if D is any diagonal matrix with positive entries, then $A = LDL^{\mathrm{T}}$ for any lower triangular L with 1's on the diagonal—or any invertible L—is positive definite.

Moreover, A has an LDU decomposition as above, and if D's diagonal entries are $d_1, \ldots, d_n$, then for any $k \in [n]$, the product $d_1 \ldots d_k$ is just the determinant of the upper left $k \times k$ minor of A. This explains the condition regarding (1).

2. Basic Notation, Graphs, and Markov Chains

In this section we give some basic notation about matrices, graphs, Markov chains, give some examples explaining our interest in symmetric matrices. One of the surprising facts is that any $m \times n$ real matrix—not necessarily a square matrix–is best approximated by low rank matrices using the theory of symmetric matrices.

In Spring 2021 we use the textbook *Matrix Analysis* by Horn and Johnson for results on linear algebra and matrices. You should look over Sections 0.0–0.6 to review the ideas we need.

2.1. Basic Notation for Vectors and Matrices.

We typically work with the vector spaces $\mathbb{R}^n$ (over the reals, $\mathbb{R}$) and $\mathbb{C}^n$ (over the complex numbers $\mathbb{C}$). We use the notation $\mathbf{v} = (v_1, \ldots, v_n)$ for the elements of $\mathbb{R}^n$ or $\mathbb{C}^n$, i.e., boldface for the vector (e.g., $\mathbf{v}$) and regular typeface for its components (e.g., $v_1, \ldots, v_n$). [([HJ85, HJ13]uses regular typeface for both.] We will also use $(v_1, \ldots, v_n)$ to denote the $n \times 1$ column vector whose entries are v_i, and at times the $1 \times n$ row vector when the meaning is clear.

Our notation for an $m \times n$ matrix is usually (see [HJ], Section 0.3)

$$A = [a_{ij}] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

for $A \in M_{m,n}(\mathbb{R}) = \mathbb{R}^{m \times n}$, which gives rise to a linear map $\mathbb{R}^n \to \mathbb{R}^m$ via the map $\mathbf{v} \mapsto A\mathbf{v}$. We will also write a_{ij} as $(A)_{ij}$ at times. With these conventions the usual

convention to multiply $A \in M_{m,n}(\mathbb{R})$ by $B \in M_{n,p}$ is given by the formula

$$(AB)_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk}$$

(for all $i \in [m] = \{1, \ldots, m\}$ and all $k \in [p]$).

We warn the reader that one uses the MAP ON ROW VECTORS for directed graphs and Markov chains. (This can cause no end of confusion, especially for reversible Markov chains...) More precisely, an $A \in M_{m,n}(\mathbb{R}) = \mathbb{R}^{m \times n}$ can be viewed as the map $\mathbb{R}^m \to \mathbb{R}^n$ given by $\mathbf{u}^{\mathrm{T}} \mapsto \mathbf{u}^{\mathrm{T}} A$ as a map of row vectors, which is the same thing as the map $\mathbf{u} \mapsto A^{\mathrm{T}} \mathbf{u}$ as column vectors. We shall now see examples of this in directed graphs and Markov chains.

2.2. **The Standard Basis and Indicator (or Characteristic) Vectors.** When $\mathbb{R}^n$ or $\mathbb{C}^n$ is understood, for $i \in [n]$ we use $\mathbf{e}_i$ to denote the i-th standard basis vector, i.e., the vector that is everywhere 0 except 1 in the i-th component ([HJ85, HJ13], Section 0.1.7). Notice that for $A \in M_{m,n}(\mathbb{R}) = \mathbb{R}^{m \times n}$, $\mathbf{e}_i^T A \mathbf{e}_j^T$ equals a_i (here $\mathbf{e}_i$ lies in $\mathbb{R}^m$ or $\mathbb{C}^m$). If $I \subset [n]$, we use

$$\mathbf{e}_I = \sum_{i \in I} \mathbf{e}_i$$

which lies in $\mathbb{R}^n$ or $\mathbb{C}^n$; hence $\mathbf{e}_I$ is 1 in coordinates of I and 0 elsewhere, which is sometimes called the *indicator (or characteristic) vector of* I. It follows that for $I \subset [m]$, $J \subset [n]$

$$\mathbf{e}_I^{\mathrm{T}} A \mathbf{e}_J = \sum_{i \in I, j \in J} a_{ij}.$$

2.3. **Directed Graphs.**

Definition 2.1. A *directed graph* is a 4-tuple $G = (V_G, E_G, t_G, h_G)$ where V_G and E_G are sets—the *vertex set* and *edge set*—and t_G, h_G are maps $E_G \to V_G$—the *tails* and *heads* map. One (usually) defines the *adjacency matrix* of G to be the matrix $A = A_G$ indexed on the vertex set, i.e., $A \in M_n(\mathbb{R}) = M_{n,n}(\mathbb{R}) = \mathbb{R}^{n \times n}$ where $n = |V_G|$, whose entry a_{ij} are given as the number of edges from vertex i to vertex j (i.e., whose tail is vertex i and head is vertex j). A *walk* in G is an alternating sequence of vertices and edges,

$$w = (v_0, e_1, v_1, e_2, v_2, \ldots, v_{k-1}, e_k, v_k) \tag{2}$$

for some integer $k \geq 0$—called the *length* of w—such that for all $i \in [k] = \{1, \ldots, k\}$ we have $t_G(e_i) = v_{i-1}$ and $h_G(e_i) = v_i$; w is *closed* if $v_0 = v_k$.

We will use common graph theoretic terminology (without rigorously defining each term in these notes): for example, the walk w above *originates* in v_0 and *terminates* in v_k; also, if $t_G(e) = v$ and $h_G(e) = w$, then we say that e is an edge *from* v *to* w; most of the terminology should be intuitive.

It is easy to see that for any integer $k \geq 2$, the ij-th entry of A_G^k (the product of A_G k times) is the number of *walks* from i to j. It is immediate that if $I, J \subset [n]$, then $\mathbf{e}_I^{\mathrm{T}} A_G \mathbf{e}_J$ is the number of edges from a vertex in I to a vertex in J (where we identify $[n]$ with V).

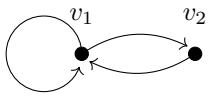

FIGURE 1. The Fibonacci (Directed) Graph

Example 2.2. The Fibonacci graph (see Figure 1) has adjacency matrix

$$A_{\text{Fib}} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix};$$

it has two vertices, v_1, v_2, and (1) a "self-loop" from v_1 to itself, (2) one edge from v_1 to v_2, and (3) one from v_2 to v_1. It turns out that (see Exercise **??**) for any $k \in \mathbb{Z}$ we have

$$A_{\text{Fib}}^k = \begin{bmatrix} F_{k+1} & F_k \\ F_k & F_{k-1} \end{bmatrix} \tag{3}$$

where $F_0, F_1, F_2, \ldots$ are the *Fibonacci numbers*, i.e., $F_0 = 0$, $F_1 = 1$, and $F_{k+2} = F_{k+1} + F_k$ for all k (hence $F_k = F_{k+2} - F_{k-1}$ defines this sequence inductively for k negative); hence

$$\ldots, F_{-7} = 13, F_{-6} = -8, F_{-5} = 5, F_{-4} = -3, F_{-3} = 2, F_{-2} = -1, F_{-1} = 1,$$
$$F_0 = 0, F_1 = 1, F_2 = 1, F_3 = 2, F_4 = 3, F_5 = 5, F_6 = 8, F_7 = 13, F_8 = 21, F_9 = 34, F_{10} = 55, \ldots$$

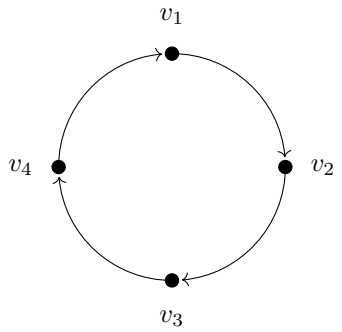

FIGURE 2. The (Directed) Cycle of Length 4

Example 2.3. The cycle graph of length 4 (see Figure 2) has adjacency matrix

$$C_4 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

Notice that the graph indicates motion from v_1 to v_2 to v_3 etc., and indeed $(C_4)_{ij} = 1$ iff $j = i + 1$ (modulo 4). And notice that

$$\begin{bmatrix} \alpha & \beta & \gamma & \delta \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} \delta & \alpha & \beta & \gamma \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} b \\ c \\ d \\ a \end{bmatrix}.$$

Hence, if you think of this directed graph as indicating motion from v_1 to v_2 to v_3 etc., then you need C_4 to act ROW VECTORS, NOT COLUMN VECTORS.

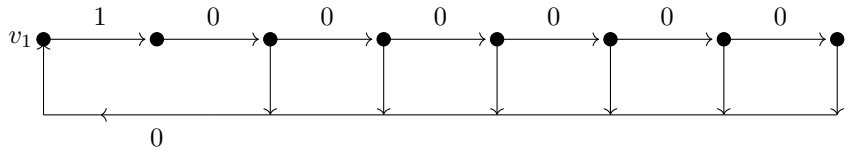

FIGURE 3. The $(2,7)$-Constrained Binary Data Graph

Example 2.4. The $(2,7)$-constrained binary data graph (see Figure 3) has adjacency matrix that is an 8×8 matrix (that we won't bother to write down). The set of walks of any length k beginning at v_1 corresponds to words over $\{0,1\}$ that begin in 1 and have between 2 and 7 0's between an two occurrences of 1's; the labels above the edges indicate this correspondence between walks and words over $\{0,1\}$. We will motivate and study similarly "(d,k) run-length constrained data" (here $d=2$ and $k=7$) in Subsection **??**.

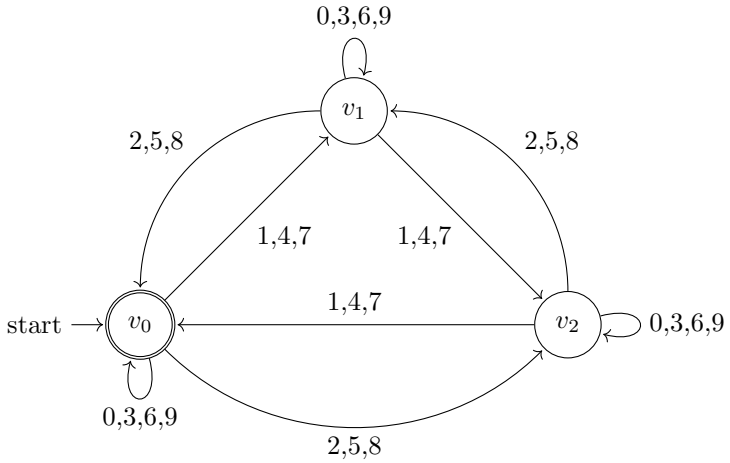

FIGURE 4. The DFA Describing Strings "Divisible By 3."

Example 2.5. Any DFA, i.e., (deterministic) finite automaton, can be viewed as a directed graph with some extra information: (1) one vertex is called the *initial state*, (2) there is a subset of vertices called the set of *final states* or *accepting states*, and (3) each edge comes with a label from a finite alphabet, Σ, such that each vertex has exactly one outgoing edge (i.e., whose tail is this vertex) labelled with each letter in Σ. The set of walks of length k in the DFA that begin with the initial vertex corresponding to the words of length k in Σ.

In Figure 4 we give a DFA for the strings over $\Sigma = \{0,1,\ldots,9\}$ that represent integers divisible by 3 (v_0 is the initial state and the unique accepting state). [This DFA accepts the empty string and allows for leading 0's.]

2.4. **(Undirected) Graphs.** An *undirected graph*, or simply a *graph*, is essentially a graph when each directed edge is paired with another edge in the opposite direction.

Definition 2.6. A *graph* (or *undirected graph* for emphasis) is a tuple $G = (V_G, E_G^{\text{dir}}, t_G, h_G, \iota_G)$ where $(V_G, E_G^{\text{dir}}, t_G, h_G)$ is a directed graph—called the *underlying directed graph of* G—and ι_G is an orientation reversing involution of E_G^{dir}, i.e., $\iota_G \colon E_G \to E_G$ is a bijection with $\iota_G = \iota_G^{-1}$ and $t_G \iota_G = h_G$ (and therefore also $h_G \iota_G = t_G$).

The notion of a *walk* in a graph, G, and its *adjacency matrix*, A_G, is simply that notion in the underlying directed graph. Therefore if G is a graph, A_G is a symmetric matrix (and does not depend on the choice of ι_G).

Notice that if G has multiple edges, i.e., edges with the same heads and tails, or two *self-loops about a vertex* $v \in V_G$, i.e., two $e \in E_G^{\text{dir}}$ with $t_G(e) = h_G(e) = v$, then there is more than one choice of ι_G. Indeed, if e is a self-loop, we may have $\iota_G e = e$, in which case we say that e is a *half-loop*, and otherwise a *whole-loop*.

Definition 2.7. If $G = (V_G, E_G^{\text{dir}}, t_G, h_G, \iota_G)$ is a graph, we define its *edge set*, denoted E_G, to be the set of "orbits of ι_G," i.e., consisting of two-element sets $\{e, \iota_G e\}$ for $e \in E_G^{\text{dir}}$ with $\iota_G e \neq e$, and one-element sets $\{e\}$ for each half-loop e.

It follows that for $v \in V_G$, the diagonal entry $(A_G)_{vv}$ equals twice the number of whole-loops plus one times the number of half-loops about v. Similarly, $|E_G^{\text{dir}}|$ equals the number of half-loops plus twice the number of edges that are not half-loops.

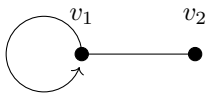

FIGURE 5. The Undirected Fibonacci Graph

Example 2.8. The Fibonacci (directed graph) of Example 2.2 is a graph whose edge set consists of one half-loop at v_1 and one edge joining v_1 and v_2. We depict it as a graph by replacing the two directed edges between v_1 and v_2 with a line joining v_1 and v_2 without an arrow; we leave the half-loop with an arrow.

Example 2.9. The (directed) cycle of length 4 in Example 2.3 is not a graph, but one can turn it into a graph by adding an oppositely oriented directed edge. We depict it as in Figure 6; its adjacency matrix is

$$A = C_4 + C_4^{-1} = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}.$$

2.5. **Sample Problems about Digraphs and Graphs where Eigenvalues/vectors are Sometimes Useful.** In some of the examples above, especially Figure 3 of Example 2.4 (I'll explain in class, or see [ACH83] or later papers [AFKM86, HMS91, BBM+10]), one wants to know for $A \in M_n(\mathbb{R})$ what is the rough value of $f(k) = f_{ij}(k) = (A^k)_{ij}$ for some fixed values of $i, j \in [n]$ and $k \to \infty$.

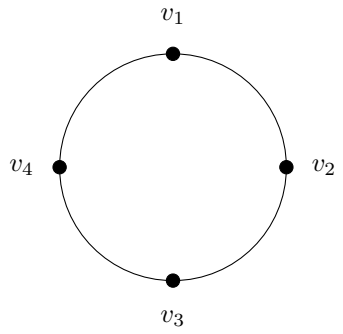

FIGURE 6. The (Undirected) Cycle of Length 4

Question 2.10. Given a graph and its adjacency matrix $A = A_G \in \mathcal{M}_n(\mathbb{R})$, for fixed $i, j \in [n]$, describe $f = f_{ij} \colon \mathbb{N} \to \mathbb{R}$ given by $f_{ij}(k) = (A_G^k)_{ij}$, i.e., the number of walks of length k from i to j. Either describe $f(k)$ exactly or a good approximation to f.

If A has all positive entries, then for each i, j there is a $c_{ij} > 0$ such that $f_{ij}(k) = c_{ij}\lambda_1^k(1 + o(1))$ as $k \to \infty$, where λ_1 is the largest, real eigenvalue of A, known as the *Perron-Frobenius eigenvalue* of A.

Question 2.11. Given a graph and its adjacency matrix $A_G \in \mathcal{M}_n(\mathbb{R})$, and given $I, J \subset [n]$ (we identify $[n]$ with V_G), describe (either exactly or approximately)

$$|E(I, J)| = e_I^{\mathrm{T}} A_G e_J.$$

For a given $a, b \in [n]$, bound (from above and below) the value of $|E(I, J)|$ for all I, J with $|I| = a$ and $|J| = b$.

For example, a graph with a "cluster" is a vertex set $I \subset [n]$ such that if $I^{\mathrm{comp}} = [n] \setminus I$ (identifying V_G with $[n]$), then $|E(I, I^{\mathrm{comp}})|$ is "unusually small." A graph has *good expansion (in one sense of this term)* if $|E(I, I^{\mathrm{comp}})|$ is "roughly what we would expect from a random graph."

[If there are no edges whose head and tail lie in I, then I is called an *independent set* (of vertices of the graph). One expects some independent set, I, for graph with many vertices and, say, each vertex having a fixed, small degree; however one does not expect to have a large independent set. A *bipartite graph* is one where for some I, all edges have one endpoint in I and one in I^{comp}; in this case I and I^{comp} are both independent sets and $|E(I, I^{\mathrm{comp}})|$ is, roughly speaking, "unusually large."]

It turns out that eigenvalues/vectors can give both the exact and approximate values of $f(k)$ and $|E(I, J)|$ above. In practical applications with n large, often the exact values are too costly to compute; in theoretical applications, at times it is too difficult to prove theorems about all the eigenvalues/vectors of a graph. At times the approximate values we get from eigenvalues/vectors solves our problem satisfactorily, at times not.

2.6. **Markov Chains.** At this point we introduce some basic notation for Markov chains, following [LP17], which is very standard notation.

We say that $\mathbf{p} = (p_1, \dots, p_n) \in \mathbb{R}^n$ is *stochastic* if $p_i \geq 0$ for all $i \in [n]$, and $p_1 + \cdots + p_n = 1$; intuitively such a $\mathbf{p}$ can represent the probability in being in

one of n "states of a system." Again, we caution the reader that in the theory of Markov chain, we typically use this when referring to ROW VECTORS, not COLUMN VECTORS.

By a *(row) stochastic matrix* ([HJ85, HJ13], Section 8.2) or sometimes *Markov matrix* we mean a $P \in M_n(\mathbb{R})$ each of whose rows is stochastic.

One can define a (finite, discrete time) Markov chain to be a pair $(\mathcal{X}, P)$ consisting of a finite *set of states*, $\mathcal{X} = \{s_1, \ldots, s_n\}$, and a stochastic matrix, $P = [p_{ij}] \in M_n(\mathbb{R})$, where $n = |\mathcal{X}|$. The interpretation is that p_{ij} represents the probability of transitioning from state s_i to state s_j. Again, P acts TO THE RIGHT OF ROW VECTORS, unless you define p_{ij} as the probability of transitioning from state j to state i...

Example 2.12. Say that everyone's favourite TV show (in a certain fixed population of people) is either The Expanse (s_1) or The Mandalorian (s_2). Each month you determine that 1% of the people who prefer The Expanse change their mind, and prefer The Mandalorian, and each month 2% change their mind and prefer The Expanse. The matrix

$$P = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} = \begin{bmatrix} .99 & .01 \\ .02 & .98 \end{bmatrix}$$

is a way of organizing the constants p_{ij} represents the probability moving from state i to state j. If, at month 0, 90% of the population prefer The Expanse, and 10% The Mandalorian, then at time $t \in \mathbb{N}$ the number of people in either state equals

$$\begin{bmatrix} 0.9 & 0.1 \end{bmatrix} \begin{bmatrix} .99 & .01 \\ .02 & .98 \end{bmatrix}^t.$$

Remark 2.13. Here is the "fine print." The textbook [LP17] has a similar 2 state Markov chain example, with a frog jumping between two lilly pads in Chapter 1. However, the notation in Chapter 1 there is more elaborate: a Markov chain is a sequence of random variables $X_0, X_1, \ldots$ that take values in $\mathcal{X} = \{s_1, \ldots, s_n\}$ (with $n = 2$ for a frog jumping between two lilly pads or people preferring The Expanse to The Mandalorian at various itmes); the Markov condition is that for all $t = 0, 1, \ldots$, $\mathbf{P}[X_{t+1} = s_j] = \mathbf{P}[X_t = s_i]\, p_{ij}$ (where $\mathbf{P}$ stands for "probability," and where $P = [p_{ij}]$ is a Markov matrix, i.e., a row stochastic matrix). Working with a sequence of random variables $X_0, X_1, \ldots$ is extremely useful—e.g., to give a formula for the stationary distribution in terms of the expected return time—and crucial for discussing "stopping times" and related notations; "coupling methods" use another sequence $Y_0, Y_1, \ldots$ on the same space, and again it is crucial to work with the X_t and Y_t rather than the row vectors of probabilities of the events at a given time.

To rigorously explain what you mean by a "sequence of random variables $X_0, X_1, \ldots$" you want an underlying *probability space*, $(\Omega, \mathbf{P})$, where Ω is a set, and $\mathbf{P}$ is a probability measure; then $X_0, X_1, \ldots$ are (measurable) random variables $X_t \colon \Omega \to \mathcal{X} = \{s_1, \ldots, s_n\}$. Once you've take a course that discusses measure theory—or not—you can ignore the underlying foundations and think of the (measurable) random variables $X_0, X_1, \ldots$. Often you can even ignore the $X_0, X_1, \ldots$, and just consider the stochastic row vectors of probabilities; for example, in the above example,

$$\begin{bmatrix} \mathbf{P}(X_t = s_1) & \mathbf{P}(X_t = s_2) \end{bmatrix} = \begin{bmatrix} 0.9 & 0.1 \end{bmatrix} \begin{bmatrix} .99 & .01 \\ .02 & .98 \end{bmatrix}^t.$$

If you are interested in only finite time $t = 0, 1, \ldots, T$ for some fixed T and finite Markov chain, then you avoid measure theory and build $(\Omega, \mathbf{P})$ as a finite probability space; e.g., Ω can be taken to be the set of functions X from $\{0, \ldots, T\}$ to $\{s_1, s_2\}$, where $X(t) = s_1$ represents the people who prefer The Expanse at time t. Otherwise Ω need to be at least as large (generally speaking) as the set of functions X from $\{0, 1, \ldots\}$ to $\{s_1, s_2\}$; the probability measure $\mathbf{P}$ is only defined on certain "measurable" subsets of Ω; if you have another set of random variables $Y_0, Y_1, \ldots$ defined

on Ω, then Ω has to be larger. In this subsection we view Markov chains as Markov matrices, and—at least for now—ignore the X_t.

Remark 2.14. If G is a graph, i.e., an undirected graph, then A_G is symmetric and so the difference between the meaning of a_{ij} (versus a_{ji}) can be ignored. However, if P is a *reversible* Markov chain—which is an analog of an undirected graph—then $p_{ij} \neq p_{ji}$ and the difference cannot be ignored. (Of course, if $p_{ij} = p_{ji}$—a condition that occurs more rarely—then P is symmetric.)

2.7. **Sample Problems about Markov Chains where Eigenvalues/vectors are Sometimes Useful.** As in digraphs and graphs, if $P \in \mathcal{M}_n(\mathbb{R})$ is a Markov matrix, many questions about Markov chains concern the exact or, more likely, approximate value of P^t for $t \in \mathbb{Z}$ and t large.

A specific example is the notion of the *mixing time* of a Markov chain.

Definition 2.15. We say that a Markov matrix, $P \in \mathcal{M}_n(\mathbb{R})$ (or an associated Markov chain) is *irreducible* if for any $i, j \in [n]$, for some $t \in \mathbb{N}$ we have $(P^t)_{ij} > 0$. A *stationary distribution* for P (e.g., [LP17], Section 1.5) is a stochastic vector, $\boldsymbol{\pi} \in \mathbb{R}^n$, such that $\boldsymbol{\pi}^{\mathrm{T}} P = \boldsymbol{\pi}^{\mathrm{T}}$.

It is useful to associate to each Markov matrix, $P \in \mathcal{M}_n(\mathbb{R})$, its *underlying directed graph*, which is the graph with vertex set $[n]$ and an edge from i to j iff $p_{ij} > 0$. Then $(P^t)_{ij} > 0$ iff there is a walk from i to j of length t in the underlying directed graph. Furthermore P is irreducible iff the underlying digraph is *strongly connected*, meaning that for any i, j there is some path from i to j (in a digraph, *weakly connected* means that there is either a path from i to j or a path from j to i).

Theorem 2.16. *If $P \in \mathcal{M}_n(\mathbb{R})$ is an irreducible Markov matrix, then it has a unique stationary distribution.*

This follows from the Perron-Frobenius theorem. Alternatively, one easily shows that a stationary distribution, $\boldsymbol{\pi}$, can be described by fixing any $j \in [n]$ and setting

$$\tilde{\pi}(i) \stackrel{\text{def}}{=} \mathbf{E}_{s_j}[\text{number of visits to } s_i \text{ before the first return to } s_j], \tag{4}$$

where $\mathbf{E}_{s_j}$ is the expected value of the Markov chain under the condition $\mathbf{P}[X_0 = s_j] = 1$, i.e., $\mathbf{p}_0 = \mathbf{e}_j$, the j-th standard basis vector. One then scales $\tilde{\pi}$ to get a stochastic vector $\boldsymbol{\pi}$. (One then proves that $\boldsymbol{\pi}$ is unique; see [LP17], Subsection 1.5.4.) For this reason you really want to the probability distributions:

$$\mathbf{P}[X_0 = i_0, \dots, X_t = i_t]$$

and you will want to take $t \to \infty$; to define what (4) really means, morally you are defining a "stopping time" τ_{s_j}, of a process that begins at s_j at time 0 (i.e., where $\mathbf{P}[X_0 = s_j] = 1$, i.e., $\mathbf{p}_0 = \mathbf{e}_j$, the j-th standard basis vector).

Definition 2.17. The *mixing time* of a Markov matrix, $P \in \mathcal{M}_n(\mathbb{R})$, etc. ([LP17]).

Question 2.18. Given a Markov matrix, $P \in \mathcal{M}_n(\mathbb{R})$, find its mixing time.

Here is a question of a "hidden Markov chain."

Question 2.19. Say that $P \in \mathcal{M}_n(\mathbb{R})$ is unknown, but we can observe the approximate (or exact) values of $(P^t)_{ij}$. How well can we find P approximately or exactly?

2.8. **Reversible Markov Chains.** See [LP17], Section 1.6.

Definition 2.20. Let $P \in M_n(\mathbb{R}) = M_{n,n}(\mathbb{R}) = \mathbb{R}^{n\times n}$ be the transition matrix of an irreducible Markov chain, and $\boldsymbol{\pi}$ its unique stationary distribution. Then P (or the Markov chain) is *reversible* if any of the following equivalent conditions hold:

(1) For all $i, j \in [n]$, $\pi_i p_{ij} = \pi_j p_{ji}$.
(2) For all $i, j \in [n]$ and $k \in \mathbb{N}$, $\pi_i (P^k)_{ij} = \pi_j (P^k)_{ji}$.

To understand why such a P is called *reversible*, we need to view the Markov chain as a sequence of random variables $X_0, X_1, \ldots$ on some probability space, where each X_t takes its values in the n states of the Markov chain; when we write $P \in \mathcal{M}_n(\mathbb{R})$, we are identifying $[n]$ with the states of the Markov chain. (This is discussed in [LP17] and will be explained in class.)

The condition that $\{X_t\}_{t=0,1,\ldots}$ is a Markov chain is that

$$\text{Prob}[X_{t+1} = j] = \text{Prob}[X_t = i]\, p_{ij}$$

for all $t = 0, 1, \ldots$ and $i, j \in [n]$. Under this condition, setting

$$\mathbf{p}_t^{\mathrm{T}} = [\text{Prob}(X_t) = 1 \ \cdots \ \text{Prob}(X_t) = n],$$

we have

$$\mathbf{p}_t^{\mathrm{T}} = \mathbf{p}_0^{\mathrm{T}} P^t$$

for any $t \geq 0$. It follows that for any t and $i_0, \ldots, i_t \in [n]$,

$$\text{Prob}[X_0 = i_0, \ldots, X_t = i_t]$$

depends only on $\mathbf{p}_0$. If $\boldsymbol{\mu} \in \mathbb{R}^n$ is any stochastic vector, we use $\text{Prob}_{\boldsymbol{\mu}}$ to denote these probabilities when $\mathbf{p}_0 = \boldsymbol{\mu}$.

It then follows that a Markov chain is reversible iff for any $t \geq 0$ and $i_0, \ldots, i_t \in [n]$, we have

$$\text{Prob}_{\boldsymbol{\pi}}[X_0 = i_0, X_1 = i_1, \ldots X_t = i_t] = \text{Prob}_{\boldsymbol{\pi}}[X_0 = i_t, X_1 = i_{t-1}, \ldots X_t = i_0].$$

It turns out that any reversible Markov matrix P has real eigenvalue, and that P can be viewed as a symmetric matrix in two different senses. To understand this most simply, we should understand the theory of self-adjoint matrices; we'll do this later.

2.9. **The Higman-Sims Technique.** Let $A \in \mathcal{M}_n(\mathbb{R})$ be symmetric and written in block form as

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1k} \\ A_{21} & A_{22} & \cdots & A_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ A_{k1} & A_{k2} & \cdots & A_{kk} \end{bmatrix},$$

and let $B \in \mathcal{M}_k(\mathbb{R})$ be the matrix $B = [b_{ij}]$ where b_{ij} is the average row sum of A_{ij}. Then—like reversible Markov chains—the eigenvalues of B are real, but B is not generally symmetric. And like reversible Markov chains, the point is that B is "symmetric with respect to the inner product weighted by the dimensions of the blocks." Again, this is easier to understand in terms of self-adjoint operators with respect to an inner product.

It will turn out that the eigenvalues of B *interlace* with those of A, in a certain sense, exploited by [HH71][2], which has been known as the "Higman-Sims technique." For example, if $A = A_G$ is the adjacency matrix of a graph, and $U \subset V_G$ is a subset of vertices of G, then writing A_G in the 2×2 blocks indexed by U and its complement, one has $b_{11} = 0$ iff U is an independent set (called a *coclique*, see [Hae78], [Hae80][3]).

The above theorem also works when $A \in \mathcal{M}_n(\mathbb{C})$ is Hermitian, i.e., $A^{\mathrm{H}} = A$.

2.10. **Other Places Where Symmetric Matrices Arise.**

2.10.1. *Least Squares and Symmetric Matrices.* Given n data points,

$$(x_1, y_1), \dots, (x_n, y_n) \in \mathbb{R}^2,$$

there are a, b, c such that

$$E(a,b,c) \stackrel{\text{def}}{=} \sum_{i=1}^{n} \bigl(y_i - a - bx_i - cx_i^2\bigr)^2$$

is minimized, given by the "normal equations"

$$\begin{bmatrix} n & \sum_i x_i & \sum_i x_i^2 \\ \sum_i x_i & \sum_i x_i^2 & \sum_i x_i^3 \\ \sum_i x_i^2 & \sum_i x_i^3 & \sum_i x_i^4 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} \sum_i y_i \\ \sum_i x_i y_i \\ \sum_i x_i^2 y_i \end{bmatrix}$$

More generally, to project $\mathbf{y}$ onto the space spanned by $\mathbf{u}_1, \dots, \mathbf{u}_n$, the projection is given by $a_1\mathbf{u}_1 + \cdots + a_n\mathbf{u}_n$ where

$$\begin{bmatrix} \mathbf{u}_1 \cdot \mathbf{u}_1 & \cdots & \mathbf{u}_1 \cdot \mathbf{u}_n \\ \vdots & \ddots & \vdots \\ \mathbf{u}_n \cdot \mathbf{u}_n & \cdots & \mathbf{u}_n \cdot \mathbf{u}_n \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} \mathbf{u}_1 \cdot \mathbf{y} \\ \vdots \\ \mathbf{u}_n \cdot \mathbf{y} \end{bmatrix}$$

2.10.2. *Other Symmetric Matrices: Variance-Covariance, Discrete Approximation of Laplacian, Etc.* See Section 7.0 of Horn and Johnson.

2.10.3. *Data Compression and SVD (Singular-Value Decomposition): Every Matrix can be "Approximated" using Symmetric Matrices.* The SVD (Singular-Value Decomposition) gives you the best low rank approximations to any real $m \times n$ matrix, where we measure approximation, i.e., distance between two matrices, with the Frobenius norm, i.e., sum of squares of entries.

We will discuss this type of approximation in more detail later. For now we may look at low rank approximations to:

(1) a two-dimensional picture, such as the classic clown picture (whose underlying data can be found in MATLAB) whose pixel values are a matrix which one seeks to approximate (see, e.g., Ascher-Greif, page 232), (or https://sites.math.washington.edu/~morrow/498_13/demmelsvd.pdf, pages 115, 116),

[2] As of January 2021, this article did not seem available on the internet or through the UBC Library; however, anyone with a UBC CWL can access its Mathematical Reviews summary (e.g., via searching on "Hestenes and Higman"), or, if you are connected to the UBC VPN or on the UBC network, via this link, and more specifically this Math. Reviews summary.

[3] Currently available at https://pure.tue.nl/ws/files/1721530/41103.pdf

(2) The "eigenfaces" example (the current Wikipedia page is reasonable), where the pixels values of a each of large number of face pics become the rows of the matrix.
(3) More generally, one wants to compress any m-points of $\mathbb{R}^n$ into a low rank approximation.

Cases (2),(3) create a matrix that you want to approximate by a low rank approximation, whereas (1) is done assuming that a 2-dim array of pixel values can be well-approximated by low rank approximations.

Theorem 2.21. *Let $A \in M_{m,n}(\mathbb{R}) = \mathbb{R}^{m\times n}$, and $q = \min(m,n)$. There are $\sigma_1(A) \geq \cdots \sigma_q(A) \geq 0$ and orthonormal $u_1, \ldots, u_q$ and $v_1, \ldots, v_q$ such that for any k,*

$$\sum_{i=1}^{k} u_i^{\mathrm{T}} \sigma_i v_i$$

is the best rank-k approximation to A in the Frobenius norm (i.e., square root of sum of squares of all entries). In more detail, $\sigma_i(A)$ is the square root of the i-th eigenvalue in decreasing order (i.e., $\sigma_1(A)$ is the largest) of both AA^{T} and $A^{\mathrm{T}}A$, and u_i, v_i are corresponding right eigenvectors of these matrices. (And σ_i, u_i, v_i are unique, the latter two up to $\pm$, whenever σ_i has multiplicity one of AA^{T} and/or $A^{\mathrm{T}}A$.)

Remark 2.22. If $A = A^{\mathrm{T}}$ is symmetric, then $AA^{\mathrm{T}} = A^{\mathrm{T}}A = A^2$, and hence $\sigma_i(A)$ is just the i-th largest absolute value of all eigenvalues of A.

Remark 2.23. There are algorithms to determine the top k top singular values of a matrix, which work especially well when the $\sigma_1, \ldots, \sigma_k$ are "reasonably seperated" and there is large gap between σ_k and σ_{k+1}. This idea occurs in many applications, especially in expanding graphs.

Remark 2.24. In the case of m-points, $\mathbf{x}_1, \ldots, \mathbf{x}_m \in \mathbb{R}^n$, the SVD determines, for any k, the minimum $\mathbf{x}'_1, \ldots, \mathbf{x}'_m \in \mathbb{R}^k$ and linear function $L\colon \mathbb{R}^k \to \mathbb{R}^n$ such that

$$E = E(L, \mathbf{x}'_1, \ldots, \mathbf{x}'_n) = \sum_{i=1}^{n} \|\mathbf{x}_i - L\mathbf{x}'_i\|_2^2$$

is as small as possible. However, you may be more interested in minimizing

$$E = E_p(L, \mathbf{x}'_1, \ldots, \mathbf{x}'_n) = \max_{i\in[n]} \|\mathbf{x}_i - L\mathbf{x}'_i\|_p,$$

i.e., the "worst-case L^p-norm" error; this can be solved for $p = 1, \infty$ by (more laboriously) solving a linear program. The advantage of the SVD is that it is easier to solve, especially when k is small and you are interested in only the first k terms of the SVD.

2.10.4. *Closing Remarks: Positive Semidefinite, Graph Laplacians, Inertia, Complex Matrices.*

(1) Many of the above matrices are postive semidefinite. (Give examples on homework?)
(2) Define graph Laplacian: $\Delta = D - A$ given by incidence matrix equation $\Delta = M_{v\to e}^{\mathrm{T}} M_{v\to e}$, which is therefore positive semidefinite.

(3) The reason why computing the signs of $f_{x_1x_1}$ and $f_{x_1x_1}f_{x_2x_2} - f_{x_1x_2}^2$ is sufficient to determine positive definite versus negative definite versus neither is due to "Sylver's Law of Inertia," which says that for any square A, B, M with A, B symmetric, $A = MBM^{\mathrm{T}}$, and M invertible, A, B have the same number of positive, negative, and zero eigenvectors. (This is related to the importance of an "$A = LDU$ decomposition" of matrices, where if A is symmetric, then $U = L^{\mathrm{T}}$.)

(4) Much of this theory holds over the complex numbers. Here A^{T} is replaced with A^{H}, the conjugate transpose, and the dot product is replaced with

$$\langle \mathbf{u}, \mathbf{v} \rangle \stackrel{\text{def}}{=} \mathbf{v}^{\mathrm{H}}\mathbf{u} = \sum_i u_i \overline{v_i}$$

(which is linear in $\mathbf{u}$ but conjugate linear in $\mathbf{v}$, i.e., $\langle \mathbf{u}, \alpha\mathbf{v} \rangle = \overline{\alpha}$ (for a scalar α, i.e., $\alpha \in \mathbb{C}$). For example, $A \in M_n(\mathbb{C}) = \mathbb{C}^{n\times n}$ is *Hermitian* if $A^{\mathrm{H}} = A$, in which case A has real eigenvalues and a corresponding orthonormal basis (where orthonrmality is defined wrt $\langle \cdot, \cdot \rangle$).

(5) Warning: Just as often in the literature one defines

$$\langle u, v \rangle \stackrel{\text{def}}{=} \sum_i \overline{u_i} v_i = \overline{\mathbf{v}^{\mathrm{H}}\mathbf{u}} = \mathbf{u}^{\mathrm{H}}\mathbf{v}.$$

This difference can be confusing.

(6) Warning: Usually a matrix operates on a column vector, so $A \in M_{m,n}(\mathbb{F}) = \mathbb{F}^{m\times n}$ is a map $\mathbb{F}^n \to \mathbb{F}^n$. However, in Markov chains, one typically sets $P = \{p_{ij}\}$ to be the matrix where p_{ij} is the probability of transitioning from state i to state j. In this case P acts (to the right of) row vectors. This can be confusing.

3. Eigenvalues and the Perron-Frobenius Eigenvalue

In this section we review the ideas behind eigenvalues and prove the Perron-Frobenius theorem regarding non-negative matrices that is behind almost everything we do with the adjacency matrices and Markov matrices. (See [HJ85, HJ13], Chapters 1 and 8.)

The motivation from the previous section is that $A \in M_n(\mathbb{R})$ and we want to approximate A^k for k large; however, the of eigenvalues/vectors of A can tell us a lot more about A. For applications to graphs and Markov chains A will have non-negative entries, which will make things easier to understand in terms of the

3.1. **Similarity.** We say that $A, B \in M_n(\mathbb{R})$ are *similar* if $A = S^{-1}BS$ for some invertible $S \in M_n(\mathbb{R})$. In this case we have

$$A^2 = (S^{-1}BS)(S^{-1}BS) = S^{-1}BSS^{-1}BS = S^{-1}B^2S,$$
$$A^3 = (S^{-1}BS)(S^{-1}BS)(S^{-1}BS) = S^{-1}B^3S,$$

and similarly $A^k = S^{-1}B^kS$ for any $k \in \mathbb{N}$.

It is helpful to think of these kinds of formulas as "stories." For example, the formula $(AB)^{-1} = B^{-1}A^{-1}$ can be viewed as "putting on socks first, and then shoes" but "taking off shoes first, then socks." (Notice that if A is "putting on socks," then A is acting on row vectors, whereas if B is "putting on socks," then B is acting on column vectors.) Similarly the formula

$$(S^{-1}B_kS)\dots(S^{-1}B_2S)(S^{-1}B_1S) = S^{-1}B_k\dots B_2B_1S$$

for any B_i and invertible S of the same dimension can be thought of as saying—acting on column vectors—S is "go to the coffee shop (or office)," B_1 is "solve first math problem," B_2 is "solve second math problem," etc.

3.2. **Diagonalization.** To *diagonalize* A is to write it as $S^{-1}DS$, where D is a diagonal matrix, i.e.,

$$D = \begin{bmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_n \end{bmatrix} \tag{5}$$

(in our usual convention we would write d_{ii} instead of d_i, but we usually omit the second i for matrix understood to be diagonal). The advantange of diagonal matrices is that we easily see that for any $k \in \mathbb{N}$,

$$D^k \begin{bmatrix} d_1^k & 0 & \dots & 0 \\ 0 & d_2^k & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_n^k \end{bmatrix}$$

and so

$$A^k = S^{-1} \begin{bmatrix} d_1^k & 0 & \dots & 0 \\ 0 & d_2^k & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_n^k \end{bmatrix} S,$$

which allows us to more easy "understand" A^k. (Diagonalization is viewed is a way of "decoupling" an $n \times n$ matrix, an $n \times n$ system of equations, an $n \times n$ ODE, etc.)

Similarly, if $p(x) = c_0 + c_1 x + \ldots + c_s x^s$ is any polynomial (with real or complex c_i), we define

$$p(A) = c_0 I + c_1 A + \ldots + c_s A^s$$

for any $A \in M_n(\mathbb{R})$, where $I = I_n$ is the $n \times n$ identity matrix. If $A = S^{-1}BS$, then we similarly see that

$$p(A) = S^{-1} p(B) S,$$

and if $B = D$ is diagonal as above, then

$$p(D) = \begin{bmatrix} p(d_1) & 0 & \dots & 0 \\ 0 & p(d_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & p(d_n). \end{bmatrix}$$

You may likely have also seen the matrix power series

$$e^{At} = I + At + \frac{(At)^2}{2} + \frac{(At)^3}{3!} + \cdots$$

in a previous class. The point is that typical ODE's with constant coefficients can be written in vector/matrix form as $(d/dt)\mathbf{y}(t) = A\mathbf{y}(t)$, whose solution is given by

$\mathbf{y}(t) = e^{At}\mathbf{y}(0)$. In this case if $A = S^{-1}DS$, then

$$e^{At} = S^{-1}\begin{bmatrix} e^{d_1 t} & 0 & \dots & 0 \\ 0 & e^{d_n t} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & e^{d_n t}. \end{bmatrix} S$$

Returning to the problem of "understanding" A^k for large k, if $A = S^{-1}DS$ with $D = \mathrm{diag}(d_1, \dots, d_n)$—this is shorthand for (5)—if $|d_1|$ is larger than all $|d_i|$ with $i \geq 2$, then we can write

$$A^k = d^k S^{-1}\begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & (d_2/d_1)^k & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & (d_n/d_1)^k \end{bmatrix} S,$$

which means that for large k

$$A^k = d^k S^{-1}\begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & o(1) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & o(1) \end{bmatrix} S,$$

where $o(1)$ is shorthand for any function of k that tends to 0 as $k \to \infty$. One can equivalently write

$$\lim_{k\to\infty} A^k/d_1^k = S^{-1}\begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & o(1) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & o(1) \end{bmatrix} S,$$

where the limit—i.e., convergence—means entry-by-entry.

3.3. **Diagonalization: Eigenvalues and Eigenvectors.** Say we can write $A = SDS^{-1}$ for matrices $A, S, D \in M_n(\mathbb{R})$ with S invertible and D diagonal, $D = \mathrm{diag}(d_1, \dots, d_n)$. In this case $AS = DS$, and it follows that if $\mathbf{v}_i$ is the i-th column of S, then $A\mathbf{v}_i = d_i\mathbf{v}_i$.

Definition 3.1. Say that $A \in M_n(\mathbb{C})$. By an *eigenvalue of* A we mean a $\lambda \in \mathbb{C}$ such that $A\mathbf{v} = \lambda\mathbf{v}$ for some non-zero $\mathbf{v} \in \mathbb{C}^n$; we call $\mathbf{v}$ a *eigenvector of* A *(corresponding to the eigenvalue* λ*)*; we also call $\mathbf{v}$ and λ an *eigenvector/value pair* or simply an *eigenpair* of A.

Definition 3.2. We say that $A \in M_n(\mathbb{C})$ is *diagonalizable* if there is a basis $\mathbf{v}_1, \dots, \mathbf{v}_n$ of $\mathbb{R}^n$ such that for each i, there is a $\lambda_i \in \mathbb{C}$ such that $A\mathbf{v}_i = \lambda_i\mathbf{v}_i$. We call $\mathbf{v}_1, \dots, \mathbf{v}_n$ an *eigenbasis* of A.

We will now give a number of examples.

3.4. **The Characteristic Polynomial.** For any $A \in \mathcal{M}_n(\mathbb{R})$ or $\mathcal{M}_n(\mathbb{C})$, we define the *characteristic polynomial of* A (Section 1.2 [HJ85, HJ13])

$$p_A(t) = \det(tI - A),$$

which is a polynomial of degree n (sometimes one defines this as $\det(A - tI) = (-1)^n \det(tI - A)$).

Theorem 3.3. *An $A \in \mathcal{M}_n(\mathbb{C})$ has λ as an eigenvalue iff $p_A(\lambda) = 0$. Moreover $A\mathbf{v} = \lambda\mathbf{v}$ iff $\mathbf{v}$ belongs to the nullspace (in the usual sense, i.e., the right nullspace) of $\lambda I - A$ or of $A - \lambda I$.*

We similarly define $\mathbf{v} \neq \mathbf{0}$ to be a *left eigenvector* of A if $\mathbf{v}^{\mathrm{T}}A = \lambda\mathbf{v}^{\mathrm{T}}$ for some $\lambda \in \mathbb{C}$ which we call the corresponding eigenvalue; one can equivalently write $A^{\mathrm{T}}\mathbf{v} = \lambda\mathbf{v}$. Since

$$\det(tI - A) = \det\big((tI - A)^{\mathrm{T}}\big) = \det\big(tI - A^{\mathrm{T}}\big),$$

we have that the eigenvalues of A are exactly those (with multiplicity) of A^{T}.

3.5. **The Trace and Determinant in the Characteristic Polynomial.** We have

$$p_A(t) = t^n + r_1 t^{n-1} + \ldots + r_{n-1}t + r_n,$$

so that $p_A(t)$ has n roots, $\lambda_1, \ldots, \lambda_n$; since

$$(t - \lambda_1)\cdots(t - \lambda_n) = p_A(t) = t^n + r_1 t^{n-1} + \ldots + r_{n-1}t + r_n,$$

we have

$$-r_1 = \sum_{i=1}^{n} \lambda_i, \quad r_2 = \sum_{i<j} \lambda_i\lambda_j, \ \ldots \ (-1)^n r_n = \lambda_1\lambda_2\ldots\lambda_n.$$

Since $p_A(0) = \det(-A)$, we see that

$$\lambda_1\lambda_2\ldots\lambda_n = \det(A). \tag{6}$$

One easily sees that

$$r_1 = \mathrm{Trace}(A) \stackrel{\mathrm{def}}{=} a_{11} + a_{22} + \cdots + a_{nn},$$

and hence

$$\lambda_1 + \cdots + \lambda_n = \mathrm{Trace}(A). \tag{7}$$

More generally, each $r_\ell = r_\ell(a_{ij})$ is a polynomial in the a_{ij} which is a sum of the determinants of the $\ell \times \ell$ *principal minors* of A.

The formulas (6) and (7) can be useful for a number of reasons, including so-called *trace methods* for computing exactly or approximating large eigenvalues, which are based on the formula

$$\lambda_1^k + \cdots + \lambda_n^k = \mathrm{Trace}(A^k) \tag{8}$$

for all $k \in \mathbb{N}$.

3.6. **Examples of Eigenpairs of Matrices.** Recall that a matrix $A \in M_n(\mathbb{R})$ or $M_n(\mathbb{C})$ is *diagonalizable* if there is a *basis* $\mathbf{v}_1, \ldots \mathbf{v}_n$ of $\mathbb{C}^n$ such that $M\mathbf{v}_i = \lambda_i\mathbf{v}_i$ for some $\lambda_i \in \mathbb{C}$. We say A is *orthonormally diagonalizable* if A has an eigenbasis $\mathbf{v}_1, \ldots, \mathbf{v}_n$ that is *orthonormal*, i.e., $\mathbf{v}_i \cdot \mathbf{v}_j = 0$ for $i \neq j$.

Example 3.4.

$$\begin{bmatrix} a & b \\ b & a \end{bmatrix},$$

We have $\det(A - \lambda I) = (\lambda - a)^2 - b^2$. Setting this equal to zero gives $\lambda = a \pm b$. We can check that $[1 \ \pm 1]^{\mathrm{T}}$ given a corresponding eigenbasis.

Example 3.5.

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix},$$

$(\lambda - a)^2 = -b^2$; $\lambda_i = a \pm ib$. Eigenvectors $[1 \ \pm i]$. **Note that these eigenvectors are not orthogonal if you just naively extend the real "dot product" to the complex numbers.** For this reason the complex analog of the dot product involved complex conjugation.

Example 3.6. 90 degree rotation: $\mathbf{e}_1 \to \mathbf{e}_2$, $\mathbf{e}_2 \to -\mathbf{e}_1$.

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

$\lambda_{1,2} = \pm i$.

Example 3.7. Rotation by θ radians:

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix};$$

$\lambda_{1,2} = \cos\theta \pm i \sin\theta = e^{\pm\theta i}$.

Example 3.8. Fibonacci directed graph:

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

$\lambda_{1,2} = (1 \pm \sqrt{5})/2$, i.e., the "golden ratio" and its "conjugate."

Example 3.9. Constant matrix: $a = b = 1$ above, and

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix},$$

Since the row sums are all equal, $\mathbf{m}.\mathbf{e}.1$ is an eigenvector with eigenvalue 1. Since all rows are equal and non-zero, we see that the kernel of this matrix—and therefore the multiplicity of 0 as an eigenvalue—is one less than the number of columns in this matrix (i.e., $3 - 1 = 2$). See Exercise **??**.

Example 3.10. Cyclic shift:

$$C_3 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix},$$

and similarly C_n for any $n \in \mathbb{N}$. Note that C_3 acts on row vectors if you think of a digraph where vertex 1 moves to vertex 2, etc.; otherwise C_3 acts on column vectors moving vertex 2 to vertex 1, vertex 3 to vertex 2. For any $n \in \mathbb{Z}$, we easily see that C_n has an eigenvalue for each the n-th roots of unity, i.e., the $\zeta^n = 1$, with eigenvector $\chi_\zeta = [1 \ \zeta \ \dots \zeta^{n-1}]^{\mathrm{T}}$ (see Exercise 3.14 for more on this). As of January 2021 we have all seen the following matrix too often:

$$\big(C_n + C_n^2 + \cdots + C_n^7\big)/7,$$

i.e., the 7-day moving average—thinking of n as large—acting on row vectors, or

$$\big(C_n^{-1} + C_n^{-2} + \cdots + C_n^{-7}\big)/7$$

acting on column vectors. Note that sometimes one wishes to work with a shift that is not cyclic, where the last vertex "maps to **0**" (i.e., disappears); in this case one works with

$$\tilde{C}_3 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix},$$

which is a "Jordan matrix" denotes $J_3(0)$ (see the next subsection), all of whose eigenvalues are 0, so that $\tilde{C}_3^3 = 0$.

3.7. **Examples of Non-Diagonlizable Matrices.** We shall see that "most" (in a sense made precise below) square matrices ove $\mathbb{R}$ or $\mathbb{C}$ are diagonalizable, so it is good to keep in mind some cases of non-diagonalizable matrices.

Example 3.11. If

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

then $p_t(A) = (t-b)^2$, so $\lambda_1 = \lambda_2 = b$ are the eigenvalues of A. But we easily see that $Ib - A$ is only one dimensional. Hence this A is not diagonalizable.

Example 3.12. For generally (see [HJ85, HJ13], Section 3.1), a $k \times k$ matrix of the form

$$J_k(\lambda) \stackrel{\text{def}}{=} \begin{bmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{bmatrix}$$

(where a blank space implies a 0) for some $\lambda \in \mathbb{R}$ (or $\mathbb{C}$) is called the $k \times k$ *Jordan block with eigenvalues* λ. We easily see that $J_k(\lambda)$ has the eigenvalue λ with multiplicity k, but has only a one dimensional space of eigenvalues. We also see that for any $m \in \mathbb{N}$

$$\big(J_3(\lambda)\big)^m = \begin{bmatrix} \lambda^m & m\lambda^{m-1} & \binom{m}{2}\lambda^{m-2} \\ & \lambda^m & m\lambda^{m-1} \\ & & \lambda^m \end{bmatrix} \tag{9}$$

and similarly for $J_k(\lambda)$ for any $k \in \mathbb{N}$ (see Exercise 3.12); see also Exercise 3.21 for $\lambda \neq 0$ and $m \in \mathbb{Z}$ with $m < 0$.

Example 3.13. We shall see that "most" (in a sense made precise below) square matrices ove $\mathbb{R}$ or $\mathbb{C}$ are diagonalizable, but there are applications where non-diagonalizable matrices do occur in practice. For example, if $G_0, G_1 \in \mathbb{R}$ and G_n for $n \geq 2$ is given by

$$G_n + a_1 G_{n-1} + a_2 G_{n-2} = 0,$$

then the general formula for G_n is $c_1 r_1^n + c_2 r_2^n$ provided that the equation $x^2 + a_1 x + a_2 = 0$ has distinct roots r_1, r_2. (E.g., consider the Fibonacci numbers.) On the other hand, for a recurrence like

$$G_n - 2G_{n-1} + G_{n-2} = 0,$$

in which case $x^2 - 2x + 1 = 0$ has roots $x = 1, 1$, the general solution for G_n is $G_n = c_1 + nc_2$. However, we easily see that for $n \geq 2$ we have

$$\begin{bmatrix} G_{n+1} \\ G_n \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix}^n \begin{bmatrix} G_1 \\ G_0 \end{bmatrix}$$

for this recurrence equation, and it turns out that this is precisely why the have matrix has eigenvalues $\lambda_1 = \lambda_2 = 1$ and is not diagonalizable.

Example 3.14. More generally, if $G_0, G_1, \dots$ is a sequence of integers satisfying the recurrence equation $G_n = b_1 G_{n-1} + \cdots + b_k G_{n-k}$ for all $n \geq k$ and fixed (k and) $b_1, \dots, b_k$, then for any $n \geq 0$ we have

$$\begin{bmatrix} G_{n+k-1} \\ G_{n+k-2} \\ \vdots \\ G_n \end{bmatrix} = A^n \begin{bmatrix} G_{k-1} \\ G_{k-2} \\ \vdots \\ G_0 \end{bmatrix}, \quad \text{where} \quad A = \begin{bmatrix} b_1 & b_2 & \cdots & b_{k-1} & b_k \\ 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \tag{10}$$

The above recurrence equation has general solution $c_1 r_1^n + \cdots + c_k r_k^n$, provided that A has distinct eigenvalues $r_1, \dots, r_k$, and is a bit more complicated when there are multiple roots (see Exercise 3.11).

3.8. **Generalized Eigenspaces.** When a matrix is not diagonalizable, it can be useful to work with *generalizaed eigenspaces*: if λ is an eigenvalue of $A \in \mathcal{M}_n(\mathbb{R})$, we define the *eigenspace of λ (in A)* to be

$$\mathrm{Eig}_A(\lambda) = \ker(A - \lambda I)$$

(which is the set of eigenvectors with eigenvalue λ plus $\mathbf{0}$), and the *generalized eigenspace of λ (in A)* to be

$$\mathrm{GenEig}_A(\lambda) = \bigcup_{k \in \mathbb{Z}} \ker\big((A - \lambda I)^k\big).$$

One can show that if the multiplicity of λ is m, then the genearlized eigenspace equals

$$\mathrm{GenEig}_A(\lambda) = \ker\big((A - \lambda I)^m\big).$$

In fact, the Jordan canonical form theorem is proven by the observation that the nested sequences of subspaces

$$V_1 = \mathrm{Eig}_A(\lambda) = \ker(A - \lambda I) \subset V_2 = \ker\big((A - \lambda I)^2\big) \subset V_3 = \ker\big((A - \lambda I)^3\big) \subset \dots$$

has to stop growing at some point, since the dimension of any subspace of $\mathbb{R}^n$ is at most n. If we take the largest k such that

$$V_{k-1} = \ker\big((A - \lambda I)^{k-1}\big) \neq \ker\big((A - \lambda I)^k\big) = V_k,$$

then we choose a vector v in $V_k \setminus V_{k-1}$, we have $(A - \lambda I)^k v = 0$, but one easily shows that the "Jordan chain"

$$v, (A - \lambda I)v, (A - \lambda I)^2 v \dots, (A - \lambda)^{k-1} v, \tag{11}$$

are all linearly independent (see Exercise 3.13). Now consider W, the span of the above vectors of the Jordan chain; if $W = \mathrm{GenEig}_A(\lambda)$, then we are done. Otherwise we consider the largest k' such that

$$W + V_{k'-1} \neq W + V_{k'},$$

choose a $v' \in (W + V_{k'}) \setminus (W + V_{k'-1})$, and extract a second chain

$$v', (A - \lambda I)v', (A - \lambda I)^2 v' \ldots, (A - \lambda)^{k-1} v',$$

set W' to be the span of these vectors and W, and continue similarly to extract chains until the span of all chains exhaust $\mathrm{GenEig}_A(\lambda)$.

Definition 3.15. The *geometric multiplicity* of an eigenvalue, λ, of $A \in \mathcal{M}_n(\mathbb{R}, \mathbb{C})$ is the dimension of $\ker(A - \lambda I)$, i.e., of $\mathrm{Eig}_A(\lambda)$ (the set of eigenvectors where we include $\mathbf{0}$). The *algebraic multiplicity*, of λ is its multiplicity as a root of $p_A(t)$, which equals the dimension of $\mathrm{GenEig}_A(\lambda)$. We usually are interested in the *algebraic multiplicity*, so we refer to this as, more simply, the *multiplicity*.

The following proposition is very useful.

Proposition 3.16. *Let $A \in \mathcal{M}_n(\mathbb{R}, \mathbb{C})$ have eigenvalues $\lambda_1, \ldots, \lambda_n$ listed with their multiplicities (i.e., as roots of $p_A(t)$). Then for any $m \in \mathbb{N}$ the eigenvalues of A^m are $\lambda_1^m, \ldots, \lambda_n^m$.*

Proof. The main idea is evident in the case $m = 2$. The eigenvalues of $-A$ are the roots of $p_{-A}(t) = \det(tI + A) = \det((-t)I - A)(-1)^n$, which are therefore $-\lambda_1, \ldots, -\lambda_n$. Since determinants factor through matrix multiplication,

$$\det(It^2 - A^2) = \det\big((It + A)(It - A)\big) = \det(I + At)\det(I - At),$$

and the right-hand-side has roots $\pm\lambda_1, \ldots, \pm\lambda_n$ counted with multiplicity (for each occurrence as roots of $\det(It - A)$). It follows that the roots of $\det(Iu - A^2)$ are $\lambda_1^2, \ldots, \lambda_n^2$, each with multiplicity two.

The general case can be proven by considering the roots of $\det(Iu - A^m)$; since

$$\det(It^m - A^m) = \prod_{\zeta^m = 1} \det(It - \zeta A)$$

has roots $\zeta\lambda_i$ for each i and each $\zeta^m = 1$, $\det(Iu - A^m)$ has roots $\lambda_1^m, \ldots, \lambda_n^m$. □

One can give another proof by approximating A by a matrix A' with distinct eigenvalues and taking limits, although one needs to take care that A' has eigenvalues that are not only distinct, but that also have their m-th powers distinct.

3.9. **Similar Matrices have the Same Characteristic Polynomial, Eigenvalues, and Similar Eigenvectors.** If $A, B, S \in \mathcal{M}_n(\mathbb{C})$ with $A = SBS^{-1}$ and $A\mathbf{v} = \lambda\mathbf{v}$ for some $\mathbf{v} \in \mathbb{C}^n$ and $\lambda \in \mathbb{C}$, then setting $\mathbf{v}' = S^{-1}\mathbf{v}$ we have $B\mathbf{v}' = \lambda\mathbf{v}'$. Hence there is a one-to-one correspondence (namely $v \mapsto S^{-1}\mathbf{v}$) between the eigenvectors of A with those of B, with the same underlying eigenvalue. Hence the geometric multiplicity of A and B are the same.

Moreover,

$$p_t(A) = \det(tI - A) = \det(tI - SBS^{-1}) = \det(S(tI - B)S^{-1}) = \det(S)p_t(B)\det(S^{-1}) = p_t(B).$$

Hence when we count the multiplicity of an eigenvalue according to its characteristic polynomial, the multiplicities of each eigenvalue of A is the same as that of B.

3.10. **A Matrix with Distinct Eigenvalues is Diagonalizable.** See Exercise 3.5

3.11. **An Upper Triangular Matrix with Distinct Diagonal Entries is Diagonalizable.** If $A \in \mathcal{M}_n(\mathbb{R},\mathbb{C})$ is an *upper triangular*, i.e., $a_{ij}=0$ for $i>j$, then $\det(A)=a_{11}\dots a_{nn}$ (since for any permutation, σ, of $[n]$ we have $i>\sigma(i)$ for some $i\in[n]$ unless σ is the identity permutation).

It follows (EXERCISE) that the eigenvalues of A are $a_{11},\dots,a_{nn}$, and the i-th eigenvector can be taken to be e_i plus some linear combination of $e_1,\dots,e_{i-1}$. Hence the eigenvectors form a basis of $\mathbb{R}^n,\mathbb{C}^n$ (EXERCISE).

It also follows that the Jordan matrix $J_k(\lambda)$ has the eigenvalue λ occuring with geometric multiplicity k. The fact that every matrix can be brought into Jordan canonical form (the argument was sketched above), shows that the algebraic multiplicity of an eigenvalue, λ, is precisely the sum of the dimensions of the Jordan blocks $J_k(\lambda)$ of the canonical form. This is one way to show that the algebraic multiplicity of an eigenvector, λ, in $A\in\mathcal{M}_n(\mathbb{R},\mathbb{C})$ is precisely the dimension of the associated generalized eigenspace of λ, $\mathrm{GenEig}_A(\lambda)$.

3.12. **Application to the Stationary Distribution.**

Example 3.17. If $P\in\mathcal{M}_n(\mathbb{R})$ is a Markov matrix, i.e., a row stochastic matrix, then $P\mathbf{e}_i=1$ for all $i\in[n]$, since $P\mathbf{e}_i$ is just the i-th row sum of P (i.e., the sum of the entries of the i-th row of P). Hence $P\mathbf{1}=\mathbf{1}$, where $\mathbf{1}$ is the vector whose components are $\mathbf{1}$. Hence 1 is an eigenvalue of P (it turns out to be one of the largest in absolute value, and the largest when P is not *periodic*). Hence for some $\boldsymbol{\pi}\in\mathbb{R}^n$ we have $\boldsymbol{\pi}^{\mathrm{T}}P=\boldsymbol{\pi}$. When P has all positive entries, then there is a unique such $\boldsymbol{\pi}$ if we normalize $\boldsymbol{\pi}$ to have its sum of entries equal 1 and we call this $\boldsymbol{\pi}$ the *stationary distribution* of P; it turns out that $\boldsymbol{\pi}\ge\mathbf{0}$ (i.e., $\pi_i\ge 0$ for all i). A similar theorem holds if P represents an *irreducible* Markov chain (see below).

See also Section 1.5 of [LP17] for an introduction to the stationary distribution of a Markov chain.

3.13. **Examples: The Boolean Cube, Cartesian Product of Graphs, and Tensor Product (of Graphs and Matrices).**

Remark 3.18. In CPSC 531F, Spring 2021, we discussed Section 4.1 simultaneously with examples in this subsection, in order to get used to the formula $A=\sum_i \lambda_i \mathbf{v}_i\mathbf{v}_i^{\mathrm{T}}$ for these examples.

In this section we give more examples of graphs and their eigenpairs. Many interesting examples arise from the *cartesian product* and *tensor product* of graphs; these can be understood in terms of the *tensor product* of matrices. Let us explain.

The *Boolean n-cube*, $\mathbb{B}^n$ is the graph whose vertex set is $V_{\mathbb{B}^n}=\{0,1\}^n$, such that for each i and $x_1,\dots,x_n\in\{0,1\}$, there is one edge from

$$(x_1,\dots,x_{i-1},x_i,x_{i+1},\dots,x_n)\quad\text{to}\quad(x_1,\dots,x_{i-1},1-x_i,x_{i+1},\dots,x_n)$$

(this describes a directed graph that we view as a graph by identifying pairs of directed edges in the only way possible). Hence

$$A_{\mathbb{B}^1}=\begin{bmatrix}0&1\\1&0\end{bmatrix}$$

(which equals C_2) and has eigenpairs $1, (1,1)$ and $-1, (1,-1)$. Also notice that

$$A_{\mathbb{B}^2} = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix} = C_4 + C_4^{-1}$$

which therefore has eigenpairs $\zeta + \zeta^{-1}, (1, \zeta, \zeta^2, \zeta^3)$ ranging over all 4-th roots of unity ζ (i.e., $\zeta^4 = 1$, i.e., $\zeta = \pm 1, \pm i$); hence $A_{\mathbb{B}^2}$ has eigenvalues 2 ($\zeta = 1$), 0 with multiplicity 2 ($\zeta = \pm i$), -2 ($\zeta = -1$). [For $n \geq 3$, $A_{\mathbb{B}^2}$ is no longer a circulant matrix, and hence is unrelated to sums of powers of C_{2^n}.] Some properties of Boolean functions on n variables can be studied by looking at the eigenpairs of $A_{\mathbb{B}^n}$; this is also called the "Fourier analysis" of the Boolean cube. Let us explain.

To find the eigenpairs of $A_{\mathbb{B}^n}$ we define the *cartesian product*, $G \times H$, of any two directed graphs, G, H: this is the following graph on the vertex set $V_{G\times H} = V_G \times V_H$: informally, $G \times H$ has directed edges of two kinds: (1) for each $e \in E_G^{\text{dir}}$ and each $v \in V_H$, $G \times H$ has a directed edge from $(t_G(e), w)$ to $(h_G(e), w)$, and (2) similarly for each $v \in V_G$ and $e \in V_H$. More formally,

$$E_{G\times H}^{\text{dir}} = E_G^{\text{dir}} \times V_H \amalg V_G \times E_H^{\text{dir}},$$

(where $\amalg$ is the disjoint union[4]) where t, h map $(e, w) \in E_G^{\text{dir}} \times V_H$ to $(t(e), w), (h(e), w)$, respectively, and similarly for directed edges in $V_G \times E_H^{\text{dir}}$.

The cartisian product is quite common. For example, the product of two 1-dimensional "grid graphs" (EXPLAIN) is a 2-dimensional "grid graph." Similarly, the product of two "cycle graphs" is a 2-dimensional "torus graph," a grid graph with "wrap around." The Boolean n-cube is the n-fold product of the Boolean 1-cube.

It turns out that if $|V_G| = n$ and $|V_H| = m$, and A_G, A_H respectively have eigenpair bases $\lambda_i, \mathbf{u}_i$ (so $i \in [n]$) and $\nu_j, \mathbf{w}_j$, then $A_{G\times H}$ has a basis of eigenpairs

$$\lambda_i(G) + \lambda_j(H), \mathbf{u}_i \otimes \mathbf{w}_j, \quad i \in [n],\ j \in [m]. \tag{12}$$

where for any $\mathbf{u} \in \mathbb{C}^{V_G}$ and $\mathbf{w} \in \mathbb{C}^{V_H}$ we define

$$(\mathbf{u} \otimes \mathbf{w})(v_1, v_2) = u(v_1)w(v_2)$$

(see Exercise 3.15).

It follows that $A_{\mathbb{B}^2}$ has eigenvalues $2, 0, 0, -2$ with eigenvectors $(1, \pm 1) \otimes (1, \pm 1)$ (draw this). Furthermore, the normalized eigenvectors of $\mathbb{B}^1$, arranged as an orthogonal matrix is

$$H_1 \stackrel{\text{def}}{=} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix},$$

and the eigenpairs of $\mathbb{B}^n$ are $(H_1)^{\otimes n} = H_1 \otimes \cdots \otimes H_1$, where we define the *tensor product* of two matrices, $A \in \mathcal{M}_{m_1,n_1}$ and $B \in \mathcal{M}_{m_2,n_2}$ to be $A \otimes B \in \mathcal{M}_{m_1 m_2, n_1 n_2}$ as having its rows indexed on pairs in $[m_1] \times [m_2]$ and its columns indexed on pairs $[n_1] \times [n_2]$, whose $(i_1, i_2), (j_1, j_2)$ entry is $a_{i_1 j_1} b_{i_2 j_2}$.

Here are some other useful ways to look at the cartesian product of graphs, in terms of tensor products of matrices and *tensor products of graphs*. The following statements are left as exercises.

[4] The *disjoint union*, $A \amalg B$, of two sets is, informally, their union when we regard them as not intersecting, and is therefore of size $|A| + |B|$; more formally, it is a limit that is defined only up to unique isomorphism, which can be taken to be $A \times \{0\} \cup B \times \{1\}$.

(1) If G, H are directed graphs, then

$$A_{G\times H} = A_G \otimes I_H + I_G \otimes A_H, \tag{13}$$

where I_G, I_H are identity matrices (Exercise 3.16).

(2) If A, B are diagonalizable matrices, with eigenpair bases $\lambda_i, \mathbf{u}_i$ and $\nu_j, \mathbf{w}_j$, then

$$\lambda_i\nu_j, \mathbf{u}_i \otimes \mathbf{w}_j \tag{14}$$

is an eigenpair basis for $A \otimes B$ (Exercise 3.17).

(3) For G, H directed graphs, there is a natural way to define a graph $G \otimes H$ with the property that $A_{G\otimes H} = A_G \otimes A_H$ (Exercise 3.18).

(4) If G, H have the same vertex set $V_G = V_H$, there is a natural way to define the *superposition* of G and H, $G \amalg H$, as a graph where $A_{G\amalg H} = A_G + A_H$ (Exercise 3.19).

(5) One can view $G \times H$ as a "superposition" of (1) G times a (very simple) graph with vertex set H, and (2) a (very simple) graph with vertex set G times H. How? (Exercise 3.20.)

3.14. EXERCISES.

Exercise 3.1. Verify the eigenvalue/vector computations in Examples 3.4 and 3.5.

Exercise 3.2. Let n be the *undirected cycle of length* n, i.e., whose adjacency matrix is $C_n + C_n^{-1}$ with C_n as in Example 3.10.

3.2(a) Give an expression for the eigenvalues of $C_n + C_n^{-1}$.

3.2(b) Show that if we arrange the eigenvalues as

$$\lambda_n \le \cdots \le \lambda_2 \le \lambda_1$$

then $\lambda_1 = 2$ and $\lambda_2 = 2 + C/n^2 + O(1/n^4)$ for n large and some (negative) constant C; what is C?

3.2(c) Show that $\lambda_n = -2$ iff n is even, and that if n is odd then $\lambda_n = \lambda_{n-1} = -2 + C'/n^2 + O(1/n^4)$ for n large and some constant C'; what is C'?

Exercise 3.3. Show that the eigenvalues of $E \in \mathcal{M}_n(\mathbb{R})$ given by

$$E = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix}$$

are n—with multiplicity 1, and 0 with multiplicity $n-1$. Do this by giving a basis for $\mathrm{Eig}_E(0)$ (the space of eigenvectors of E with eigenvalue 0) of $n-1$ vectors; make sure to prove that your $n-1$ vectors are really a basis.

Exercise 3.4. Let $A \in \mathcal{M}_n(\mathbb{R})$ be any matrix of the form:

$$A = \mathbf{u}\mathbf{w}^{\mathrm{T}}$$

where $\mathbf{u}, \mathbf{w}$ are nonzero vectors in $\mathbb{R}^n$. (Such an A is a matrix of rank 1, i.e., the dimension of its image is 1; the matrix E of Exercise 3.3.) Show that A has an eigenvalue $\mathbf{u}\cdot\mathbf{w}$ with multiplicity 1, and eigenvalue 0 with multiplicity $n-1$. [You may quote any theorems you like regarding rank, kernel (nullspace), etc.]

Exercise 3.5. Let $A \in \mathcal{M}_n(\mathbb{R}, \mathbb{C})$, and $\mathbf{v}_1, \ldots, \mathbf{v}_k$ be non-zero eigenvectors of A with distinct corresponding eigenvalues (i.e., $A\mathbf{v}_j = \lambda_j\mathbf{v}_j$ with λ_j distinct).

3.5(a) Show that $\mathbf{v}_1, \ldots, \mathbf{v}_k$ are linearly independent. [Hint: use induction on k; for $k \geq 2$, we may assume that

$$\alpha_1 \mathbf{v}_1 + \cdots + \alpha_k \mathbf{v}_k = \mathbf{0}$$

and that $\alpha_j \neq 0$ for all $j \in [k]$; multiply the above equation by A, and use the new equation and original one to eliminate $\mathbf{v}_k$.]

3.5(b) If A has n distinct eigenvalues, show that any set of (non-zero) corresponding eigenvectors is a basis of $\mathbb{R}^n$. Show that in this case A is diagonalizable.

Exercise 3.6. Prove the two statements labelled (EXERCISE) in the second paragraph of Subsection 3.11.

Exercise 3.7. Let notation be as in Example 2.2.

3.7(a) Show that for any k we have

$$\begin{bmatrix} F_{k+1} & F_k \\ F_k & F_{k-1} \end{bmatrix} = A_{\text{Fib}} \begin{bmatrix} F_k & F_{k-1} \\ F_{k-1} & F_{k-2} \end{bmatrix}$$

3.7(b) Show that

$$A_{\text{Fib}}^k = \begin{bmatrix} F_{k+1} & F_k \\ F_k & F_{k-1} \end{bmatrix}$$

for all integers $k \geq 2$.

3.7(c) Show that A_{Fib} is invertible, and that the above formula for A_{Fib}^k holds for all integers k.

3.7(d) Let $\{G_k\}_{k \in \mathbb{Z}}$ be any doubly infinite sequence of reals such that $G_{k+2} = G_{k+1} + G_k$ for all k. What is

$$A_{\text{Fib}}^n \begin{bmatrix} G_1 \\ G_0 \end{bmatrix}?$$

Exercise 3.8. Let $x_1, \ldots, x_n$ be distinct real or complex numbers. Show that

$$\det \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{bmatrix} = \prod_{i<j} (x_j - x_i).$$

The above matrix is called a *Vandermonde matrix*, and its determinant a *Vandermonde determinant*. [Hint: Use induction on n; for the inductive step view x_n as a variable $t = x_n$, and show that

$$p(t) = \det \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & t & t^2 & \cdots & t^{n-1} \end{bmatrix}$$

is a polynomial of degree $n-1$; what are the roots of this polynomial, and what is its leading coefficient?]

Exercise 3.9. 3.9(a) Show that the recurrence equation $x_n - 3x_{n-1} + 3x_{n-2} - x_{n-3} = 0$ is satisfied by $x_n = c_0 + c_1 n + c_2 n^2$ for any $c_0, c_1, c_2 \in \mathbb{R}$.

3.9(b) Show that for distinct i, j, k, and any given values of x_i, x_j, x_k, there is a unique solution to the above recurrence equation. (You can use the result of Exercise 3.8.)

3.9(c) What is the connection between the above recurrence equation and the matrix

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

3.9(d) What are the values for all $k \in \mathbb{Z}$ of

$$A^k \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad A^k \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}, \quad A^k \begin{bmatrix} 9 \\ 4 \\ 1 \end{bmatrix}?$$

3.9(e) Show that for $\mathbf{v} = [9\ 4\ 1]^{\mathrm{T}}$,

$$\mathbf{v}, (A-I)\mathbf{v}, (A-I)^2\mathbf{v}$$

is a basis for $\mathbb{R}^3$. Show that A with respect to this basis equals $J_3(1)$.

3.9(f) Show the same for $v = [1\ 0\ 0]^{\mathrm{T}}$.

3.9(g) Would the same hold for $v = [3\ 2\ 1]^{\mathrm{T}}$?

3.9(h) Give an expression for

$$A^k \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

[Hint: Explain (with little or no calculation, ideally) why we have

$$A \begin{bmatrix} \binom{n}{2} \\ \binom{n-1}{2} \\ \binom{n-2}{2} \end{bmatrix} = \begin{bmatrix} \binom{n+1}{2} \\ \binom{n}{2} \\ \binom{n-1}{2} \end{bmatrix}$$

for any $n \in \mathbb{Z}$ (or, for that matter, $n \in \mathbb{R}$).]

Exercise 3.10. Consider the recurrence equation $x_n - 4x_{n-1} + 6x_{n-2} - 4x_{n-3} + x_{n-4} = 0$, or equivalently the equation

$$(\sigma - 1)^4 x_n = 0,$$

where σ is the "upward shift operator," i.e., $\sigma x_n = x_{n+1}$. [One can view this is a formal way of manipulating subscripts; however, σ can be viewed more precisely as a map on sequences $\mathbf{x} = \{x_n\}_{n\in\mathbb{Z}}$ taking X to the sequence $\sigma\mathbf{x}$ given by $(\sigma\mathbf{x})_n = x_{n+1}$.] Consider the associated (via (10))

$$A = \begin{bmatrix} 4 & -6 & 4 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

3.10(a) For any $k \in \mathbb{N}$, find a simple expression for $(\sigma - 1)\binom{n}{k}$.

3.10(b) Characterize the solutions to the above recurrence equation.

3.10(c) For which vectors $\mathbf{v} \in \mathbb{R}^4$ is

$$\mathbf{v}, (A-I)\mathbf{v}, (A-I)^2\mathbf{v}, (A-I)^3\mathbf{v}$$

a Jordan chain for A, and for which v is not?

Exercise 3.11. Consider a general recurrence equation with constant coefficients in $\mathbb{R}, \mathbb{C}$ (or really any field) $p(\sigma)x_n = 0$, where σ is the shift operator in the previous exercise (i.e., $\sigma x_n = x_{n+1}$), and p is a polynomial of degree k with leading coefficient 1; i.e., if

$$p(t) = t^k + c_1 t^{k-1} + c_2 t^{k-2} + \cdots + c_k,$$

then the recurrence equation is

$$x_{n+k} + c_1 x_{n+k-1} + \cdots + c_k x_n = 0,$$

or equivalently

$$x_n + c_1 x_{n-1} + \cdots + c_k x_{n-k} = 0.$$

3.11(a) Show that if $c_k = p(0) \neq 0$, and $p(t)$ has roots $r_1, \ldots, r_s$ with r_j occurring with multiplicity m_j, then a general solution to this recurrence is given by

$$x_n = \sum_{j=1}^{s} r_j^n p_j(n)$$

where p_j is any polynomial of degree at most $m_j - 1$.

3.11(b) Describe the Jordan form of the matrix (10) with $b_i = -c_i$.

Exercise 3.12. Fix a $k \in \mathbb{N}$ and let $J = J_k(0)$.

3.12(a) Draw a directed graph whose adjacency matrix is J. [Hint: you might try $k = 2, 3$ if you don't see the general construction.]

3.12(b) Using the above directed graph, describe how $J^2, \ldots, J^k$ acts on row vectors; use this to describe the matrices $J^2, \ldots, J^k$.

3.12(c) Explain why $A, B \in \mathcal{M}_n(\mathbb{R}, \mathbb{C})$ commute, i.e., $AB = BA$, then for any $m \in \mathbb{N}$,

$$(A + B)^m = A^m + \binom{m}{1} A^{m-1} B + \cdots + B^m.$$

3.12(d) Use the fact that $J_k(\lambda) = \lambda I + J_k(0)$ to show that (9) holds.

Exercise 3.13. Let $A \in \mathcal{M}_n(\mathbb{R}, \mathbb{C})$ such that for some $\lambda \in \mathbb{R}, \mathbb{C}$ and nonzero $\mathbf{v} \in \mathbb{R}^n, \mathbb{C}^n$, $(A - \lambda I)^k \mathbf{v} = \mathbf{0}$ and $(A - \lambda I)^{k-1} \mathbf{v} \neq \mathbf{0}$. Show that

$$\mathbf{v}, (A - \lambda I)\mathbf{v}, \ldots, (A - \lambda I)^{k-1}\mathbf{v}$$

are linearly independent. [Hint: Consider the greatest common factor of the set of polynomials $p(x)$ such that $p(A)\mathbf{v} = \mathbf{0}$.]

Exercise 3.14. Let $n \in \mathbb{N}$. For each n-th root of unity, ζ, i.e., $\zeta^n = 1$, we eigenvector of C_n (as in Example 3.10), $\chi_\zeta = [1 \ \zeta \ \ldots \zeta^{n-1}]^{\mathrm{T}}$, with eigenvalue ζ. Since the n values of ζ are distinct, we know that the χ_ζ are linearly independent.

3.14(a) Use Exercise 3.8 to give an alternate proof that the χ_ζ are linearly independent.

3.14(b) Show that if $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{v}^H \mathbf{u}$ is the standard complex inner (or "dot") product, then any two distinct χ_ζ are orthogonal. [Hint: Since $\zeta^n - 1$ can be factored as $(\zeta - 1)(\zeta^{n-1} + \cdots + 1)$, we have $\zeta^{n-1} + \cdots + 1 = 0$ for $\zeta^n = 1$ with $\zeta \neq 1$.]

Exercise 3.15. Show that (12) gives a basis of eigenpairs of $G \times H$, in the notation of that paragraph.

Exercise 3.16. Show that (13) holds.

Exercise 3.17. Show that (14) gives a basis of eigenpairs with the notation and assumptions of that paragraph.

Exercise 3.18. Give a general definition of a digraph $G \otimes H$ for G, H digraphs with the property that $A_{G\otimes H} = A_G \otimes A_H$; we call $G \otimes H$ the *tensor product* of G and H or G *tensored with* H.

Exercise 3.19. If G, H have the same vertex set $V_G = V_H$, define the *superposition* of G and H, $G \amalg H$, in a way that makes $A_{G\amalg H} = A_G + A_H$.

Exercise 3.20. Explain how $G \times H$ can be viewed as a superposition of (1) G tensored with with a (very simple) graph with vertex set V_H, and (2) a (very simple) graph with vertex set V_G tensored with H.

Exercise 3.21. Show that (9) holds for any $\lambda \neq 0$ and $m \in \mathbb{Z}$. Then prove the analogous result for $J_k(\lambda)$ any $k \in \mathbb{N}$. [Hint: obtain a formula for $(I - N)^{-m}$ for N nilpotent (i.e., $N^t = 0$ for some $t \in \mathbb{N}$) based on a power series for $(1 - x)^{-m}$, which you can obtain by differentiating $m - 1$ times the power series $(1 - x)^{-1} = 1 + x + x^2 + \cdots$.]

4. Orthonormal Eigenbases, Expanders, Reversible Markov Chains, Rayleigh Quotients, and the SVD (Singular-Value Decomposition)

4.1. **Orthnormal Eigenbases, Orthogonal and Unitary Matrices.** Recall that a matrix $Q \in \mathcal{M}_n(\mathbb{R})$ is called *orthogonal* if $QQ^{\mathrm{T}} = I$ (where I is the $n \times n$ identity matrix). Here is how these matrices arise.

Proposition 4.1. *Let $\mathbf{v}_1, \dots, \mathbf{v}_n \in \mathbb{R}^n$. Then the following are equivalent:*

(1) $\mathbf{v}_1, \dots, \mathbf{v}_n$ are orthonormal;
(2) the matrix, Q, whose columns are $\mathbf{v}_1, \dots, \mathbf{v}_n$ satisfies $QQ^{\mathrm{T}} = I$;

if so, then

(1) $\mathbf{v}_1, \dots, \mathbf{v}_n$ form a basis for $\mathbb{R}^n$;
(2) for any $\mathbf{u} \in \mathbb{R}^n$ we have

$$\mathbf{u} = \sum_{i=1}^{n} (\mathbf{u} \cdot \mathbf{v})\mathbf{v};$$

(3) $Q^{-1} = Q^{\mathrm{T}}$, $Q^{\mathrm{T}}Q = I$;
(4) Q^{T} is also an orthogonal matrix;
(5) the rows of Q also form an orthonormal eigenbasis;
(6) for any $\mathbf{u}, \mathbf{w} \in \mathbb{R}$, $(Q\mathbf{u}) \cdot (Q\mathbf{w}) = \mathbf{u} \cdot \mathbf{w}$;
(7) the map $\mathbf{u} \mapsto Q\mathbf{u}$ preserves the lengths of vectors and the angle between vectors.

Here are some examples:

(1) $\mathbf{e}_1, \dots, \mathbf{e}_n$, the standard basis; $Q = I$;
(2) some permutation of $\mathbf{e}_1, \dots, \mathbf{e}_n$; Q is a permutation matrix;
(3) $(1, 1)/\sqrt{2}$, $(1, -1)/\sqrt{2}$ in $\mathbb{R}^2$;
(4) $(1, 1, 1)/\sqrt{3}$, $(1, -1, 0)/\sqrt{2}$, $(1, 1, -2)/\sqrt{6}$ in $\mathbb{R}^3$.

Remark 4.2. If you are teaching in linear algebra at some point in your career, when you teach the diagonalization of matrices you are likely to use matrices such as

$$E_3 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \quad E_3 - I = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}, \quad E_3 + 2I$$

as an example where you have a repeated eigenvalue. In this case mechanical Gaussian elimination on the eigenspace of dimension two will produce (up to $\pm$) the eigenbasis $(1,1,1)/\sqrt{3}$, $(1,-1,0)/\sqrt{2}$, $(1,1,-2)/\sqrt{6}$, yielding the orthogonal matrix

$$Q = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & -1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & 0 & -2/\sqrt{6} \end{bmatrix}$$

which looks rather awkward compared to the eigenvectors for C_3 (or any circulant 3×3 matrix) yielding the unitary matrix

$$U = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 1/\sqrt{3} & \omega/\sqrt{3} & \omega^2/\sqrt{3} \\ 1/\sqrt{3} & \omega^2/\sqrt{3} & \omega/\sqrt{3} \end{bmatrix}$$

where ω is a *primitive* third root of unity (a *primitive* n-th root of unity is a ζ such that $\zeta^n = 1$ but $\zeta^m \neq 1$ for integers $1 \leq m < n$).

Comparing Q, U above should convince you that it is sometimes nicer to work with a complex set of eigenvectors. Hence you should note the following proposition.

A matrix $U \in \mathcal{M}_n(\mathbb{C})$ with $UU^{\mathrm{H}} = I$ is called *unitary*.

Proposition 4.3. *Proposition 4.1 holds if everywhere you replace* $\mathbb{R}$ *with* $\mathbb{C}$, $^{\mathrm{T}}$ *with* $^{\mathrm{H}}$, *the word "orthogonal" with "unitary," and the complex dot product* $\mathbf{u} \cdot \mathbf{v} = \mathbf{v}^{\mathrm{H}}\mathbf{u}$.

Note that some authors use $\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^{\mathrm{H}}\mathbf{v}$, which makes the dot product linear (under scalar multiplication, i.e., multiplication by $\mathbb{C}$) in $\mathbf{v}$ and "skew linear" in the first variable; in these notes will (try to) follow the conventions in [HJ85, HJ13], that take $\mathbf{u} \cdot \mathbf{v} = \mathbf{v}^{\mathrm{H}}\mathbf{u}$. [A bilinear form $\mathbb{C}^m \times \mathbb{C}^n \to \mathbb{C}$ that is linear in one variable and skew-linear in the other variable is typically called *sesquilinear*.]

Of course, an orthogonal matrix is just a unitary matrix with purely real entries. So Proposition 4.1 is a special case of Proposition 4.3. Also, note that the orthonormal basis:

$$\begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}, \quad \begin{bmatrix} 1/\sqrt{3} \\ \omega/\sqrt{3} \\ \omega^2/\sqrt{3} \end{bmatrix}, \quad \begin{bmatrix} 1/\sqrt{3} \\ \omega^2/\sqrt{3} \\ \omega\sqrt{3} \end{bmatrix}$$

fails to be orthonormal (in its last two vectors) if you use the real dot product instead of the complex one.

Proposition 4.4. *Let* $A \in \mathcal{M}_n(\mathbb{R})$, *and* $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{R}^n$ *be orthnormal, and* $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ *Then the following are equivalent:*

(1) $A\mathbf{v}_i = \lambda_i$ *for all* $i \in [n]$*;*
(2) $A = \sum_{i=1}^n \lambda_i \mathbf{v}_i \mathbf{v}_i^{\mathrm{T}}$.

More generally, the above hold with $\mathbb{R}$ *replaced by* $\mathbb{C}$ *and* $^{\mathrm{T}}$ *replaced with* $^{\mathrm{H}}$.

Note that this last proposition applies to:

(1) circulant matrices (i.e., a sum of powers of C_n);
(2) A_G for the graphs we have seen in Subsection 3.13;
(3) the matrix

$$E_3 = I + C_3 + C_3^2 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}.$$

Notice that it is a bit awkward to find explicit, real $\mathbf{v}_2, \mathbf{v}_3$ that are orthonormal and are in the kernel of E_3, i.e., are eigenvectors with eigenvalue 0; since $\mathbf{v}_1 = (1,1,1)/\sqrt{3}$ has eigenvalue 3, one is naturally lead to the example above of an orthonormal eigenbasis in $\mathbb{R}^3$. The complex eigenbasis with $\mathbf{v}_2 = (1, \omega, \omega^2)/\sqrt{3}$ and $\mathbf{v}_3 = (1, \omega^2, \omega)/\sqrt{3}$ is a bit better.

4.2. **Symmetric Matrices and the L^2-Operator Norm.** At this point we need the following theorem.

Theorem 4.5. *Let $A \in \mathcal{M}_n(\mathbb{R})$ be a symmetric matrix. Then A has real eigenvalues and an orthonormal eigenbasis in $\mathbb{R}^n$. Similarly Hermitian $A \in \mathcal{M}_n(\mathbb{C})$ (i.e., $A^{\mathrm{H}} = A$), (real eigenvalues, but an orthonormal eigenvalues in $\mathbb{C}^n$).*

There are a number of different proofs in these notes.

(1) Verify this for A with distinct eigenvalues: the general fact that $(\mathbf{Au}) \cdot \mathbf{w} = \mathbf{u} \cdot (A^* \mathbf{w})$ shows that if $A\mathbf{v}_1 = \lambda_1 \mathbf{v}_1$ and $A\mathbf{v}_2 = \lambda_2 \mathbf{v}_2$, then $(\lambda_1 - \overline{\lambda_2})(\mathbf{v}_1 \cdot \mathbf{v}_2) = 0$. This easily implies the theorem when A has distict eigenvalues. When A does not have distinct eigenvalues, one takes a sequence $A_1, A_2, \ldots$ of symmetric (or Hermitian) matrices with distinct eigenvalues whose limit is A, and passes to a subsequence where an orthonormal eigenbasis converges.
(2) Use the Rayleigh quotient

$$\mathcal{R}_A(\mathbf{u}) = \frac{(A\mathbf{u}) \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{u}} = \frac{(A\mathbf{u}) \cdot \mathbf{u}}{|\mathbf{u}|_2^2}$$

to prove many properties of the eigenpairs of A, including a number of theorems regarding the "interlacing of eigenvalues (between two matrices)." This will be extremely important to us.
(3) Prove that any matrix $N \in \mathcal{M}_n(\mathbb{C})$ with $NN^{\mathrm{H}} = N^{\mathrm{H}}N$ has an orthonormal eigenbasis in $\mathbb{C}^n$. Then view symmetric and Hermitian matrices as special cases.

All of these proofs are important.

The upshot of the above theorem is that if A is symmetric (or, more generally, Hermitian), then there is an orthonormal basis with respect to which A is given by $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$; since the basis is orthonormal, i.e., $A = QDQ^{-1}$ where Q is orthogonal (or $A = UDU^{-1}$ where U is unitary in the complex case), the transformations Q, Q^{-1} (or U, U^{-1}) preserve the dot product, and therefore lengths of vectors and angles between them. Hence we easily see from the diagonal matrix D that

$$\|A\|_{L^2} \stackrel{\text{def}}{=} \max_{\mathbf{v} \neq \mathbf{0}} \frac{|A\mathbf{v}|_2}{|\mathbf{v}|_2} = \max_i |\lambda_i|.$$

In terms of the Rayleigh quotient, etc.

4.3. **The Rayleigh Quotient and Related Inequalities.** [See also Section 4.2 [HJ85, HJ13].]

At this point we are assuming that every (real) symmetric matrix, A, can be written as QDQ^{-1} where D is a real, diagonal matrix and Q is orthogonal (i.e., $QQ^* = I$), or, equivalently, A has an orthonormal eigenbasis with real eigenvalues (more generally, any Hermitian matrix, A, can be written as UDU^{-1} with D real, diagonal and U unitary, or, equivalently A has an orthonormal eigenbasis with real eigenvalues).

Theorem 4.6. *Let $A \in \mathcal{M}_n(\mathbb{R})$ be symmetric (or $A \in \mathcal{M}_n(\mathbb{C})$ be Hermitian), i.e., $A^* = A$, and let*

$$\lambda_n \leq \cdots \leq \lambda_2 \leq \lambda_1$$

be its eigenvalues. Then

(1) for any $\mathbf{w} \in \mathbb{R}^n$ (or $\mathbb{C}^n$ in the Hermitian case),

$$\lambda_n(\mathbf{w} \cdot \mathbf{w}) \leq \mathbf{w} \cdot (A\mathbf{w}) \leq \lambda_1(\mathbf{w} \cdot \mathbf{w}); \tag{15}$$

(2) for any $\mathbf{w} \in \mathbb{R}^n$ (or $\mathbb{C}^n$),

$$\|A\mathbf{w}\|_2 \leq (\max_i |\lambda_i|)\|\mathbf{w}\|_2; \tag{16}$$

(3) for any $\mathbf{u}, \mathbf{w} \in \mathbb{R}^n$ (or $\mathbb{C}^n$),

$$|\mathbf{u} \cdot A\mathbf{w}| \leq \big(\max_i |\lambda_i|\big)\|\mathbf{u}\|_2 \, \|\mathbf{w}\|_2. \tag{17}$$

Proof. First consider the case when A is a diagonal matrix $A = \operatorname{diag}(d_1, \ldots, d_n)$ (hence the $d_i \in \mathbb{R}$ in either the symmetric or Hermitian case). In this case A's eigenvalues are the d_i. Also

$$\mathbf{w} \cdot A\mathbf{w} = d_1 w_1^2 + \cdots + d_n w_n^2,$$

so

$$\mathbf{w} \cdot (A\mathbf{w}) \leq (\max_i d_i)(w_1^2 + \ldots + w_n^2) = \lambda_1 \mathbf{w} \cdot \mathbf{w},$$

and equality holds iff for all j with $w_j \neq 0$ we have $d_j = \lambda_1$, and in this case $A\mathbf{w} = \lambda_1 \mathbf{w}$. Similarly for the ineqality $\lambda_n(\mathbf{w} \cdot \mathbf{w}) \leq \mathbf{w} \cdot A\mathbf{w}$. Also

$$\|A\mathbf{w}\|^2 = \|d_1\mathbf{w}_1 + \cdots + d_n\mathbf{w}_n\|^2$$

$$= (d_1\mathbf{w}_1 + \cdots + d_n\mathbf{w}_n) \cdot (d_1\mathbf{w}_1 + \cdots + d_n\mathbf{w}_n) = d_1^2 w_1^2 + \cdots + d_n w_n^2,$$

which implies (16). The last inequality follow from Cauchy-Schwartz:

$$|\mathbf{u} \cdot A\mathbf{w}| \leq \|\mathbf{u}\|_2 \, \|A\mathbf{w}\|_2.$$

Next consider the case where $A = QDQ^{-1}$ with D real diagonal and Q orthogonal (respectively, $A = UDU^{-1}$, U unitary). Since Q (respectively U) preserves the dot product, and D and A have the same eigenvalues, (15)–(17) hold—this is an important general principle. For example, since $D = Q^{-1}AQ = Q^*AQ$, we have

$$\mathbf{w} \cdot (Q^*AQ)\mathbf{w} \leq \lambda_1 \mathbf{w} \cdot \mathbf{w}$$

and hence

$$(Q\mathbf{w}) \cdot A(Q\mathbf{w}) \leq \lambda_1 \mathbf{w} \cdot \mathbf{w} = \lambda_1 (Q\mathbf{w}) \cdot (Q\mathbf{w}).$$

But since Q is an isomorphism of $\mathbb{R}^n$, we have that the above equation is equivalent to saying that for all $\mathbf{w}'$,

$$\mathbf{w}' \cdot A\mathbf{w}' \leq \lambda_1 \mathbf{w}' \cdot \mathbf{w}'.$$

□

Let us derive two corollaries of Theorem 4.6.

Recall that if $A \in \mathcal{M}_{m,n}(\mathbb{R},\mathbb{C})$, then we define the L^2-operator norm of A to be

$$\|A\|_{L^2} = \max_{\mathbf{w}\neq\mathbf{0}} \frac{\|A\mathbf{w}\|_2}{\|\mathbf{w}\|_2}, \tag{18}$$

Corollary 4.7. *If A is real symmetric (more generally, Hermitian), then*

$$\|A\|_{L^2} = \max_i(|\lambda_i|).$$

Definition 4.8. Let $A \in \mathcal{M}_n(\mathbb{R})$ be symmetric and $\mathbf{w} \in \mathbb{R}^n$ be nonzero (or $A \in \mathcal{M}_n(\mathbb{C})$ be Hermitian and $\mathbf{w} \in \mathbb{C}^n$). We define the *Rayleigh quotient of A at $\mathbf{w}$* to be

$$\mathcal{R}_A(\mathbf{w}) = \frac{\mathbf{w}\cdot A\mathbf{w}}{\mathbf{w}\cdot\mathbf{w}}.$$

Corollary 4.9. *Let $A \in \mathcal{M}_n(\mathbb{R})$ be symmetric (or Hermitian) with eigenvalues $\lambda_n \le \ldots \le \lambda_1$. Then the maximum value of $\mathcal{R}_A$ is λ_1, and attained on eigenvectors corresponding to eigenvalue λ_1. Similarly with "minimum" and λ_n replacing "maximum" and λ_1.*

To connect "clustering" with eignevalues we will need a generalization of the above corollary, which follows from the proof of Theorem 4.6.

Theorem 4.10 (The Max-Min Principle)**.** *Let $A \in \mathcal{M}_n(\mathbb{R})$ be symmetric (or Hermitian) with eigenvalues $\lambda_n \le \ldots \le \lambda_1$. Then for $i \in [n]$,*

$$\lambda_i = \max_{W\subset\mathbb{R}^n,\,\dim(W)=i} \left(\min_{\mathbf{w}\in W,\,\mathbf{w}\neq\mathbf{0}} \mathcal{R}_A(\mathbf{w}) \right). \tag{19}$$

Similarly

$$\lambda_{n-i} = \min_{W\subset\mathbb{R}^n,\,\dim(W)=i} \left(\max_{\mathbf{w}\in W,\,\mathbf{w}\neq\mathbf{0}} \mathcal{R}_A(\mathbf{w}) \right).$$

Note that the statement about λ_{n-i} about A is equivalent to the statement about λ_i for $-A$; hence it suffices to verify the first.

Proof. First we verify this for A real diagonal: if $A = \mathrm{diag}(\lambda_1,\ldots,\lambda_n)$, then for each $W \subset \mathbb{R}^n$ of dimension i, we claim there is an $\mathbf{w} \neq \mathbf{0}$ with $\mathbf{w} \in W$ and $w_1,\ldots,w_{i-1} = 0$: indeed, if $\mathbf{u}_1,\ldots,\mathbf{u}_i$ is a basis for W, then for any $a_1,\ldots,a_i$ we have that

$$\mathbf{w} = a_1\mathbf{u}_1 + \cdots + a_i\mathbf{u}_i$$

has its first $i-1$-coordinates equal to 0 iff for $j \in [i-1]$ we have

$$0 = \mathbf{e}_j\cdot\mathbf{w} = a_1(\mathbf{e}_j\cdot\mathbf{u}_1) + \cdots + a_i(\mathbf{e}_j\cdot\mathbf{u}_i),$$

which gives $i-1$ equations linear for $a_1,\ldots,a_i$, and therefore ther is a non-trivial solution.

But for $\mathbf{w} \in W$ with $\mathbf{w} \neq \mathbf{0}$ and $w_1,\ldots,w_{i-1}$, we have

$$\mathbf{w}\cdot A\mathbf{w} = \lambda_n w_n^2 + \cdots + \lambda_i w_i^2 \le \lambda_i(\mathbf{w}\cdot\mathbf{w}).$$

Hence

$$\lambda_i \le \min_{\mathbf{w}\in W,\,\mathbf{w}\neq\mathbf{0}} \mathcal{R}_A(\mathbf{w}).$$

If $W = \mathrm{Span}(\mathbf{e}_1,\ldots,\mathbf{e}_i)$, then for all $\mathbf{w} \in W$ we have

$$\mathbf{w}\cdot A\mathbf{w} = \lambda_1 w_1^2 + \cdots + \lambda_i w_i^2 \ge \lambda_i(w_1^2 + \cdots + w_i^2) = \lambda_i \mathbf{w}\cdot\mathbf{w}.$$

Hence we have

$$\lambda_i = \min_{\mathbf{w}\in W,\, \mathbf{w}\neq \mathbf{0}} \mathcal{R}_A(\mathbf{w})$$

for this particular choice of W, and (19) follows.

The theorem follows for $A = QDQ^{-1}$, since Q and Q^{-1} act on $\mathbb{R}^n$ in a way that preserves dot products and the dimension of subspaces (or U and U^{-1} on $\mathbb{C}^n$). [Alternatively, we can repeat the same argument, replacing all occurrences of $\mathbf{e}_1, \ldots, \mathbf{e}_{i-1}, \mathbf{e}_i$ with $\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_i$ which are orthonormal eigenvectors of A with respect to the eigenvalues $\lambda_1, \ldots, \lambda_i$ (the role of the components w_i of $\mathbf{w}$ is played by $\mathbf{v}_i \cdot \mathbf{w}$, which are the component of $\mathbf{w}$ wrt the $\mathbf{v}_1, \ldots, \mathbf{v}_n$ in the sense that $\mathbf{w} = \sum_{j\in[n]} (\mathbf{v}_j \cdot \mathbf{w})\mathbf{v}_j$. □

4.4. **Regular Graphs, Expanders, and the L^2-Operator Norm.** Some of the examples of matrices we have seen at this point have been symmetric with entries that are non-negative integers. Such matrices are adjacency matrices of graphs. In order to get some intuition regarding what their eigenvalues mean, it is simplest to look at the case of *regular* graphs.

Definition 4.11. A graph, G, is d-regular for a $d \in \mathbb{N}$ if each vertex is incident upon d edges, counted with multiplicity, i.e., if each row sum (and therefore column sum) of A_G is d.

Theorem 4.12. *Let G be a d-regular graph, and let the eigenvalues of A_G be arranged*

$$\lambda_n(G) \leq \ldots \leq \lambda_2(G) \leq \lambda_1(G).$$

Then:

1. *We have $\lambda_1(G) = d$ with the constant vector $\mathbf{1}$ being a corresponding eigenvector; the multiplicity of d is the number of connected components of G;*
2. *G is* bipartite *(see below) iff $\lambda_n(G) = -d$, and if so then (n is even and) for every i we have $\lambda_i(G) = -\lambda_{n+1-i}(G)$;*
3. *setting $\mathbf{v}_1 = \mathbf{1}/\sqrt{n}$, and letting $\mathbf{v}_2, \ldots, \mathbf{v}_n$ be any orthonormal eigenbasis for A_G (with $A\mathbf{v}_i = \lambda_i \mathbf{v}_i$),*

$$A_G = \frac{d}{n}\begin{bmatrix} 1 & \ldots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \ldots & 1 \end{bmatrix} + \mathcal{E}, \quad \text{where} \quad \mathcal{E} = \sum_{i\geq 2} \lambda_i \mathbf{v}_i \mathbf{v}_i^{\mathrm{T}} \tag{20}$$

4. *hence (i) $\mathcal{E}\mathbf{1} = \mathbf{0}$, (ii) the image of $\mathcal{E}$ lies in $\mathbf{1}^\perp$ (the orthogonal complement of $\mathbf{1}$), and (iii) for any $\mathbf{v} \in \mathbb{R}^n$ we have*

$$\mathcal{E}\mathbf{v} = \mathcal{E}\big(\mathrm{Proj}_{\mathbf{1}^\perp}(\mathbf{v})\big);$$

and (iv) $\|\mathcal{E}\|_{L^2} = \max_{i\geq 2} |\lambda_i|$;

5. *for any $U, W \subset V_G$ we have*

$$|E(U,W)| = \frac{d}{n}|U|\,|W| + \mathbf{e}_U^{\mathrm{T}} \mathcal{E} \mathbf{e}_W \tag{21}$$

(where $\mathbf{e}_U, \mathbf{e}_W$) are the indicator functions of U, W, and hence

$$|\mathbf{e}_U^{\mathrm{T}} \mathcal{E} \mathbf{e}_W| \leq \big(\max_{i\geq 2} |\lambda_i|\big) \sqrt{\frac{|U|\,(n-|U|)}{n}} \sqrt{\frac{|W|\,(n-|W|)}{n}}, \tag{22}$$

(6) for any $U \subset V_G$ we have

$$\lambda_n \frac{|U|\,(n-|U|)}{n} \le \mathbf{e}_U^{\mathrm{T}} \mathcal{E} \mathbf{e}_U \le \lambda_2 \frac{|U|\,(n-|U|)}{n} \tag{23}$$

and, with $U^{\mathrm{c}} = [n] \setminus U$ (i.e., the complement of U),

$$\lambda_n \frac{|U|\,|U^{\mathrm{c}}|}{n} \le \left(-\mathbf{e}_U^{\mathrm{T}} \mathcal{E} \mathbf{e}_{U^{\mathrm{c}}}\right) \le \lambda_2 \frac{|U|\,|U^{\mathrm{c}}|}{n}. \tag{24}$$

In (20), the all 1's matrix arises from noting that $A\mathbf{1} = d(\mathbf{1})$, so setting $\mathbf{v}_1 = \mathbf{1}/\sqrt{n}$ we have

$$\lambda_1 \mathbf{v}_1 \mathbf{v}_1^{\mathrm{T}} = \frac{d}{n} \begin{bmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{bmatrix}.$$

One way to understand some global properties of "expanders" and graphs with "bad clustering" is in terms of (20): any d-regular graph is d/n times the all 1's matrix plus an error term, $\mathcal{E}$. This may seem bizarre, since if d is fixed and n large, then most of the entries of $\mathcal{E}$ are $-d/n$, and a small number are $1 - d/n$. And yet, all the formulas obtained in the above theorem result from this point of view, and the fact that this "error term" $\mathcal{E}$ lives on $\mathbf{1}^{\perp}$ and has eigenvalues $\lambda_2, \dots, \lambda_n$ there.

Studying graphs by exploiting the eigenvalues of their adjacency matrices is often called "algebraic graph theory" or "spectral graph theory" and has a rich literature. The bounds (24), (23), and (21) imply

$$\lambda_n \frac{|U|\,|U^{\mathrm{c}}|}{n} \le \left(|E(U,U)| - \frac{d}{n}|U|\,|U| \right) \le \lambda_2 \frac{|U|\,|U^{\mathrm{c}}|}{n},$$
$$\lambda_n \frac{|U|\,|U^{\mathrm{c}}|}{n} \le \left(\frac{d}{n}|U|\,|U^{\mathrm{c}}| - |E(U,U^{\mathrm{c}})| \right) \le \lambda_2 \frac{|U|\,|U^{\mathrm{c}}|}{n},$$

which can be viewed as a 2×2-block case of the "Higman-Sims technique" (see Subsection 2.9 and Exercise **??**) of [HH71] (see also [Hae78], [Hae80], Section 2.1)), although bounds based on (21)–(24) likely are implicit elsewhere. The bounds (22) and (21) imply

$$\left| |E(U,W)| - \frac{d}{n}|U|\,|W| \right| \le (\max_{i \ge 2} |\lambda_i|) \sqrt{\frac{|U|\,(n-|U|)}{n}} \sqrt{\frac{|W|\,(n-|W|)}{n}},$$

which began appearing more prominently and explicitly in the 1980's literature on *expanders* and *generalized polygons* [Tan84, AM85, AM84, Buc86, AC88]. These days the simpler implication

$$\left| |E(U,W)| - \frac{d}{n}|U|\,|W| \right| \le (\max_{i \ge 2} |\lambda_i|) \sqrt{|U|\,|W|},$$

is known as the *Expander Mixing Lemma*; see Section 2.4 of [HLW06] for a discussion and a converse of Bilu and Linial [BL06].

4.5. **Mixing Times, Refinement, and Reversibility in Markov Chains.** Consider Example 2.12 regarding the Markov matrix

$$P = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} = \begin{bmatrix} .99 & .01 \\ .02 & .98 \end{bmatrix}$$

Recall that P represents people in state s_1, meaning that the prefer The Expanse to The Mandalorian, and s_2 the reverse. Let us estimate the mixing times of this

Markov chain; we briefly recall the definition; see Sections 4.4 and 4.5 of [LP17] for details. The rough idea is as follows: P above has distinct eigenvalues $1, 0.97$, and using this a computation shows that as integer $t \to \infty$

$$\begin{bmatrix} .99 & .01 \\ .02 & .98 \end{bmatrix}^t \to \begin{bmatrix} 2/3 & 1/3 \\ 2/3 & 1/3 \end{bmatrix} = \mathbf{1}\boldsymbol{\pi}^{\mathrm{T}}$$

where $\boldsymbol{\pi}^{\mathrm{T}}$ is the stationary distibution $[2/3\ 1/3]$. The mixing time tells us how fast P^m converges to $\mathbf{1}\boldsymbol{\pi}^{\mathrm{T}}$.

4.5.1. *The Stationary Distribution.* Let $P \in \mathcal{M}_n(\mathbb{R})$ be a Markov matrix (i.e., a row stochastic matrices). The fact that $P\mathbf{1} = \mathbf{1}$ and the Perron-Frobenius theorem (see Section 6) implies that if P is irreducible (i.e., the associated digraph of positive probability transitions is strongly connected) then P has a unique eigenvector $\boldsymbol{\pi}$ such that $\boldsymbol{\pi}^{\mathrm{T}}P = \boldsymbol{\pi}^{\mathrm{T}}$ up to scalar multiple, i.e., of eigenvalue 1, and that all components of $\boldsymbol{\pi}$ are strictly positive or strictly negative (depending on which scalar multiple we choose). When $\boldsymbol{\pi}$ is chosen to be stochastic, then $\boldsymbol{\pi}$ is unique and called the *stationary distribution* of P (or of the Markov chain).

The *convergence theorem* states that as $t \in \mathbb{Z}$ tends to infinity, then $P^m \to \mathbf{1}\boldsymbol{\pi}^{\mathrm{T}}$, i.e., for all $i \in [n]$, $\mathbf{e}_i P^t \to \boldsymbol{\pi}$, provided that P is *aperiodic*; let us define this term.

Definition 4.13. If G is a digraph, then the *period* of G is the GCD (greatest common divisor) if the lengths of all closed walks in G. If $P \in \mathcal{M}_n(\mathbb{R})$, then the *period* of P is the period of the graph on vertex set $[n]$ where there is an edge from i to j when $p_{ij} > 0$. In other words, the period of P is the GCD of all k such that P^k has a nonzero diagonal element. We say that an irreducible Markov matrix—or a strongly connected graph—is *aperiodic* if its period is 1.

When working with Markov matrices or directed graphs that are not irreducible, then the period should be defined on the irreducible or strongly connected components (and the GCD over all states or all vertices is not usually of interest). If P is irreducible, we say that P is *aperiodic*

4.6. **The Convergence Theorem.** The *convergence theorem* states that if P is an aperiodic, irreducible Markov matrix, then the Perron-Frobenius theorem implies that if λ is the second largest eigenvalue of P in absolute value, then $|\lambda| < 1$; from this it will follow that

$$\max_i \|\mathbf{e}_i^{\mathrm{T}} P^t - \boldsymbol{\pi}^{\mathrm{T}}\| \leq C t^k |\lambda|^t \tag{25}$$

for some constant C, where k is the largest multiplicity of an eigenvalue of P of absolute value. (See also Theorem 4.9 of [LP17], which is slightly weaker.) In other words, $P^t \to \mathbf{1}\pi^{\mathrm{T}}$ as $t \to \infty$. However, C can be arbitrarily large, as the example below shows, and getting concrete bounds on C requires some extra information on P. Such information is provided by $\boldsymbol{\pi}$ when the Markov chain is *reversible*.

Example 4.14. Let $P \in \mathcal{M}_n(\mathbb{R})$ be given by $1 = p_{12} = p_{23} = \cdots = p_{n-1,n} = p_{nn} = 1$. (It suffices to think of n as fixed and at least 3, but n varying is also worthwhile to consider.) Then, as a block matrix,

$$P = \begin{bmatrix} J_{n-1}(0) & \\ & 1 \end{bmatrix},$$

and so P has eigenvalues 0 with (geometric) multiplicity $n-1$ and 1 with multiplicity 1, and $\mathbf{e}_1 P^k = \mathbf{e}_{k+1}$ for $k \le n-1$. If for $\epsilon > 0$—think of ϵ as very small—we set

$$P_\epsilon = P(1-\epsilon) + Q\epsilon$$

where Q is any irreducible Markov matrix with distinct eigenvlues, then for sufficiently small ϵ, the eigenvalues of P_ϵ, except for 1, are distinct and arbitrarily near 0. Also, by continuity, as $\epsilon \to 0$, any limit point of the stationary distributions of P_ϵ tends to the unique left (i.e., row) eigenvector of P with eigenvalue 1, which is $\mathbf{e}_n$. Hence, taking any $t < n-1$ in (25), for fixed $n \ge 2$, $C = C(n,\epsilon)$ in (25) tends to infinity.

4.6.1. *Mixing Time.* For any $\epsilon > 0$, we define *ϵ-mixing time* of an irreducible Markov matrix $P \in \mathcal{M}_n(\mathbb{R})$ with stationary distribution $\boldsymbol{\pi}$ as follows: in keeping with the notation of [LP17], Section 4.5, let $\mathcal{P} = \mathcal{P}_n$ denote the set of stochastic (row) vectors in $\mathbb{R}^n$. The *ϵ-mixing time of P* is the smallest $t = t_{\mathrm{mix}}(\epsilon) \in \mathbb{N}$ such that

$$\max_{\mu \in \mathcal{P}} \|\mu P^t - \pi\|_1 \le 2\epsilon, \tag{26}$$

where $\|\cdot\|_1$ is the L^1-norm, i.e., $\|\mathbf{u}\|_1 = |u_1| + \cdots + |u_n|$. Let us explain where the 2 in the 2ϵ of (26) comes from.

In Markov chains, it is more convenient to work with the total variation distance between two stochastic vectors, $\mu, \nu \in \mathcal{P}$, given by

$$\|\mu - \nu\|_{\mathrm{TV}} = \max_{A \subset [n]} |\mathbf{P}_\mu(A) - \mathbf{P}_\nu(A)|, \tag{27}$$

where $\mathbf{P}_\mu(A) = \sum_{a \in A} \mu_a$ (i.e., we view μ as giving a "probability measure" on subsets of $[n]$). Setting

$$\overline{d}(t) = \max_{\mu,\nu \in \mathcal{P}} \|\mu P^t - \nu P^t\|_{\mathrm{TV}},$$

we have $\overline{d}(t)$ is nonincreasing in t and for any $s, t \in \mathbb{N}$ we have

$$\overline{d}(t+s) \le \overline{d}(t)\overline{d}(s); \tag{28}$$

this is Lemma 4.11 of [LP17], and the proof makes use of (27). It is not hard to see that

$$\max_{\mu,\nu \in \mathcal{P}} \|\mu P^t - \nu P^t\|_{\mathrm{TV}} = \max_{i,j \in [n]} \|\mathbf{e}_i P^t - \mathbf{e}_j P^t\|_{\mathrm{TV}},$$

which gives an alternate way to express $\overline{d}(t)$.

The relation between the total variation distance and the L^1-distance is simply that

$$\|\mu - \nu\|_{\mathrm{TV}} = (1/2)\|\mu - \nu\|_1. \tag{29}$$

Of course, (28) can be stated in terms of the L^1 norm, but the factor of $1/2$ in (29) would make things more awkward to state. Moreover, the simplicity of (28) and (29) explains the factor of 2 in (26).

The usual definition of mixing time is obtained by setting

$$d(t) = \max_{\mu \in \mathcal{P}_n} \|\mu P^t - \boldsymbol{\pi}\|_{\mathrm{TV}}$$

(which similarly equals $\max_i \|\mathbf{e}_i P^t - \boldsymbol{\pi}\|_{\mathrm{TV}}$), which is of more direct interest than $\overline{d}(t)$, since we are interested in how part P^t converges to $\mathbf{1}\boldsymbol{\pi}^{\mathrm{T}}$. However $d(t)$ does not generally satisfy $d(t+s) \le d(t)d(s)$, which makes $d(t)$ more awkward to work

with than $\overline{d}(t)$. However one can easily show that $d(t) \leq \overline{d}(t) \leq 2d(t)$; it follows from (28) that $d(t+s) \leq 2d(t)d(s)$.

Definition 4.15. The *mixing time*, $t_{\text{mix}} = t_{\text{mix}}(P)$ of an irreducible, aperiodic Markov matrix, P, is the ϵ-mixing time of P with $\epsilon = 1/4$.

(This is the usual definition of mixing time, but taking any $\epsilon < 1/2$ would work.) In this way we have $\overline{d}(t_{\text{mix}}) \leq 1/2$, and hence

$$d(\ell t_{\text{mix}}) \leq \overline{d}(\ell t_{\text{mix}}) \leq 1/2^{\ell}$$

for any $\ell \in \mathbb{N}$. Hence the mixing time gives us an upper bound, for any stochastic μ, of how quickly $\mu^{\mathrm{T}} P^t$ converges to $\boldsymbol{\pi}$.

4.6.2. *A Discerning Dolphin.* Let us return to the Markov matrix

$$P = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} = \begin{bmatrix} .99 & .01 \\ .02 & .98 \end{bmatrix}$$

Consider the following though experiment: imagine that although P is observed by TV content providers, there is a more highly attuned being, say a dolphin, that can observe more information about people in these two states: namely, the dolphin observes that the people in state s_1 fall into two groups, s_1' and s_1'', which are, respectively, those people who prefer Star Trek to Star Wars (in some well-defined sense). Furthermore, the dolphin observes that each month 1% of the people in state s_1' switch to state s_1'', and vice versa. Hence the dolphin observes the Markov chain

$$P' = \begin{bmatrix} p_{1'1'} & p_{1'1''} & p_{1'2} \\ p_{1''1'} & p_{1''1''} & p_{1''2} \\ p_{21'} & p_{21''} & p_{22} \end{bmatrix} = \begin{bmatrix} .98 & .01 & .01 \\ .01 & .98 & .01 \\ .01 & .01 & .98 \end{bmatrix}$$

We know that P' is orthonormally diagonalizable, and in fact

$$P' = (.97)I + (.01)E_3,$$

where E_3 is the all 1's matrix; the dolphin knows that P' is orthonormally diagonalizable, with eigenvalues $1, .97, .97$, and that

$$(P')^m = (1/3)E_3 + (.97)^m \mathcal{E},$$

where $\mathcal{E}$ is the identity on $\mathbf{1}^{\mathrm{T}}$. It follows that for any $i \in [3]$ we have

$$\mathbf{e}_i^{\mathrm{T}} (P')^m = [1/3\ 1/3\ 1/3] + (.97)^m \mathbf{v}_i,$$

where $\mathbf{v}_i$ is the projection of $\mathbf{e}_i$ into $[1\ 1\ 1]$, and therefore $\|\mathbf{v}_i\|_2 = \sqrt{2/3}$.

The dolphin now seeks to compare the information on this 3-state Markov chain with the 2-state chain observed by TV content providers. Note that the later observe the eigenvalues $1, 0.97$ in their 2-state chain, and the dolphin observes these and one more eigenvalue (the fact that it happens to be a repeat of 0.97 is a coincidence). Put some EXERCISES here regarding refinement of Markov chains and pulling back eigenfuctions; perhaps add in covering maps of graphs.

4.6.3. *Reversible Markov Chains.* There is a simpler way to express what the dolphin of the last subsection is observing, in a way that allows to work with no extra states of the Markov chain.

Let P be an irreducible Markov matrix $P \in \mathcal{M}_n(\mathbb{R})$ with stationary distribution $\boldsymbol{\pi}$; we define the *time reversal* of P to be the Markov matrix $\tilde{P}$ given by $\hat{p}_{ij} = p_{ji}\pi_j/\pi_i$; we easily see that $\tilde{P}$ is a Markov matrix; the justification for calling $\hat{P}$ the time reversal of P is given in Proposition 1.23 of [LP17]. We say that P is *reversible* if $\hat{P} = P$, or, equivalently it satisfies

$$\pi_i p_{ij} = \pi_j p_{ji} \text{ for all } i, j \in [n] \tag{30}$$

(these equations are called the *detailed balance equations*).

For example, the stationary distribution of

$$P = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} = \begin{bmatrix} .99 & .01 \\ .02 & .98 \end{bmatrix}$$

is easily seen to be $\boldsymbol{\pi} = [2/3\ 1/3]$, so $\pi_1 P_{12} = \pi_2 P_{21}$.

Proposition 4.16. *Let $P \in \mathcal{M}_n(\mathbb{R})$ be an irreducible Markov matrix. Then if $\nu \in \mathbb{R}^n$ is any vector with positive entries such that $\nu_i p_{ij} = \nu_j p_{ji}$ for all $i, j \in [n]$. Then P is a reversible Markov matrix, and the stationary distribution of P is $\boldsymbol{\pi}$ given by*

$$\pi_i = \nu_i/(\boldsymbol{\nu} \cdot \mathbf{1}).$$

For a proof, see Exercise 4.4.

Example 4.17. Let G be a digraph on n vertices; associate to G the Markov matrix, P, such that for each i, $p_{ij} = 0$ if there is no edge from i to j, and otherwise $p_{ij} = e_{ij}/d_i$, where d_i is the degree of G and e_{ij} is the number of edges from i to j. If we further assume that G is a connected graph, then P is irreducible, and if i, j are adjacent then $d_i p_{ij} = e_{ij} = e_{ji} = d_j p_{ji}$, and otherwise $d_i p_{ij} = 0 = d_j p_{ji}$. Proposition 4.16 implies that in this case P is reversible with stationary distribution $\boldsymbol{\pi}$ given by $\pi_i = d_i/D$, where $D = \sum_i d_i$.

Note that for some G that are digraphs but not graphs, i.e., e_{ij} is not always equal to e_{ji}, the Markov matrix associated to G can still be a reversible Markov chain; e.g., if $n = 2$ and G is 100-regular with 1 edge from state 1 to state 2, and 2 edges from state 2 to state 1, then P above is the matrix

$$\begin{bmatrix} .99 & .01 \\ .02 & .98 \end{bmatrix},$$

which is a reversible Markov matrix.

Below we adopt the notation in Section 3.2 of [LP17].

Example 4.18. Let $\Psi \in \mathcal{M}_n(\mathbb{R})$ be an arbitrary irreducible Markov matrix, and $\boldsymbol{\pi}$ an arbitrary stochastic vector with all positive components. Then the *Metropolis (et al.*[5]*) matrix based on π and Ψ* (or associated Markov chain) is the Markov matrix P given by

$$p_{ij} = \psi_{ij} \min\Big(1, \pi_j \psi_{ji} / \big(\pi_i \psi_{ij}\big)\Big) \tag{31}$$

[5] In the literature, people often append other names to Metropolis to refer to this or some special case of this construction.

for $i \neq j$; since $\psi_{ij} \leq \psi_{ij}$ for $j \neq i$, there is a unique $p_{ii} \geq \psi_{ii}$ making the i-th row of P stochastic. It is easy to check (Exercise 4.9) using Proposition 4.16 that P is an irreducible Markov chain with stationary distribution $\boldsymbol{\pi}$.

Remark 4.19. If $P = I \in \mathcal{M}_n(\mathbb{R})$ (which is a Markov matrix that is not irreducible for $n \geq 2$), then (1) any stochastic π is a left eigenvector, i.e., $\boldsymbol{\pi}^{\mathrm{T}} P = \boldsymbol{\pi}^{\mathrm{T}}$, and (2) for any such $\boldsymbol{\pi}^{\mathrm{T}}$ we have $\pi_i p_{ij} = \pi_j p_{ji}$ for all i, j (simply becauase $p_{ij} = 0$ for $i \neq j$). In particular, we do need to check that a Markov matrix is irreducible before applying, say, the Metropolis algorithm. See also Exercise **??**, to see that even if P is irreducible, the Metropolis algorithm can yield slowly converging Markov chains, i.e., chains with large mixing times, especially when P has large mixing times.

Example 4.20. Consider the Metropolis Markov matrix, P, based on the Markov matrix, Φ, associated to a graph, G. Then we have $\psi_{ij} = e_{ij}/d_i$ where e_{ij} is the number of edges from i to j and d_i is the degree of i. Hence

$$p_{ij} = (e_{ij}/d_i) \min(1, \pi_j d_i/(\pi_i d_j)\big).$$

If G is a d-regular graph for some d, then also

$$p_{ij} = (e_{ij}/d_i) \min(1, \pi_j/\pi_i).$$

When minimizing a function $f\colon [n] \to \mathbb{R}$, where n is very large and $[n]$ typically represents a type of "configuration space" (e.g., all subsets of a given set, all subgraphs of a graph, all k-colourings of a graph, all matchings of a graph), one often takes π_i proportional to $\lambda^{f(i)}$; in this case $\pi_i = \lambda^{f(i)}/Z_\lambda$ where $Z_\lambda = \sum_i \lambda^{f(i)}$. While Z_λ can be impractical to compute, the Metropolis matrix only cares about the ratio of $\pi_j/\pi_i = \lambda^{f(j)-f(i)}$ (and d_i/d_j if the graph isn't regular), which makes it easy to determine $p_{ij} = p_{ij}(\lambda)$ for a given λ, i, j.

The term *simulated annealing* is used when one considers $p_{ij}(\lambda)$ for various values of λ; this is the case of trying to minimize $f\colon [n] \to \mathbb{R}$ by taking various values of λ: the extreme values are $\lambda = 1$ (so that f plays no role, and one has a random walk on G), and $\lambda = +\infty$, so that one always follows a random descent direction.

In actual annealing one heats and then cools a metal object to obtain a structure with lower "potential energy" (measured in some sense), i.e., that is "stronger;" this is analogous to *tempering* chocolate, so that a thin layer of chocolate can cover a pastry, making it crisp, shiny, and more resistant to melting by the temperature of your fingers. With metals, one often heats and cool a number of times; here the idea is that each cycle you should reach a lower potential energy upon cooling, in which case the same amount of heat at the next cycle will not move you as far. These days one often speaks of *stochastic gradient descent* as a way of similarly minimizing a function by a gradient descent that is blended with some random movement (depending on a parameter analogous to λ).

Section 3.3 of [LP17] discusses of *Glauber dynamics*.

4.7. **Symmetric Matrices and Linear Maps on Inner Product Spaces.** One important aspect of reversible Markov matrices, P, is that they are symmetric with respect to a generalized notion of dot product. Let us summarize the main points.

By an *inner product* over $V = \mathbb{R}^n$ for some $n \in \mathbb{N}$, or, more generally over any finite dimensional real vector space, V (see [HJ85, HJ13], Section 0.1, and maybe write some EXERCISES), we mean a map $V \times V \to \mathbb{R}$, denoting $(\mathbf{u}, \mathbf{v})$ by the inner product applied to $\mathbf{u}, \mathbf{v} \in V$, that is

(1) bilinear, i.e., for $\mathbf{u}_1, \mathbf{u}_2, v \in V$ and $\alpha, \beta \in \mathbb{R}$,
$$(\alpha\mathbf{u}_1 + \beta\mathbf{u}_2, \mathbf{v}) = \alpha(\mathbf{u}_1, \mathbf{v}) + \beta(\mathbf{u}_2, \mathbf{v}),$$
(2) symmetric, i.e., $(\mathbf{u}, \mathbf{v}) = (\mathbf{v}, \mathbf{u})$ for all $\mathbf{u}, \mathbf{v} \in V$; and
(3) positive definite, i.e., $(\mathbf{v}, \mathbf{v}) \geq 0$ with equality iff $\mathbf{v} = \mathbf{0}$.

For $V = \mathbb{C}^n$, or more generally any vector space, V, over $\mathbb{C}$, we give the same definition, except that we use $\langle \mathbf{u}, \mathbf{v} \rangle$ instead of $(\mathbf{u}, \mathbf{v})$ and we replace the symmetric condition with skew-symmetry condition
$$\langle \mathbf{v}, \mathbf{u} \rangle = \overline{\langle \mathbf{u}, \mathbf{v} \rangle}.$$
Hence the usual dot product over $\mathbb{R}^n$ and over $\mathbb{C}^n$ [6] are, respectively, a real and a complex inner product.

Remark 4.21. In some applications, notably relativity, one replaces the positive semidefinite condition above with the more general *non-degeneracy* condition, meaning that (1) for each $\mathbf{u} \in V$ there is some $\mathbf{w} \in V$ such that $(\mathbf{u}, \mathbf{w}) \neq 0$, or, equivalently, (2) any linear functional on V, i.e., any map $L\colon V \to \mathbb{R}$, can be represented uniquely as the map $\mathbf{v} \mapsto (\mathbf{v}, \mathbf{w})$ for some $\mathbf{w} \in V$. For example, the non-degeneracy holds for the bilinear form
$$\big((x_1, y_1, z_1, t_1), (x_2, y_2, z_2, t_2)\big) = x_1x_2 + y_1y_2 + z_1z_2 - c^2t_1t_2,$$
where c is the speed of light (and the t_i represents time, and x_i, y_i, z_i represent 3-dimensional space coordinates).

Remark 4.22. If W is a vector space, then its dual space, denoted W^*, is the set of *linear functionals* on W, i.e., maps $W \to \mathbb{R}$. If W is an inner product space, then each $\mathbf{w} \in W$ gives rise to the linear functional $\mathbf{u} \mapsto (\mathbf{u}, \mathbf{w})$, and conversely any element of W^* is of this form (this is also true if $(\cdot, \cdot)$ is non-degenerate rather than positive definite). See Exercise 4.16 for details.

The above remark—in the much wider context of Hilbert spaces (which are typically infinite dimensional)—is called the *Riesz representation theorem*.

Remark 4.23. For any linear map $L\colon V \to W$ for real inner product spaces V, W, of a real vector space, V, and any $\mathbf{w} \in W$, the map $\mathbf{v} \mapsto (L\mathbf{v}, \mathbf{w})_W$ (we use the subscript W to emphasize we are using the inner product on W) is a linear functional on V; from Remark 4.22, there is a unique vector $L^*\mathbf{v}$ such that
$$(L\mathbf{v}, \mathbf{w})_W = (\mathbf{v}, L^*\mathbf{w})_V; \tag{32}$$
we easily see that L^* is a linear map $W \to V$, and we call L^* the *adjoint* of L. (See Exercise 4.14 for an alternate proof that L^* exists.) For example, for the standard inner product on $\mathbb{R}^m$ and $\mathbb{R}^n$, if L is given by the matrix $A \in \mathcal{M}_{m,n}(\mathbb{R})$ acting on column vectors, i.e., $\mathbf{v} \mapsto A\mathbf{v}$, then L^* is given by the matrix A^{T}. Similarly for $A \in \mathcal{M}_n(\mathbb{C})$ and A^{H}.

Remark 4.24. If V is a real vector space endowed with an inner product, then V^* is generally much larger than the set of functionals of the form $\mathbf{v} \to (\mathbf{v}, \mathbf{w})$. For this reason we insist that V be finiite dimensional. By constrast, if V is a *Hilbert space*, i.e., a real vector space endowed with an inner product such that V is *complete* with respect with inner product distance, and V^* is the set of *bounded* linear functionals,

[6] Here we view $\mathbb{C}^n$ as a vector space over $\mathbb{C}$; one can view $\mathbb{C}^n$ as $2n$-dimensional vector space over $\mathbb{R}$.

then V^* is the same as the set of functionals of the forma $\mathbf{v} \to (\mathbf{v}, \mathbf{w})$ (the Ritz representation theorem). This is the usual setting of infinite dimensional inner product spaces. (The notion of *span* of a set has to be modified to mean the *closure* of the set of finite linear combinations of elements of the set.)

Definition 4.25. If V is a real inner product space, we say that a linear map $L\colon V \to V$ is *self-adjoint* (with respect to the given inner product) if $L^* = L$. Similarly for complex inner product spaces.

Example 4.26. Let $P \in \mathcal{M}_n(\mathbb{R})$ be a reversible (hence irreducible) Markov matrix, with stationary distribution $\boldsymbol{\pi} \in \mathbb{R}^n$. Define the inner product

$$(\mathbf{u}, \mathbf{v})_{\boldsymbol{\pi}} = \sum_{i=1}^{n} u_i v_i \pi_i, \tag{33}$$

and define the inner product on row vectors

$$(\mathbf{u}^{\mathrm{T}}, \mathbf{v}^{\mathrm{T}})_{1/\boldsymbol{\pi}} = \sum_{i=1}^{n} u_i v_i (1/\pi_i) \tag{34}$$

(NB: we will use the inner product (34) exclusively on row vectors). Then (see Exercise 4.10) P as an operator on column vectors (i.e., $\mathbf{v} \mapsto P\mathbf{v}$) is self-adjoint with respect to $(\cdot,\cdot)_{\boldsymbol{\pi}}$, and as an operator on row vectors is self-adjoint with respect to $(\cdot,\cdot)_{1/\boldsymbol{\pi}}$. Conversely (see Exercise 4.11) if P as an operator on column vectors is self-adjoint with respect to some inner product, or the same as an operator on row vectors, then P is reversible (and so the unique inner product with respect to which P is self-adjoint is $(\cdot,\cdot)_{\boldsymbol{\pi}}$ above.

Remark 4.27. As in [HJ85, HJ13], one can regard any real inner product over a real vector space, V, as a special case of the complex inner product

$$\langle \mathbf{u}_1 + i\mathbf{u}_2, \mathbf{v}_1 + i\mathbf{v}_j \rangle \stackrel{\text{def}}{=} (\mathbf{u}_1, \mathbf{v}_1) + i(\mathbf{u}_2, \mathbf{v}_1) - i(\mathbf{u}_1, \mathbf{v}_2) + (\mathbf{u}_2, \mathbf{v}_2)$$

defined on $V \times_{\mathbb{R}} \mathbb{C}$, which is the natural way to extend V to produce a complex vector space. From this point of view, one can mostly just work with the complex inner product.

4.8. **The Rayleigh Quotient and a "Variational" Proof of Orthonormal Eigenbases for Self-Adjoint Operators.** We now prove that a symmetric matrix has an orthonormal eigenbasis. In fact, the proof works for any self-adjoint operator on an inner product space.

Definition 4.28. Let V be a real inner product space, and $L\colon V \to V$ a linear operator. We define the *Rayleigh quotient of* L to be the real-valued function of nonzero vectors of V given by

$$\mathcal{R}_L(\mathbf{v}) \stackrel{\text{def}}{=} \frac{(L\mathbf{v}, \mathbf{v})}{(\mathbf{v}, \mathbf{v})}.$$

It is immediate that $\mathcal{R}_L(\mathbf{v})$ is invariant when we multiply $\mathbf{v}$ by a nonzero scalar.

If V is a real inner product space, then *unit vectors* are vectors $\mathbf{v} \in V$ with $(\mathbf{v}, \mathbf{v})$, and, similarly, $\mathbf{v}, \mathbf{u} \in V$ are *orthogonal* if $(\mathbf{v}, \mathbf{u}) = 0$; similarly we speak of orthonormal vectors, orthonormal bases, orthonormal eigenbases, etc.

Theorem 4.29. *Let L be a self-adjoint operator $V \to V$ on a real inner product space V. Then L has an eigenbasis that is orthonormal with respect to the inner product on V.*

To prove this theorem rigorously, we need to know the following fact; the reader is free to take this for granted, especially if they find this fact intuitively clear. If V is any *normed* $\mathbb{R}$-vector space, i.e., V is a $\mathbb{R}$-vector space endowed with a norm, $\|\cdot\|$, the *unit sphere* is the subset of V given by

$$S_1(\mathbf{0}) \stackrel{\text{def}}{=} \{\mathbf{v} \in V \mid \|\mathbf{v}\| = 1\}.$$

One similarly defines the unit sphere on any normed $\mathbb{C}$-vector space.

Lemma 4.30. *Let f be a continuous, real-valued function on the unit sphere of any normed $\mathbb{R}$-vector space, V, of finite dimension. Then f has a maximum value. Similarly for real- or complex-valued functions, on any - or $\mathbb{C}$-vector space of finite dimension.*

For a proof, see Exercise **??**. In class (in 2021) we explained why this is true for the usual Euclidean norm $\|\cdot\|_2$ on $\mathbb{R}^n$; at UBC, you would the necessary ingredients for the proof in Math 320.

Proof of Theorem 4.29. Since $\mathcal{R}_L$ is a continuous function on the vectors of unit length (i.e., $\mathbf{v} \in V$ with $(\mathbf{v},\mathbf{v}) = 1$), $\mathcal{R}_L(\mathbf{v})$ attains its maximum on some vector $\mathbf{v}$ (to prove this rigorously you need to know Lemma 4.30). We claim that if $\mathbf{u}$ is orthogonal to $\mathbf{v}$, then $L\mathbf{v}$ is also orthogonal to $\mathbf{u}$: to see this, for any $\epsilon \in \mathbb{R}$ let $\mathbf{v}_\epsilon = \mathbf{v} + \epsilon\mathbf{u}$; we have

$$(\mathbf{v}_\epsilon, \mathbf{v}_\epsilon) = (\mathbf{v},\mathbf{v}) + 2\epsilon(\mathbf{v},\mathbf{u}) + \epsilon^2(\mathbf{u},\mathbf{u}) = 1 + O(\epsilon^2)$$

while

$$(L\mathbf{v}_\epsilon, \mathbf{v}_\epsilon) = (L\mathbf{v},\mathbf{v}) + \epsilon(L\mathbf{v},\mathbf{u}) + (L\mathbf{u},\mathbf{v}) + O(\epsilon^2) = (L\mathbf{v},\mathbf{v}) + 2\epsilon(L\mathbf{v},\mathbf{u}) + O(\epsilon^2).$$

It follows that

$$\mathcal{R}_L(\mathbf{v}_\epsilon) = \frac{\big(\mathcal{R}_L(\mathbf{v}) + 2\epsilon(L\mathbf{v},\mathbf{u}) + O(\epsilon^2)\big)}{1 + O(\epsilon^2)} = \mathcal{R}_L(\mathbf{v}) + 2\epsilon(L\mathbf{v},\mathbf{u})$$

so the maximality of $\mathcal{R}_L(\mathbf{v})$ implies that $(L\mathbf{v},\mathbf{u}) = 0$. Since $L\mathbf{v}$ is orthogonal to the orthogonal complement of $\mathbf{v}$, it follows that $L\mathbf{v} = \lambda\mathbf{v}$ for some $\lambda \in \mathbb{R}$. (One can also verify that $\lambda = \mathcal{R}_L(\mathbf{v})$.)

Now let $\mathbf{v}_1, \lambda_1$ be the $\mathbf{v}, \lambda$ of the previous paragraph. Let $\mathbf{v}_2$ be the unit vector where $\mathcal{R}_L$ attains its maximum value over all vectors orthogonal to $\mathbf{v}_1$. We similarly prove that $L\mathbf{v}_2$ is orthogonal to any vector orthogonal to both $\mathbf{v}_1$ and $\mathbf{v}_2$. It follows that $L\mathbf{v}_2$ is a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$; but since $L\mathbf{v}_1$ is orthogonal to $\mathbf{v}_2$ we have that $(L\mathbf{v}_2, \mathbf{v}_1) = (\mathbf{v}_2, L\mathbf{v}_1) = 0$, and so $L\mathbf{v}_2$ is a combination of $\mathbf{v}_2$ alone, i.e., $L\mathbf{v}_2 = \lambda_2\mathbf{v}_2$.

We similar find $\mathbf{v}_3, \dots, \mathbf{v}_n$ that are eigenvectors of L with $(\mathbf{v}_i, \mathbf{v}_j) = 0$ for all $i \neq j$. □

4.9. **Another Variational Argument: The Best Rank One Approximation of A Matrix and the SVD (Singular Value Decomposition).** One obtains a singular value decomposition for any linear map $L\colon V \to W$ of real or complex inner product spaces. For each of notation, let us describe this for matrices $A \in \mathcal{M}_{m,n}(\mathbb{R})$.

For $A \in \mathcal{M}_{m,n}(\mathbb{R})$, we define its *Frobenius norm* to be

$$\|A\|_{\text{Frob}} \stackrel{\text{def}}{=} \sum_{i,j} |a_{ij}|^2$$

(where the sum is over all $i \in [m]$ and $j \in [n]$. It is easy to see that

$$\|A\|_{\text{Frob}} = \text{Trace}(AA^{\text{T}}) = \text{Trace}(A^{\text{T}}A),$$

and this turns out to be useful in simplifying many computations. More generally, $\mathcal{M}_{m,n}(\mathbb{R})$ can be endowed with the *Frobenius inner product*

$$(A, B) = \text{Trace}(A^{\text{T}}B) = \text{Trace}(A^{\text{T}}B) = \text{Trace}(B^{\text{T}}A) = \text{Trace}(BA^{\text{T}}) = \sum_{ij} a_{ij}b_{ij},$$

which is just the dot product of A and B when identified with elements of $\mathbb{R}^{mn}$. Similar norms and inner products hold for any $\mathbb{R}$ or $\mathbb{C}$ inner product spaces (see Exercise **??**).

Theorem 4.31. *Let $A \in \mathcal{M}_{m,n}(\mathbb{R})$ be a matrix. Then there exist $\mathbf{u} \in \mathbb{R}^m$ and $\mathbf{v} \in \mathbb{R}^n$ such that $\|A - \mathbf{u}\mathbf{v}^{\text{T}}\|_{\text{Frob}}$ is minimized. For any such $\mathbf{u}$ and $\mathbf{v}$ we have*

$$A\mathbf{v} = \alpha\mathbf{u},\ A^{\text{T}}\mathbf{u} = \beta\mathbf{v}, \quad \text{where} \quad \alpha = \mathbf{v}\cdot\mathbf{v},\ \beta = \mathbf{u}\cdot\mathbf{u}. \tag{35}$$

Hence $\mathbf{u}$ is an eigenvector of AA^{T} and $\mathbf{v}$ one of $A^{\text{T}}A$, both corresponding to the eigenvalue

$$\lambda = \alpha\beta = (\mathbf{v}\cdot\mathbf{v})(\mathbf{u}\cdot\mathbf{u}).$$

Moreover, we have

$$\|A - \mathbf{u}\mathbf{v}^{\text{T}}\|^2_{\text{Frob}} = \text{Trace}(AA^{\text{T}}) - \lambda,$$

and λ is the largest eigenvalue of both $A^{\text{T}}A$ and AA^{T}. If $\mathbf{x}$ is any eigenvector of $A^{\text{T}}A$ with eigenvalue ν, then $\nu \geq 0$, and if $\nu > 0$ then

(1) $\mathbf{y} = A\mathbf{x}$ is an eigenvector of AA^{T} with eigenvalue ν;
(2)

$$\min_{\gamma\in\mathbb{R}} \|A - \gamma\mathbf{y}\mathbf{x}^{\text{T}}\|^2_{\text{Frob}} = \text{Trace}(AA^{\text{T}}) - \nu,$$

and the above minimum is attained when $\gamma = 1/\|x\|_2$.

Two prove this theorem we will apply a variational principle in two different (but similar) ways.

Proof. First we claim that there exist $\mathbf{u}, \mathbf{v}$ for which $f(\mathbf{u}, \mathbf{v}) = \|A - \mathbf{u}\mathbf{v}^{\text{T}}\|_{\text{Frob}}$ is minimized: by scaling, it suffices to consider the case where $\mathbf{u}$ is a unit vector; then we show that if $\mathbf{u}$ is a unit vector, there exists a B such that $\|\mathbf{v}\| \geq B$ then $f(\mathbf{u}, \mathbf{v}) \geq 1 + f(\mathbf{0}, \mathbf{0})$ (see Exercise **??** for details). It follows that to find the minimum of $f(\mathbf{u}, \mathbf{v})$ it suffices to consider vectors vectors $\mathbf{u}, \mathbf{v}$ of some bounded length, and then the same compactness argument of Lemma 4.30 shows that the infimum of f is attained somewhere. For details, see Exercise **??**.

So consider any $\mathbf{u}, \mathbf{v}$ at which $\|A - \mathbf{u}\mathbf{v}^{\text{T}}\|_{\text{Frob}}$ is minimized. Let $\mathbf{w} \in \mathbb{R}^n$ be an arbitrary vector $\mathbf{v}$ and for $\epsilon\mathbb{R}$ let

$$g(\epsilon) = \|A - \mathbf{u}\mathbf{v}_\epsilon^{\text{T}}\|^2_{\text{Frob}}, \quad \text{where} \quad \mathbf{v}_\epsilon = \mathbf{v} + \epsilon\mathbf{w}.$$

Since $g(\epsilon)$ has a maximum at $\epsilon = 0$, we have $g'(0) = 0$. To compute $g'(0)$ we note that

$$g(\epsilon) = \text{Trace}\Big(\big(A - \mathbf{u}(\mathbf{v} + \epsilon\mathbf{w})^{\text{T}}\big)\big(A^{\text{T}} - (\mathbf{v} + \epsilon\mathbf{w})\mathbf{u}^{\text{T}}\big)\Big) = c_0 + \epsilon c_1 + \epsilon^2 c_2$$

where

$$c_1 = -\operatorname{Trace}\big((A-\mathbf{u}\mathbf{v}^{\mathrm{T}})\mathbf{w}\mathbf{u}^{\mathrm{T}}\big) - \operatorname{Trace}\big(\mathbf{u}\mathbf{w}^{\mathrm{T}}(A^{\mathrm{T}}-\mathbf{v}\mathbf{u}^{\mathrm{T}})\big),$$

and using $\operatorname{Trace}(C) = \operatorname{Trace}(C^{\mathrm{T}})$ we see that the two terms on the right are equal, and hence

$$c_1 = -2\operatorname{Trace}\big((A-\mathbf{u}\mathbf{v}^{\mathrm{T}})\mathbf{w}\mathbf{u}^{\mathrm{T}}\big).$$

Hence $c_1 = 0$ iff (for any $\mathbf{w}$)

$$\operatorname{Trace}(A\mathbf{w}\mathbf{u}^{\mathrm{T}}) = \operatorname{Trace}(\mathbf{u}\mathbf{v}^{\mathrm{T}}\mathbf{w}\mathbf{u}^{\mathrm{T}}) = (\mathbf{v}\cdot\mathbf{w})\operatorname{Trace}(\mathbf{u}\mathbf{u}^{\mathrm{T}})$$

The fact that $\operatorname{Trace}(BC) = \operatorname{Trace}(CB)$ for any $B \in \mathcal{M}_{k,\ell}(\mathbb{R})$ and $C \in \mathcal{M}_{\ell,k}(\mathbb{R})$ can be applied to both the left- and right-hand-sides above, to write these equations as

$$\operatorname{Trace}(\mathbf{u}^{\mathrm{T}}A\mathbf{w}) = (\mathbf{v}\cdot\mathbf{w})\operatorname{Trace}(\mathbf{u}^{\mathrm{T}}\mathbf{u}),$$

but these are both traces of 1×1 matrices, so these are equivalent to

$$(A^{\mathrm{T}}\mathbf{u})\cdot\mathbf{w} = (\mathbf{v}\cdot\mathbf{w})(\mathbf{u}\cdot\mathbf{u}). \tag{36}$$

Since this equation must hold for all $\mathbf{w}\in\mathbb{R}^n$, we may

(1) apply (36) to all $\mathbf{w}$ orthogonal to $\mathbf{v}$, and conclude that $A^{\mathrm{T}}\mathbf{u}$ is orthogonal to all vectors orthogonal to $\mathbf{v}$, and therefore $A^{\mathrm{T}}\mathbf{u} = \alpha\mathbf{v}$ for some $\alpha\in\mathbb{R}$, and
(2) apply (36) to $\mathbf{w}=\mathbf{v}$ to conclude $\alpha(\mathbf{v}\cdot\mathbf{v}) = (\mathbf{v}\cdot\mathbf{v})(\mathbf{u}\cdot\mathbf{u})$;

hence (36) implies

$$A^{\mathrm{T}}\mathbf{u} = \alpha\mathbf{v}, \quad \alpha = \mathbf{u}\cdot\mathbf{u}.$$

Since

$$\|A-\mathbf{u}\mathbf{v}^{\mathrm{T}}\|_{\mathrm{Frob}} = \|A^{\mathrm{T}}-\mathbf{v}\mathbf{u}^{\mathrm{T}}\|_{\mathrm{Frob}},$$

the same variational argument shows that $A\mathbf{v} = \beta\mathbf{u}$ for $\beta = \mathbf{v}\cdot\mathbf{v}$. Hence we have established (35).

In light of this, we have

$$AA^{\mathrm{T}}\mathbf{u} = A(\alpha\mathbf{v}) = \alpha\mathbf{v} = \alpha\beta\mathbf{u},$$

and similarly $A^{\mathrm{T}}A\mathbf{v} = \alpha\beta\mathbf{v}$. So $\lambda = \alpha\beta$ is an eigenvalue of both $A^{\mathrm{T}}A$ and AA^{T}.

It remains to show that λ is the largest eigenvalue of $A^{\mathrm{T}}A$ and to verify the claims regarding any eigenvector, $\mathbf{x}$ of $A^{\mathrm{T}}A$. The claims regarding $\mathbf{x}$ are an EXERCISE. It then follows that the minimum of $\|A-\mathbf{u}\mathbf{v}^{\mathrm{T}}\|_{\mathrm{Frob}}^2$ is always at least as small as $\operatorname{Trace}(AA^{\mathrm{T}})-\lambda$ where λ is the largest eigenvalue of $A^{\mathrm{T}}A$, and hence $\mathbf{u}$ must be an eigenvector corresponding to the largest eigenvalue. □

Theorem 4.32 (The SVD (Singular-Value Decomposition))**.** *Let $A\in\mathcal{M}_{m,n}(\mathbb{R})$ be a matrix, and let $s=\min(m,n)$. Then there exist orthogonal vectors $\mathbf{u}_1,\dots,\mathbf{u}_s\in\mathbb{R}^m$ and orthogonal vectors $\mathbf{v}_1,\dots,\mathbf{v}_s\in\mathbb{R}^n$ such that for any $k\in[s]$, if*

$$B_k = \mathbf{u}_1\mathbf{v}_1^{\mathrm{T}} + \cdots + \mathbf{u}_k\mathbf{v}_k^{\mathrm{T}},$$

then $f(B) = \|A-B\|_{\mathrm{Frob}}$ attains its minimum over all matrices, B of rank at most k at $B=B_k$. Furthermore, $\mathbf{v}_1,\dots,\mathbf{v}_s$ may be taken to be eigenvectors of $A^{\mathrm{T}}A$ whose corresponding eigenvalues $\lambda_1,\dots,\lambda_s$ satisfy

$$\lambda_1\ge\lambda_2\ge\cdots\lambda_s\ge 0$$

and this list includes all the nonzero eigenvalues of $A^{\mathrm{T}}A$; for each i with $\lambda_i>0$, each $\mathbf{u}_i$ is given as $(\mathbf{v}_i\cdot\mathbf{v}_i)A\mathbf{v}_i$, and each $\mathbf{u}_i$ is an eigenvector of AA^{T} of eigenvalue

λ_i as well. Furthermore, if i is such that $\lambda_i > 0$, and we normalize the $\mathbf{v}_i$ to be a unit vector, then we may take $\mathbf{u}_i = A\mathbf{v}_i$, and we have $\|\mathbf{u}_i\|_2 = \sqrt{\lambda_i}$.

The proof follows easily from Theorem 4.31 and the idea used in Theorem 4.29; see Exercise 4.13.

Remark 4.33. One can "reverse engineer" the above approach: start with any $A \in \mathcal{M}_{m,n}(\mathbb{R})$; first we note that $A^{\mathrm{T}}A$ is symmetric, and that if $A^{\mathrm{T}}A\mathbf{v} = \lambda\mathbf{v}$, then

$$\lambda(\mathbf{v}\cdot\mathbf{v}) = (A^{\mathrm{T}}A\mathbf{v})\cdot\mathbf{v} = (A\mathbf{v})\cdot(A\mathbf{v}) \geq 0,$$

and hence $\lambda \geq 0$. Let $\hat{\mathbf{v}}_1, \ldots, \hat{\mathbf{v}}_s$ be an orthornmal eigenbasis for $A^{\mathrm{T}}A$ for the nonzero eigenvalues $\lambda_1 \geq \cdots \geq \lambda_s$; let $\mathbf{u}_i = A\hat{\mathbf{v}}_i$; we see that $\mathbf{u}_i \neq \mathbf{0}$ for all $i \in [s]$ (otherwise $A^{\mathrm{T}}A\mathbf{v}_i = \mathbf{0}$ and so $\lambda_i = 0$), and let $\hat{\mathbf{u}}_i = \mathbf{u}_i/\|\mathbf{u}_i\|$. Then we easily see that $\mathbf{u}_1, \ldots, \mathbf{u}_s$ are mutually orthogonal, and hence $\hat{\mathbf{u}}_1, \ldots, \hat{\mathbf{u}}_s$ are orthonormal, and we easily see that

$$A = \sum_{i=1}^{s} \sqrt{\lambda_i}\hat{\mathbf{u}}_i\hat{\mathbf{v}}_i^{\mathrm{T}}; \tag{37}$$

The values $\sqrt{\lambda_i}$ are known as the *singular-values* of A, and the above expression for A is known as the *singular-value decomposition* of A.

Remark 4.34. If A in (37) above is symmetric, then we easily see that its singular-values are the absolute values of the eigenvalues of A, and $\hat{\vec{u}}_i = \pm\hat{\vec{\mathbf{v}}}_i$ according to the sign of the eigenvalue of A.

Remark 4.35. If $A = I$ is the $n \times n$ identity matrix, then we may take any orthonormal basis $\mathbf{v}_1, \ldots, \mathbf{v}_n$ of $\mathbb{R}^n$, and then $\mathbf{u}_i = \mathbf{v}_i$. Hence the SVD can be non-unique; similarly, when $A^{\mathrm{T}}A$ has multiple eigenvalues, the SVD is non-unique. However, if $A^{\mathrm{T}}A$ has all its non-zero eigenvalues distinct, then the $\mathbf{v}_i, \mathbf{u}_i$ corresponding to any nonzero eigenvalue λ_i of $A^{\mathrm{T}}A$ (and therefore of AA^{T}) are unique up to scaling.

EXERCISE: Generalize Theorem 4.31 and Theorem 4.32 to matrices in $\mathcal{M}_{m,n}(\mathbb{C})$.

EXERCISE (Harder): Generalize Theorem 4.31 and Theorem 4.32 to any linear transformation $\mathcal{L}\colon V \to W$ where V, W are any two inner product spaces.

4.10. **Another Variational Argument: The Least Squares Fit.** Let $(x_1, y_1), \ldots, (x_n, y_n) \in \mathbb{R}^2$ be n points that we wish to model by an equation $y = a_1 f_1(x) + \cdots + a_m f_m(x)$; more precisely, we want to find the $\mathbf{a} = (a_1, \ldots, a_m)$ at which

$$\mathrm{Error}(\mathbf{a}) \stackrel{\mathrm{def}}{=} \sum_i |y_i - a_1 f_1(x_i) - \cdots - a_m f_m(x_i)|^2$$

is minimized. It is not hard—although not immediate—that such a minimum occurs at some $\mathbf{a} \in \mathbb{R}^m$ (by contrast, a maximum of $\mathrm{Error}(\mathbf{a})$ does not exist). Assuming we have a maximum, we can write

$$\mathrm{Error}(\mathbf{a}) = |\mathbf{y} - F\mathbf{a}|_2^2 = (\mathbf{y} - F\mathbf{a})\cdot(\mathbf{y} - F\mathbf{a});$$

hence if the minimum of the above function is attained at $\mathbf{a}$, then for any $\mathbf{b} \in \mathbb{R}^m$ we have

$$g(\epsilon) = \mathrm{Error}(\mathbf{a} + \epsilon\mathbf{b})$$

has a maximum at $\epsilon = 0$. Furthermore we can write

$$g(\epsilon) = c_0 + c_1\epsilon + c_2\epsilon^2$$

where

$$c_1 = -2(\mathbf{y} - F\mathbf{a}) \cdot (F\mathbf{b}).$$

Hence $c_1 = 0$ iff

$$0 = (\mathbf{y} - F\mathbf{a}) \cdot (F\mathbf{b}) = \big(F^{\mathrm{T}}(\mathbf{y} - F\mathbf{a})\big)\mathbf{b}$$

for all $\mathbf{b}$, i.e.,

$$F^{\mathrm{T}}F\mathbf{a} = F^{\mathrm{T}}\mathbf{y}.$$

The above equation for $\mathbf{a}$ is called the *normal equations* (viewed as an $m \times m$ system of equations) for $\mathbf{a}$.

4.11. **More Variational Arguments in Calculus.** The following is a loose description of the analog of variational equations in calculus. Our discussion makes the assumption that minima exist and that all functions are as differentiable as we like.

Consider the shortest path from a point (x_1, y_1) to another point (x_2, y_2) in $\mathbb{R}^2$, and assume that $x_1 \neq x_2$. Hence we want to find the function f such that $f(x_i) = y_i$ and among those minimizes

$$\mathcal{I}(f) = \int_{x=x_1}^{x=x_2} \big(f'(x)\big)^2\, dx.$$

Assuming that this function of f is attained at $f(x)$, then for any $h(x)$ with $h(x_1) = h(x_2) = 0$, we have

$$g(\epsilon) = \mathcal{I}(f + \epsilon h)$$

has $g'(0) = 0$. A similar computation shows that

$$2\int h'(x)f'(x)\, dx = 0$$

and integrating by parts we conclude

$$\int h(x)f''(x)\, dx = 0$$

for all such h, and hence $f''(x) = 0$ for any $x \in (x_1, x_2)$, i.e., $f(x)$ is a linear function.

Similarly to minimize a more general integral

$$\mathcal{I}(f) = \int_{x=x_1}^{x=x_2} L\big(x, f(x), f'(x)\big)\, dx$$

for any function, L, of three variables, with the same conditions on f we have

$$\int \Big(L_2\big(x, f(x), f'(x)\big)h(x) + L_3\big(x, f(x), f'(x)\big)h'(x)\Big)\ \mathrm{d}x = 0,$$

where L_2, L_3 are the partial derivatives in the second and third variables, and integrating the second integral by parts we deduce that

$$\int \left(L_2\big(x, f(x), f'(x)\big) - \frac{d}{dx}L_3\big(x, f(x), f'(x)\big)\right)h(x)\ \mathrm{d}x = 0$$

and hence for any $x \in (x_1, x_2)$ we have

$$L_2\big(x, f(x), f'(x)\big) - \frac{d}{dx}L_3\big(x, f(x), f'(x)\big) = 0,$$

which are the so-called *Euler-Lagrange equations*.

4.12. **Adjoints and Self-Adjoints.** The term "self-adjoint" comes from a more general (and extremely useful) notion of the adjoint of a linear transformation. Recall that in Remark 4.23 we explained roughly why any map $L\colon U \to V$ of inner product spaces has an *adjoint*, i.e., a map $L^*\colon V \to U$ such that (32) holds, i.e.,

$$(L\mathbf{v}, \mathbf{w})_W = (\mathbf{v}, L^*\mathbf{w})_V;$$

for a proof, see Exercise 4.14 or Exercise 4.17. In the special case of $L\colon U \to U$, i.e., where $V = U$ and the inner products are the same, then L is self-adjoint iff $L = L^*$.

A lot of what we have proven in this section regarding variational methods has analogs for inner product spaces. For example, see Exercise 4.15.

4.13. **EXERCISES.**

Exercise 4.1. Let $A, B \in \mathcal{M}_n(\mathbb{R})$, and recall the definition of the L^2-operator norm of matrices (18).

4.1(a) Show that

$$\|AB\|_{L^2} \leq \|A\|_{L^2}\|B\|_{L^2}.$$

4.1(b) Show that for any $k \in \mathbb{N}$,

$$\|A^k\|_{L^2} \leq \|A\|_{L^2}^k. \tag{38}$$

4.1(c) Use Corollary 4.7 to show that (38) holds with equality if A is symmetric.

4.1(d) Show that (38) holds with strict inequality for the Jordan block $J_n(\lambda)$ with $\lambda = 0$ and all $k \geq n$.

4.1(e) Show that if $\|A\|_{L^2} < 1$, then $I + A + A^2 + A^3 + \cdots$ is *convergent* in the sense that partial sums

$$S_m = I + A + A^2 + \cdots + A^m$$

satisfy, for $m' > m$,

$$\|S_{m'} - S_m\|_{L^2} = \|A^{m+1} + A^{m+2} + \cdots + A^{m'}\|_{L^2}$$

satisfies

$$\lim_{m',m\to\infty} \|S_{m'} - S_m\|_{L^2} = 0.$$

4.1(f) Show that if $\|A\|_{L^2} < 1$, then as $m \to \infty$, $S_m(I - A) \to I$, and hence the inverse of $I - A$ exists and equals

$$(I - A)^{-1} = I + A + A^2 + A^3 + \cdots$$

4.1(g) Show that the same formula for $(I - A)^{-1}$ holds provided that for some $r \in \mathbb{N}$, $\|A^r\|_{L^2} < 1$.

Exercise 4.2. Let $n \in \mathbb{N}$ and $\mathbf{x} \in \mathbb{R}^n$.

4.2(a) Show that

$$\|\mathbf{x}\|_1 \leq \sqrt{n}\,\|\mathbf{x}\|_2 \tag{39}$$

[Hint: use the Cauchy-Schwart inequality

$$|\mathbf{x} \cdot \mathbf{y}| \leq \|\mathbf{x}\|_2\,\|\mathbf{y}\|_2$$

with a careful choice of $\mathbf{y}$]. Find a stochastic $\mathbf{x}$ for which equality holds in (39).

4.2(b) Show that

$$\|\mathbf{x}\|_2^2 \le \|\mathbf{x}\|_1^2, \tag{40}$$

and find a stochasic $\mathbf{x}$ for which equality holds in (40).

Exercise 4.3. Let G be a d-regular graph, whose eigenvalues are

$$\lambda_n \le \cdots \le \lambda_1 = d,$$

and let $\rho = \rho(G) = \max_{i\ge 2} |\lambda_i|$. In this exercise you may assume the results of Exercise 4.2.

4.3(a) For any $k \in \mathbb{N}$, let $G[k]$ be the graph whose vertex set is V_G, and that has one edge from i to j for every walk from i to j of length k in G. Show that $A_{G[k]} = A_G^k$. Show that $G[k]$ is a d^k-regular graph, the eigenvalues of $A_{G[k]}$ are just λ_i^k.

4.3(b) Show that if

$$A_G = \frac{d}{n}\begin{bmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{bmatrix} + \mathcal{E}$$

(as in (20)), then for any k we have

$$A_{G[k]} = A_G^k = \frac{d^k}{n}\begin{bmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{bmatrix} + \mathcal{E}^k,$$

and that for any $\mathbf{v}$ we have

$$\|\mathcal{E}^k \mathbf{v}\|_2 \le \rho^k \|\mathbf{v}\|_2.$$

4.3(c) For any $i, j \in [n]$ with $i \ne j$, show that

$$\|A_G^k \mathbf{e}_i - d^k \boldsymbol{\pi}\|_2 \le \rho^k$$

where $\boldsymbol{\pi} = (1/n)\mathbf{1}$ (you could slightly improve the bound ρ^k).

4.3(d) Let $P = (1/d)A_G$, so that P is the Markov matrix associated to a graph, G, and $\boldsymbol{\pi} = (1/n)\mathbf{1}$ is the stationary vector of G. Show that for any $k \in \mathbb{N}$ we have

$$\|\mathbf{e}_i^{\mathrm{T}} P^k - \boldsymbol{\pi}^{\mathrm{T}}\|_1 \le (\rho/d)^k \sqrt{n}.$$

4.3(e) Show that if $\mathbf{v} \in \mathbb{R}^n$ is a unit vector with $\mathbf{v} \cdot \mathbf{1} = 0$, there exists an $i \in [n]$ such that $|\mathbf{v} \cdot (\mathbf{e}_i - \boldsymbol{\pi})| \ge 1/\sqrt{n}$.

4.3(f) Conclude from part (e) that for some $i \in [n]$ we have

$$\|\mathbf{e}_i^{\mathrm{T}} P^k - \boldsymbol{\pi}^{\mathrm{T}}\|_2 \ge (\rho/d)^k / \sqrt{n}.$$

4.3(g) Conclude that for any $k \in \mathbb{N}$ we have

$$(\rho/d)^k/\sqrt{n} \le \max_{i,j\in[n]} \|\mathbf{e}_i^{\mathrm{T}} P^k - \boldsymbol{\pi}\|_1 \le (\rho/d)^k \sqrt{n}.$$

4.3(h) Consider a real $\epsilon > 0$; what inequalities can you conclude about $t_{\mathrm{mix}}(\epsilon)$, the ϵ-mixing time of P, in terms of $\epsilon, \rho/d, n$ based on part (g)? Show that, as a result, we have

$$t_{\mathrm{mix}}(\epsilon) = \log_{d/\rho}(1/\epsilon) \pm O(\log_{d/\rho}(n)). \tag{41}$$

Exercise 4.4. Let $P \in \mathcal{M}_n(\mathbb{R})$ be an irreducible Markov matrix, and $\nu \in \mathbb{R}^n$ be a vector with positive entries such that $\nu_i p_{ij} = \nu_j p_{ji}$ for all $i, j \in [n]$.

4.4(a) Show that $\boldsymbol{\nu}^{\mathrm{T}} P = \boldsymbol{\nu}^{\mathrm{T}}$.
4.4(b) Let $\boldsymbol{\pi} = \boldsymbol{\nu}/(\boldsymbol{\nu} \cdot \mathbf{1})$. Explain why $\boldsymbol{\pi}$ is the stationary distribution of P.

Exercise 4.5. Let $\Psi \in \mathcal{M}_n(\mathbb{R})$ be an arbitrary irreducible Markov matrix, and $\boldsymbol{\pi}$ an arbitrary stochastic vector with all positive components. For $i, j \in [n]$ with $i \neq j$, let given by p_{ij} as in (31). Assume that for all i, j, $\psi_{ij} > 0$ iff $\psi_{ji} > 0$ (this condition is missing in [LP17]).

4.5(a) Show that

$$p_{ii} = 1 - \sum_{j \neq i} p_{ij}$$

is non-negative. Conclude that if P has entries p_{ij} for $i, j \in [n]$, then P is a Markov matrix.
4.5(b) Explain why P is irreducible.
4.5(c) Apply Proposition 4.16 to show that P is reversible with stationary distribution $\boldsymbol{\pi}$. [Hint: for any i, j, we have $\pi_j \psi_{ji} \geq \pi_i \psi_{ij}$ or $\pi_j \psi_{ji} \leq \pi_i \psi_{ij}$ (or both hold).]

Exercise 4.6. Let P be the irreducible Markov matrix:

$$\begin{bmatrix} .99 & .01 \\ .04 & .96 \end{bmatrix}.$$

4.6(a) For each eigenvalue, $\lambda_1 = 1$ and $\lambda_2 = 0.95$, the a corresponding eigenvector (i.e., $P\mathbf{v} = \lambda_i \mathbf{v}_i$) and a corresponding left eigenvector $\mathbf{u}_i^{\mathrm{T}} P = \lambda_i \mathbf{u}_i^{\mathrm{T}}$.
4.6(b) Find the stationary distribution, $\boldsymbol{\pi}$, of P.
4.6(c) Verify that $\mathbf{v}_1$ is orthogonal to $\mathbf{u}_2$.
4.6(d) Verify that $\mathbf{u}_1$ is orthogonal to $\mathbf{v}_2$.
4.6(e) Verify that $\mathbf{v}_1, \mathbf{v}_2$ are orthogonal with respect to the weighted dot product

$$(\mathbf{x}, \mathbf{y})_{\boldsymbol{\pi}} \stackrel{\text{def}}{=} x_1 y_1 \pi_1 + x_2 y_2 \pi_2.$$

4.6(f) Verify that $\mathbf{u}_1, \mathbf{u}_2$ are orthogonal with respect to the weighted dot product

$$(\mathbf{x}, \mathbf{y})_{1/\boldsymbol{\pi}} \stackrel{\text{def}}{=} x_1 y_1 / \pi_1 + x_2 y_2 / \pi_2.$$

Exercise 4.7. Show that an arbitrary irreducible 2×2 Markov matrix:

$$P = \begin{bmatrix} 1 - \alpha & \alpha \\ \beta & 1 - \beta \end{bmatrix}$$

is reversible (since P is irreducible, $0 < \alpha, \beta \leq 1$).

Exercise 4.8. Let Ψ be an irreducible Markov matrix, $\boldsymbol{\pi}$ an arbitrary stochastic vector with positive components, and let P be the Metropolis chain for $\boldsymbol{\pi}$ and Ψ. For any $\epsilon \in (0, 1)$, let

$$\Psi_\epsilon = (1 - \epsilon) I + \epsilon \Psi, \quad P_\epsilon = (1 - \epsilon) I + \epsilon P \tag{42}$$

(these are lazy versions of Ψ, P). Show that the Metropolis chain for $\boldsymbol{\pi}$ and Ψ_ϵ equals P_ϵ.

Exercise 4.9. [The point of the following exercise is to show that the Metropolis algorithm can yield Markov chains that converge very slowly, i.e., with very large mixing times.] Let

$$\Psi = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}.$$

4.9(a) Find the stationary distribution, $\boldsymbol{\nu}$, of Ψ and explicitly compute $d(t) = d_\Psi(t) = \max_{i=1,2} \|\mathbf{e}_i \Psi^t - \boldsymbol{\nu}\|_{\mathrm{TV}}$ for $t = 1, 2, \ldots$

4.9(b) Let $\pi_1 = 3/4$ and $\pi_2 = 1/4$. Find the Metropolis chain, P, for $\boldsymbol{\pi}$ and Ψ.

4.9(c) For any $\epsilon \in (0,1)$, let $\Psi_\epsilon, P_\epsilon$ be as in (42). Let $N \in \mathbb{N}$ be large, and let $\epsilon = 1/N$. Give a good approximation of $d_{\Psi_\epsilon}(N\ell)$ for $\ell = 1, 2, 3$.

Exercise 4.10. Let $P \in \mathcal{M}_n(\mathbb{R})$ be a reversible (hence irreducible) Markov matrix, with stationary distribution $\boldsymbol{\pi}$. Recall the notation for the weighted dot products (33) and (34). Show that the following are equivalent:

(1) P is reversible;
(2) for all $k, \ell \in [n]$, $(P\mathbf{e}_k, \mathbf{e}_\ell)_{\boldsymbol{\pi}} = (\mathbf{e}_k, P\mathbf{e}_\ell)_{\boldsymbol{\pi}}$;
(3) for all $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, $(P\mathbf{u}, \mathbf{v})_{\boldsymbol{\pi}} = (\mathbf{u}, P\mathbf{v})_{\boldsymbol{\pi}}$;
(4) for all $k, \ell \in [n]$, $(\mathbf{e}_k^{\mathrm{T}} P, \mathbf{e}_\ell^{\mathrm{T}})_{1/\boldsymbol{\pi}} = (\mathbf{e}_k^{\mathrm{T}}, \mathbf{e}_\ell^{\mathrm{T}} P)_{1/\boldsymbol{\pi}}$,
(5) for all $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, $(\mathbf{u}^{\mathrm{T}} P, \mathbf{v}^{\mathrm{T}})_{1/\boldsymbol{\pi}} = (\mathbf{u}^{\mathrm{T}}, \mathbf{v}^{\mathrm{T}} P)_{1/\boldsymbol{\pi}}$.

Exercise 4.11. 4.11(a) Show that if P is an irreducible Markov matrix that is self-adjoint with respect to some weighted inner product, i.e., $(P\mathbf{x}, \mathbf{y})_{\mathbf{w}} = (\mathbf{x}, P\mathbf{y})_{\mathbf{w}}$ for some $\mathbf{w}$, then P is reversible.

4.11(b) Harder: if P is an irreducible Markov matrix that is self-adjoint with respect to some inner product (not necessarily a weighted inner product), is it still true that P is necessarily reversible?

Exercise 4.12. Let $(,)$ be an arbitrary inner product on $\mathbb{R}^n$, and let $\|\mathbf{x}\| = (\mathbf{x}, \mathbf{x})^{1/2}$ denote the length of a vector with respect to this inner product. The following exercise is meant to show you that a lot of things that hold for the usual dot product also hold for inner products.

4.12(a) Show that for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$,

$$(\mathbf{x}, \mathbf{y}) \le \|\mathbf{x}\| \, \|\mathbf{y}\|.$$

[Hint: Consider the fact that $\|\mathbf{x} - t\mathbf{y}\|^2 \ge 0$ for all $t \in \mathbb{R}$.]

4.12(b) Show that if $\mathbf{u}_1, \ldots, \mathbf{u}_n \in \mathbb{R}^n$ are orthonormal with respect to $(,)$ (i.e., $(\mathbf{u}_i, \mathbf{u}_j)$ equals 1 if $i = j$ and 0 if $i \ne j$), then $\mathbf{u}_1, \ldots, \mathbf{u}_n \in \mathbb{R}^n$ is a basis for $\mathbb{R}^n$, and for any $\mathbf{w} \in \mathbb{R}^n$ we have

$$\mathbf{w} = \sum_{i=1}^n (\mathbf{u}_i, \mathbf{w}) \mathbf{u}_i.$$

4.12(c) Show that Pythagoras' theorem holds for $(,)$, in the sense that if $\mathbf{u}_1, \ldots, \mathbf{u}_n$ are orthonormal as in the previous part, then

$$(\mathbf{w}, \mathbf{w}) = \sum_{i=1}^n (\mathbf{u}_i, \mathbf{w})^2.$$

4.12(d) Show that for any linearly independent $\mathbf{v}_1, \ldots, \mathbf{v}_s \in \mathbb{R}^n$, one can find linearly independent $\mathbf{u}_1, \ldots, \mathbf{u}_s$ whose span equals that of $\mathbf{v}_1, \ldots, \mathbf{v}_s$ and

such that $\mathbf{u}_1, \ldots, \mathbf{u}_s$ are orthonormal with respect to $(,)$. [Hint: Set $\mathbf{u}_1 = \mathbf{v}_1/\|\mathbf{v}_1\|$, and for successive values $i = 2, \ldots, s$ set

$$\mathbf{v}_i' = \mathbf{v}_i - (\mathbf{u}_1, \mathbf{v}_i)\mathbf{u}_1 - \ldots - (\mathbf{u}_{i-1}, \mathbf{v}_i)\mathbf{u}_{i-1}$$

and $\mathbf{u}_i = \mathbf{v}_i'/\|\mathbf{v}_i'\|$; this is called the *Gram-Schmidt process.*]

Exercise 4.13. The point of this exercise is to prove Theorem 4.32 as a consequence of the ideas of Theorem 4.31. Let $A \in \mathcal{M}_{m,n}(\mathbb{R})$. Let $A \in \mathcal{M}_{m,n}(\mathbb{R})$ with $A \neq 0$ (i.e., A is not the zero matrix). Let the minimum of $f(\mathbf{u}, \mathbf{v}) = \|A - \mathbf{u}\mathbf{v}^{\mathrm{T}}\|_{\mathrm{Frob}}$ be attained at $\mathbf{u} = \mathbf{u}_1$ and $\mathbf{v} = \mathbf{v}_1$, and set

$$A_1 = A - \mathbf{u}_1\mathbf{v}_1^{\mathrm{T}}.$$

Assume that $A_1 \neq 0$.

4.13(a) Use Theorem 4.31 to show that $A_1\mathbf{v}_1 = \mathbf{0}$ and that $A_1^{\mathrm{T}}A_1\mathbf{v}_1 = \mathbf{0}$; show similar statements involving $\mathbf{u}_1$.

4.13(b) Let the minimum of $f(\mathbf{u}, \mathbf{v}) = \|A_1 - \mathbf{u}\mathbf{v}^{\mathrm{T}}\|_{\mathrm{Frob}}$ at $\mathbf{u} = \mathbf{u}_2$ and $\mathbf{v} = \mathbf{v}_2$. Show that $\mathbf{v}_2$ is orthogonal to $\mathbf{v}_1$. Then show that $\mathbf{u}_2$ is orthogonal to $\mathbf{u}_1$.

4.13(c) Show that $A^{\mathrm{T}}A\mathbf{v}_2 = \lambda_2 v_2$, and $AA^{\mathrm{T}}\mathbf{u}_2 = \lambda_2\mathbf{u}_2$.

4.13(d) Show that

$$\|A - \mathbf{u}_1\mathbf{v}_1^{\mathrm{T}} - \mathbf{u}_2\mathbf{v}_2^{\mathrm{T}}\|_{\mathrm{Frob}}^2 = \mathrm{Trace}(A^{\mathrm{T}}A) - \lambda_1 - \lambda_2.$$

4.13(e) Show that continuing in this way, we have that for some $r \leq n$

$$A = \sum_{i=1}^{r} \mathbf{u}_i\mathbf{v}_i^{\mathrm{T}}, \tag{43}$$

where $\mathbf{u}_1, \ldots, \mathbf{u}_r$ are nonzero and mutually orthogonal and $\mathbf{v}_1, \ldots, \mathbf{v}_r$ are as well, and show that r is the rank of A (i.e., the dimension of the image of A).

4.13(f) Show that for any $k \leq r$ we have

$$\|A - \mathbf{u}_1\mathbf{v}_1^{\mathrm{T}} - \cdots - \mathbf{u}_k\mathbf{v}_k^{\mathrm{T}}\|_{\mathrm{Frob}}^2 = \mathrm{Trace}(A^{\mathrm{T}}A) - \lambda_1 - \cdots - \lambda_k,$$

where λ_i is the i-th largest eigenvalue of $A^{\mathrm{T}}A$.

4.13(g) Show that if

$$f(\mathbf{w}_1, \ldots, \mathbf{w}_k, \mathbf{z}_1, \ldots, \mathbf{z}_k) = \|A - \mathbf{w}_1\mathbf{z}_1^{\mathrm{T}} - \cdots - \mathbf{w}_k\mathbf{z}_k^{\mathrm{T}}\|_{\mathrm{Frob}}^2$$

is attained at some particular values of $\mathbf{w}_1, \ldots, \mathbf{w}_k$ and $\mathbf{z}_1, \ldots, \mathbf{z}_k$, and if $\mathbf{z}_1, \ldots, \mathbf{z}_k$ are nonzero and mutually orthogonal, as well as $\mathbf{w}_1, \ldots, \mathbf{w}_k$, then f there equals $\mathrm{Trace}(A^{\mathrm{T}}A)$ minus some k eigenvalues of $A^{\mathrm{T}}A$ (where each eigenvalue appears no more times than its multiplicity as an eigenvalue of $A^{\mathrm{T}}A$). [Hint: it suffices to show that for all $i \in [k]$, (35) holds for $\mathbf{u} = \mathbf{w}_i$ and $\mathbf{v} = \mathbf{z}_i$, and then use the previous part. So fix an $i \in [k]$ and consider f above with a variation in $\mathbf{z}_i$ alone: i.e., choose a $\mathbf{u} \in \mathbb{R}^n$, set $\mathbf{z}_i(\epsilon) = \mathbf{z}_i + \epsilon\mathbf{u}$, let

$$g(\epsilon) = f(\mathbf{w}_1, \ldots \mathbf{w}_k, \mathbf{z}_1, \ldots, \mathbf{z}_{i-1}, \mathbf{z}_i(\epsilon), \mathbf{z}_{i+1}, \ldots, \mathbf{z}_k),$$

and determine c_1 where $g(\epsilon) = c_0 + c_1\epsilon + c_2\epsilon^2$; it may be notationally easier to write

$$g(\epsilon) = \|\tilde{A} - \mathbf{w}_i\mathbf{z}_i(\epsilon)^{\mathrm{T}}\|_{\mathrm{Frob}}^2, \quad \text{where} \quad \tilde{A} = A - \sum_{j \neq i} \mathbf{w}_j\mathbf{z}_j^{\mathrm{T}}.$$

Can you assert that $c_1 = c_1(\mathbf{u})$ must equal zero assuming that $\mathbf{u}$ is orthogonal to $\mathbf{z}_1, \dots, \mathbf{z}_k$? Why? What about if $\mathbf{u}$ is a linear combination of $\mathbf{z}_1, \dots, \mathbf{z}_k$? Why? You should be able to conclude that $\mathbf{w}_i^{\mathrm{T}} \tilde{A} = \mathbf{w}_i^{\mathrm{T}} A$ must equal $\mathbf{z}_i^{\mathrm{T}}(\mathbf{w}_i \cdot \mathbf{w}_i)$.]

4.13(h) Show the function $f(A) = \|A - B\|_{\mathrm{Frob}}^2$ attains its minimum over all B of rank at most $k \le n$ at

$$B = \sum_{i=1}^{k} \mathbf{u}_i \mathbf{v}_i^{\mathrm{T}}.$$

[Hint: according to (43), if B is of rank k then

$$B = \sum_{i=1}^{k} \mathbf{x}_i \mathbf{y}_i^{\mathrm{T}}$$

where $\mathbf{x}_1, \dots, \mathbf{x}_r$ are nonzero and mutually orthogonal and $\mathbf{y}_1, \dots, \mathbf{y}_r$ are as well.]

4.13(i) Show that if $\hat{\mathbf{u}}_i = \mathbf{u}_i / |\mathbf{u}_i|_2$ and $\hat{\mathbf{v}}_i = \mathbf{v}_i / |\mathbf{v}_i|_2$, then (37) holds, i.e.,

$$A = \sum_{i=1}^{s} \sqrt{\lambda_i} \hat{\mathbf{u}}_i \hat{\mathbf{v}}_i^{\mathrm{T}}.$$

Exercise 4.14. Show the existence of L^* as in (32) in the setting there. [Hint: To simiplify things, let $\mathbf{v}_1, \dots, \mathbf{v}_n$ be a basis for V, and $\mathbf{w}_1, \dots, \mathbf{w}_m$ one for W. Show that (32) holds iff for all $i \in [n]$ and $j \in [m]$, (32) for $\mathbf{v} = \mathbf{v}_i$ and $\mathbf{w} = \mathbf{w}_j$. This allows you to describe L^*.]

Exercise 4.15. Let V, U be two $\mathbb{R}$-inner product spaces; i.e., V is an n-dimensional $\mathbb{R}$-vector space for some n (there is no harm in assuming that $V = \mathbb{R}^n$, but no need to do so) endowed with an inner product $(,)_V$, and similarly U is an m-dimensional real inner product space. Prove that Theorem 4.31 can be generalized to this context, for any linear map $L\colon U \to V$. Your generalization should include the following steps.

(1) If $\mathbf{v} \in V$, explain why it makes sense to view $\mathbf{v}$ as a map $\mathbb{R} \to V$, and use Exercise 4.14 to explain that $\mathbf{v}^*$ should be the map from $V \to \mathbb{R}$ given by $v' \in V$ maps to $u(v, v')_V$.
(2) Your generalization should involve $L^*\colon V \to U$ as in Exercise 4.14.
(3) The Frobenius norm of $L\colon U \to V$ with respect to the inner products on U and V can be defined to be

$$\|L\|_{\mathrm{Frob}} = \mathrm{Trace}(LL^*).$$

(4) From your proof it should follow that the largest eigenvalue of LL^* is the same as that of L^*L.

Exercise 4.16. Show that if W is an inner product space, then each element of W^* is of the form $\mathbf{u} \mapsto (\mathbf{u}, \mathbf{w})$ for some $w \in W$. [Hint: there are a number of ways of doing this: one way is to first find an orthonormal basis $w_1, \dots, w_n$ of W; another way is to show that W^* is n-dimensional, and then to show that maps of the form $\mathbf{u} \mapsto (\mathbf{u}, \mathbf{w})$ are a subspace of dimension n.]

Exercise 4.17. Exercise 4.14 shows that any map $L\colon U \to V$ of real inner product spaces has an adjoint. Use Exercise 4.16 and the approach of Remark 4.22 to give second proof of this fact.

Exercise 4.18. Consider the matrices

$$D_1 = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 3 \end{bmatrix}, \quad D_2 = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

4.18(a) Determine the critical points of the Rayleigh quotient

$$\mathcal{R}_{D_1}(x, y, z) = \frac{5x^2 + 4y^2 + 3z^2}{x^2 + y^2 + z^2}$$

by computing the gradient of this function and seeing where it equals $(0, 0, 0)$.

4.18(b) Do the same for $\mathcal{R}_{D_2}$.

4.18(c) Imagine that you take the Rayleigh quotient of a general $n \times n$ matrix and compute its maximum (or minimum) by taking the gradient and setting it to $\mathbf{0}$. How would this compare to the computation performed in our proof of Theorem 4.29?

For the near future: add some exercises on the lazy Ehrenfest model and hypercube, referring to Section 2.3 of [LP17].

5. Symmetric Matrices are Orthonormally Diagonalizable with Real Eigenvalues: Two Short Proofs

In this section we prove that symmetric matrices have an orthonormal eigenbasis and have real eigenvectors. Both proofs give a more general result, and indicate much more general results.

5.1. Proof via Symmetric (and Hermitian) Matrices with Distinct Eigenvalue.

This proof in this subsection also hints as to how we can get a pair of *biorthogonal* bases of left and right eigenvectors for *any* matrix $A \in \mathcal{M}_n(\mathbb{R})$.

Recall that $\langle \mathbf{u}, \mathbf{v} \rangle$ denotes—for now—the standard complex inner product $\mathbf{v}^{\mathrm{H}}\mathbf{u}$, and hence for any $A \in \mathcal{M}_{m,n}(\mathbb{C})$ and $\mathbf{u} \in \mathbb{C}^n$, $\mathbf{v} \in \mathbb{C}^m$ we have

$$\langle A\mathbf{u}, \mathbf{v} \rangle = \mathbf{v}^{\mathrm{H}} A \mathbf{u} = (A^{\mathrm{H}}\mathbf{v})^{\mathrm{H}}\mathbf{v} = \langle \mathbf{u}, A^{\mathrm{H}}\mathbf{v} \rangle.$$

Lemma 5.1. *Let $A \in \mathcal{M}_n(\mathbb{C})$ be arbitrary, with an eigenpair $A\mathbf{u} = \lambda\mathbf{u}$ and left eigenpair $\mathbf{v}^{\mathrm{T}}A = \nu\mathbf{v}^{\mathrm{T}}$ (or, equivalently $A^{\mathrm{H}}\mathbf{v} = \nu\mathbf{v}$). Then*

$$(\lambda - \overline{\nu})\langle \mathbf{u}, \mathbf{v} \rangle = 0.$$

In particular,

1. *if $\lambda \neq \overline{\nu}$, then $\langle \mathbf{u}, \mathbf{v} \rangle = 0$;*
2. *if $A^{\mathrm{H}} = A$, i.e., A is Hermitian, in particular real symmetric, then its eigenvalues are real;*
3. *if $A^{\mathrm{H}} = A$, then eigenvectors of distinct eigenvales are orthogonal.*

Proof. The first claim follows from

$$\lambda\langle \mathbf{u}, \mathbf{v} \rangle = \langle A\mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{u}, A^{\mathrm{H}}\mathbf{v} \rangle = \overline{\nu}\langle \mathbf{u}, \mathbf{v} \rangle.$$

The second claim follows when $\mathbf{u} = \mathbf{v}$ (and hence $\lambda = \nu$), since then $\langle \mathbf{u}, \mathbf{v} \rangle \neq 0$. The third claim follows since λ, ν are purely real. □

Theorem 5.2. *Let $A \in \mathcal{M}_n(\mathbb{C})$ be Hermitian. Then A has an orthonormal eigenbasis, which is unique up to scaling when A has distinct eigenvalues.*

Proof. Lemma 5.1 implies that if A has distinct eigenvalues, then any eigenbasis normalized to have all unit vectors is an orthonormal eigenbasis. Conversely, in this case any eigenbasis of A must arise from the n unique eigenvectors, and is therefore unique up to scaling each vector (i.e., by a complex number of absolute value 1).

If A does not have distinct eigenvalues, we take a sequence, $A_1, A_2, \dots$ of symmetric or Hermitian matrices tending to A with distict eigenvalues. Since the set of (ordered) orthonormal eigenbases is compact, by passing to a subsequence the eigenbases of the A_i converges to an eigenbasis, $\mathbf{x}_1, \dots, \mathbf{x}_n$, which by continuity is an orthonormal eigenbasis for A. □

Exercise 5.1. Let $A \in \mathcal{M}_2(\mathbb{R})$ be all 0's. Then A does not have distinct eigenvalues. What happens, as real $\epsilon \to 0$, to the eigenvalues/vectors of the following approximations of A:

$$\begin{bmatrix} \epsilon & 0 \\ 0 & 2\epsilon \end{bmatrix}, \quad \begin{bmatrix} 0 & \epsilon \\ \epsilon & 0 \end{bmatrix}, \quad \begin{bmatrix} \epsilon & \epsilon \\ \epsilon & \epsilon \end{bmatrix} \quad ?$$

5.2. **Schur Decomposition and Matrices that are Normal (Unitary, Hermitian, Skew-Hermitian, etc.)** Recall that if $\mathbf{u}_1, \dots, \mathbf{u}_n$ are an orthonormal basis of $\mathbb{R}^n$ or $\mathbb{C}^n$, then the matrix U whose columns are $\mathbf{u}_1, \dots, u_n$ satisfies $UU^{\mathrm{H}} = U^{\mathrm{H}}U = I$, or equivalently $U^{-1} = U^{\mathrm{H}}$; we call such a U a *unitary matrix* in the complex case, and *orthogonal matrix* in the real (special) case.

Theorem 5.3 (Schur Decomposition)**.** *Let $A \in \mathcal{M}_n(\mathbb{C})$. Then there exists unitary matrix, U, such that $A = UTU^{\mathrm{H}}$, where T is an upper triangular matrix, T (i.e., $t_{ij} = 0$ for $i > j$). If, in addition, $A \in \mathcal{M}_n(\mathbb{R})$ has all real eigenvalues, then the U is (real) orthogonal.*

Writing A as UTU^{H} as above is called a *Schur decomposition* (sometimes *Schur factorization*) of A.

Proof. We use induction on n; for $n = 1$ this is clear. For $n \geq 2$, let $T\mathbf{u} = \lambda\mathbf{u}$ be any eigenpair of T, and extend $\mathbf{u}$ to an orthonormal basis $\mathbf{u} = \mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ of $\mathbb{C}^n$. Let $\tilde{U}$ be the matrix whose columns are $\mathbf{u}_1, \dots, \mathbf{u}_n$. Then

$$A\tilde{U} = \tilde{U}\tilde{T}, \quad \text{where} \quad \tilde{T} = \begin{bmatrix} \lambda & \tilde{t}_{12} & \cdots & \tilde{t}_{1n} \\ 0 & \tilde{t}_{22} & \cdots & \tilde{t}_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \tilde{t}_{n2} & \cdots & \tilde{t}_{nn} \end{bmatrix}$$

Now we apply the inductive argument to the matrix, T_1 whose entries are $\tilde{t}_{ij}$ with $2 \leq i, j \leq n$.

If, in addition, A is real, and has all real eigenvalues, then in the inductive step above λ and U_1 can be taken to be real; since the first column of $\tilde{T}$ has 0's after the first entry λ, expanding determinants by columns shows that

$$p_{\tilde{T}}(t) = (t - \lambda)p_{T_1}(t).$$

Hence T_1 is real and has real eigenvalues, which gives the stronger inductive hypothesis in case $A \in \mathcal{M}_n(\mathbb{R})$ has all real eigenvalues. □

Definition 5.4. We say that $N \in \mathcal{M}_n(\mathbb{R}, \mathbb{C})$ is *normal* if $NN^{\mathrm{H}} = N^{\mathrm{H}}N$, i.e., if N and N^{H} commute.

Clearly, the following matrices $A \in \mathcal{M}_n(\mathbb{R}, \mathbb{C})$ are normal:

(1) (1) real symmetric, and more generally, Hermitian matrices ($A^{\mathrm{H}} = A$);
(2) permutation matrices (A whose columns are a permutation of $e_1, \ldots, e_n$, and whose rows are therefore the inverse permutation of $e_1^{\mathrm{T}}, \ldots, e_n^{\mathrm{T}}$), and more generally, orthogonal, and more generally, unitary matrices (all satisfy $A^{-1} = A^{\mathrm{H}}$);
(3) real, skew symmetric, and more generally, (complex) skew Hermitian matrices ($A^{\mathrm{H}} = -A$);
(4) etc.

Lemma 5.5. *Let $T \in \mathcal{M}_n(\mathbb{C})$ be upper triangular and satisfy $TT^{\mathrm{H}} = T^{\mathrm{H}}T$. Then T is a diagonal matrix.*

Proof. TT^{H} is the dot product of the rows of T, and $T^{\mathrm{H}}T$ that of the columns of T. Since the first column of T has only one nonzero element, t_{11}, and the dot product of this column with itself, namely $t_{11}\overline{t_{11}}$ equals the dot product of the first row with itself, it follows that $t_{1j} = 0$ for $j \geq 2$. Given this, and in particular $t_{12} = 0$, we see that the second column now only has t_{22} as a non-zero element, and so, similarly, $t_{2j} = 0$ for $j \neq 2$. Hence the second row of T is zero except in the diagonal place. Proceeding similarly by induction on i, we see that $t_{ij} = 0$ for $j \neq i$. □

Theorem 5.6. *Let $N \in \mathcal{M}_n(\mathbb{C})$ be a normal matrix. Then N is diagonalizable with an orthonormal eigenbasis. If, in addition, $N \in \mathcal{M}_n(\mathbb{R})$, and N has only real eigenvalues, then the eigenbasis can be taken to be all real.*

Proof. Let $N = UTU^{\mathrm{H}}$ be a Schur decomposition. Since $U^{\mathrm{H}} = U^{-1}$, we have

$$NN^{\mathrm{H}} = UTT^{\mathrm{H}}U^{\mathrm{H}}, \quad N^{\mathrm{H}}N = UT^{\mathrm{H}}TU^{\mathrm{H}},$$

and hence $TT^{\mathrm{H}} = T^{\mathrm{H}}T$. Now apply Lemma 5.5, and then apply the second part of Theorem 5.3 for the statement regarding real matrices with real eigenvalues. □

6. The Perron-Frobenius Theorem

In this section we state and prove the Perron-Frobenius theorem.

6.1. **Periodicity.**

Definition 6.1. If G is a strongly connected digraph, then the *period* of G is the GCD (greatest common divisor) of the lengths of all closed walks of G. If $A \in \mathcal{M}_n(\mathbb{R}, \mathbb{C})$, then the *period* of A is the period of the digraph associated to A (i.e., whose vertices are $[n]$, and with an edge from i to j if $a_{ij} \neq 0$). We say that G and A above are *aperiodic* if their period equals one.

For example, if G is the directed cycle of length n, whose adjacency matrix is the "cyclic shift" matrix C_n of Example 3.10, then the period of G and of $A_G = C_n$ is n.

The following result is not difficult.

Proposition 6.2. *If G is a strongly connected graph of period p, then we may partition the vertices of G into sets $V_0, \ldots, V_{p-1}$ such that any edge whose tail lies in V_i has its head in V_{i+1}, understanding V_p to refer to V_0. Furthermore, for any $k \in \mathbb{N}$ that is sufficient large and divisible by p, in $v, v' \in V_i$ then there is a walk from v to v' of length k.*

The proof of the prosition is an EXERCISE. The idea is that if $v, v' \in V$, then

$$\{k \in \mathbb{N} \mid \text{there is a walk of length } k \text{ from } v \text{ to } v'\}$$

are in a single class modulo p. So to define the V_i we fix some $v \in V$ and declare $v' \in V_i$ if the above lengths k modulo p equal i.

If $A \in \mathcal{M}_n(\mathbb{R}, \mathbb{C})$ has period p, then partitioning $[n]$ into $V_0, \dots, V_{p-1}$ as above we can view A as a "block cyclic matrix,"

$$A = \begin{bmatrix} 0 & A_0 & & & \\ & 0 & A_1 & & \\ & & \ddots & \ddots & \\ & & & 0 & A_{p-1} \\ A_p & & & & 0 \end{bmatrix}.$$

We easily make the following conclusion.

Proposition 6.3. *Let $A \in \mathcal{M}_n(\mathbb{R}, \mathbb{C})$ has non-negative entries and be of period p, and let $V_0, \dots, V_{p-1}$ be a partition of $[n]$ such that $a_{ij} = 1$ and $i \in V_k$ implies that $j \in V_{k+1}$. Let ζ be a primitive p-th root of unity, i.e., $\zeta^p = 1$ and $\zeta^m \neq 1$ for $m \in \mathbb{N}$ with $m < p$. Then if $\mathbf{v}, \lambda$ are an eigenpair for A, then so is $\mathbf{v}', \lambda'$, where $\lambda' = \zeta\lambda$ and $v'_j = v_j \zeta^m$ for $j \in V_i$.*

The case $p = 1$ in the above propisition is worth noting separately.

Corollary 6.4. *Let $A \in \mathcal{M}_n(\mathbb{R}, \mathbb{C})$ have non-negative entries and be irreducible and aperiodic. Then for all $k \in \mathbb{N}$ sufficiently large, A^k has all positive entries.*

6.2. **Definition of the Perron-Frobenius Eigenvalue.**

Definition 6.5. Let $A \in \mathcal{M}_n(\mathbb{R})$ have non-negative entries. We define the *Perron-Frobenius eigenvalue* of A, denoted $\lambda_{\mathrm{PF}}(A)$ to be

$$\lambda_{\mathrm{PF}}(A) = \sup\{\lambda \in \mathbb{R} \mid A\mathbf{v} \geq \lambda\mathbf{v} \text{ for some nonzero } \mathbf{v} \geq \mathbf{0}\}.$$

Any nonzero $\mathbf{v} \geq \mathbf{0}$ such that $A\mathbf{v} \geq \lambda_{\mathrm{PF}}\mathbf{v}$ is called a *Perron Frobenius pseudo-eigenvector* of A.

As the names suggest, λ_{PF} turns out to be an eigenvalue of A, and a pseudo-eigenvector is not necessarily an eigenvector. However, we are most interested in A that are irreducible, i.e., whose associated directed graph is strongly connected. It is useful to consider a case where A is not irreducible.

Notice that if each row of A has at least one positive entry, then $A\mathbf{1}$ has positive components, and hence $\lambda_{\mathrm{PF}} > 0$.

6.3. **The Fundamental Non-Example.**

The following "non-example" will illustrate how the proof we give of the Perron-Frobenius works.

It illustrates that if A is reducible, then the Perron-Frobenius pseudo-eigenvector is not generally unique.

Example 6.6. Let

$$A = \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix},$$

which in previous sections we have denoted $J_2(3)$. In this case A is not irreducible, since in the graph associated to A, which has underlying vertex set $\{1, 2\}$, there is no edge from 2 to 1. In this case we easily see that $\lambda_{\mathrm{PF}}(A) = 3$, although

for any $v_1, v_2 \geq 0$ with at least one of them non-zero, we have that $\mathbf{v} \neq \mathbf{0}$ is a pseudo-eigenvector, since

$$A \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = A \begin{bmatrix} 3v_1 + v_2 \\ 3v_2 \end{bmatrix} \geq 3 \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}.$$

Hence if $v_2 = 0$ we get a true eigenvector, but otherwise the first component of $A\mathbf{v}$ is strictly larger than that of $3\mathbf{v}$.

Notice that we alter the above matrix A by making a_{21} positive, then $v_1 > 0$ would imply that $(A\mathbf{v})_2 = a_{21}v_1 + a_{22}v_2$ would be strictly larger than $3v_2$. In this case λ_{PF} could no longer equal 3.

One can view one part of the Perron-Frobenius theorem as saying that if $\lambda_{\mathrm{PF}}(A) > 0$, then the geometric or algebraic multiplicity of λ_{PF} as an eigenvalue can only be greater than one if A is reducible.

6.4. **A Perron-Frobenius Pseudo-Eigenvector Exists.**

Theorem 6.7. *Let $A \in \mathcal{M}_n(\mathbb{R})$ have non-negative entries; then there exists a Perron-Frobenius pseudo-eigenvector, $\mathbf{v}$, of A.*

The argument is a simple *compactness* argument.

Proof. The condition that nonzero $\mathbf{v} \geq 0$ satisfies $A\mathbf{v} \geq \lambda\mathbf{v}$ for some λ is invariant under a positive scaling of $\mathbf{v}$. Hence we may assume that $\mathbf{v}$ is stochastic (instead requiring $|\mathbf{v}|_p = 1$ for any $p \geq 1$ would work equally well). By basic facts about the least upper bound (usually discussed in Math 320 at UBC), there exist λ_m and stochastic $\mathbf{v}_m$ with $A\mathbf{v}_m \geq \lambda_m\mathbf{v}_m$ and $\lambda_m \to \lambda_{\mathrm{PF}}$ as $m \to \infty$. By compactness of the set of stochasic vectors, by passing to a subsequence we may assume that $\mathbf{v}_m$ has a limit $\mathbf{v}$. It then follows that as $m \to \infty$, $A\mathbf{v}_m \to A\mathbf{v}$ and $\lambda_m\mathbf{v}_m \to \lambda_{\mathrm{PF}}\mathbf{v}$, and hence $A\mathbf{v}_m \geq \lambda_m\mathbf{v}_m$ implies that $\mathbf{v}$ is a Perron-Frobenius pseudo-eigenvector. □

6.5. **The Aperiodic Perron-Frobenius Theorem.**

The Perron-Frobenius theorem is far simpler to state and prove for matrices, A, that are aperiodic, in which case for some $m \in \mathbb{N}$ we have that A^m has all positive entries. Notice that for such m, $A^{m+1} = AA^m$ also has all positive entries, and hence also $A^{m+2}, A^{m+3}, \ldots$ Notice also that $A^m\mathbf{1}$ has all positive entries, and hence setting $\mathbf{w} = A^m\mathbf{1}$, we have $A\mathbf{w} \geq \delta\mathbf{w}$ for some $\delta > 0$, and hence $\lambda_{\mathrm{PF}}(A) > 0$.

Discussing this special case of aperiodic matrices illustrates most of the main ideas of the Perron-Frobenius theorem. One constant fact that we will use is that if A is any matrix with real, non-negative entries, and if $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ with $\mathbf{x} - \mathbf{y} \geq \mathbf{0}$, then since A has non-negative entries we have $A(\mathbf{x} - \mathbf{y}) \geq \mathbf{0}$; in other words,

$$\mathbf{x} \geq \mathbf{y} \quad \Rightarrow \quad A\mathbf{x} \geq A\mathbf{y}. \tag{44}$$

Another useful fact is that the triangle inequality for $x, y \in \mathbb{C}$,

$$|x + y| \leq |x| + |y|$$

holds with strict inequality when x, y are nonzero, unless x, y have the same *complex argument*, i.e., "they point in the same direction," or more precisely $x = \alpha y$ where α is a positive real number. It follows that for any $x_1, \ldots, x_n \in \mathbb{C}$,

$$|x_1 + \cdots + x_n| \leq |x_1| + \cdots + |x_n|$$

when and only when all the nonzero x_i have the same complex argument; in particular, if one of $x_1, \dots, x_n$ is a positive real, then all of them must be real and non-negative.

Theorem 6.8. *Let $A \in \mathcal{M}_n(\mathbb{R})$ have positive entries. Then*

1. *any Perron-Frobenius pseudo-eigenvector of A, $\mathbf{v}$, is an actual eigenvector of A (with eigenvalue λ_{PF});*
2. *if $\mathbf{v}$ is an eigenvector of A with eigenvalue λ_{PF}, and if $\mathbf{v}$ is scaled so that it has at least one, real, positive component, than all components of $\mathbf{v}$ are positive reals;*
3. *any eigenvector, $\mathbf{v}$, of A with eigenvalue λ_{PF} is unique up to scaling;*
4. *there exist $C, C' > 0$ such that for all $k \in \mathbb{N}$ and $i, j \in [n]$ we have*
$$C' \lambda_{\mathrm{PF}}^k \le (A^k)_{ij} \le C \lambda_{\mathrm{PF}}^k;$$
5. *λ_{PF} has algebraic multiplicity 1; and*
6. *any other eigenvalue, λ, of A aside from λ_{PF} has $|\lambda| < \lambda_{\mathrm{PF}}$.*

Proof. To prove (1), say that $(A\mathbf{v})_i > \lambda_{\mathrm{PF}} v_i$ for some $i \in [n]$, i.e., $A\mathbf{v} \ge \lambda_{\mathrm{PF}}\mathbf{v} + \epsilon \mathbf{e}_i$ for some $\epsilon > 0$. Then (44) implies

$$A(A\mathbf{v}) \ge \lambda_{\mathrm{PF}}(A\mathbf{v}) + \epsilon A\mathbf{e}_i;$$

since $A\mathbf{e}_i$ has all components strictly positive, for $\mathbf{w} = A\mathbf{v}$ we have $A\mathbf{w}$ is larger than $\lambda_{\mathrm{PF}}\mathbf{w}$ at every component, and hence $A\mathbf{w} \ge (\lambda_{\mathrm{PF}} + \delta)\mathbf{w}$ for some $\delta > 0$, which contradicts the definition of λ_{PF}.

To prove (2), if $A\mathbf{v} = \lambda_{\mathrm{PF}}\mathbf{v}$, and scale $\mathbf{v}$ so that one of its components is a positive real number. Then for all $i \in [n]$ we have

$$\lambda_{\mathrm{PF}} v_i = (A\mathbf{v})_i = \sum_{j=1}^n a_{ij} v_j.$$

Setting $\mathbf{w}$ to be the vector with $w_i = |v_i|$, we have that for each $i \in [n]$

$$\lambda_{\mathrm{PF}} w_i = \sum_{j=1}^n a_{ij} w_j \ge \left| \sum_{j=1}^n a_{ij} v_j \right| = \lambda_{\mathrm{PF}} |v_i| = \lambda_{\mathrm{PF}} w_i,$$

and therefore equality holds the above inequality, i.e.,

$$\left| \sum_{j=1}^n a_{ij} v_j \right| = \sum_{j=1}^n a_{ij} w_j = \sum_{j=1}^n a_{ij} |v_j|.$$

But we have strict inequality

$$\left| \sum_{j=1}^n a_{ij} v_j \right| < \sum_{j=1}^n a_{ij} |v_j|$$

unless for all j, $a_{ij} v_j$ have the same complex argument; since one of v_j is a positive real, then they all must be non-negative reals. Furthermore since at least one v_j is nonzero, the equation

$$\lambda_{\mathrm{PF}} v_i = \sum_{j=1}^n a_{ij} v_j$$

implies that all v_i are positive reals.

To prove (3), let $\mathbf{v}, \mathbf{w}$ be two eigenvectors for λ_{PF}; after scaling both $\mathbf{v}, \mathbf{w}$ have all positive entries, and after a further scaling we have $\mathbf{v} \leq \mathbf{w}$ and the inequality holds with equality at at least one of their components, i.e., $v_i = w_i$ for some $i \in [n]$. But then $\mathbf{v} = \mathbf{w}$, for otherwise since A has all positive entries, $A(\mathbf{w} - \mathbf{v})$ has all positive entries, which is impossible since then $(A\mathbf{w})_i > (A\mathbf{u})_i$, which is impossible since

$$(A\mathbf{w})_i = \lambda_{\mathrm{PF}} w_i = \lambda_{\mathrm{PF}} v_i = (A\mathbf{v})_i.$$

To prove (4), let $\mathbf{v}$ be the Perron-Frobenius eigenvector scaled so that $\mathbf{v} \geq \mathbf{1}$; then

$$A^k \mathbf{1} \leq \lambda_{\mathrm{PF}}^k \mathbf{v},$$

and hence for any $i, j \in [n]$ we have

$$(A^k)_{ij} \leq (A^k \mathbf{1})_j \leq \lambda_{\mathrm{PF}}^k v_j$$

which establishes the upper bound on $(A^k)_{ij}$ for $C = \max_j v_j$. To obtain a lower bound we write

$$(A^k)_{ij} = \mathbf{e}_i^{\mathrm{T}} A^{k-1} (A\mathbf{e}_j),$$

and since $A\mathbf{e}_j$ has all positive components, we can scale $\mathbf{v}$ to get an eigenvector $\mathbf{v}'$ such that $\mathbf{v}' \leq A\mathbf{e}_j$ for all j. Hence

$$(A^k)_{ij} \geq \mathbf{e}_i^{\mathrm{T}} A^{k-1} \mathbf{v}' \geq \mathbf{e}_i^{\mathrm{T}} \lambda_{\mathrm{PF}}^{k-1} \mathbf{v}' \geq \lambda_{\mathrm{PF}}^{k-1} v_i'$$

which establishes the lower bound for $C' = \min_i v_i' / \lambda_{\mathrm{PF}}$.

To prove (5), we have that for any invertible $M \in \mathcal{M}_n(\mathbb{C})$, MA^kM^{-1} is bounded by $C''\lambda_{\mathrm{PF}}^k$ in view of the upper bound on the entries of A^k. But then the single Jordan block corresponding to λ_{PF} has to be a 1×1 block, since an $m \times m$ Jordan block $J = J_m(\lambda_{\mathrm{PF}}$ has J^k having its immediate off diagonal entries equal to $k\lambda_{\mathrm{PF}}^{k-1}$, which exceeds $C''\lambda_{\mathrm{PF}}^k$ for k sufficiently large. Hence λ_{PF} has algebraic multiplicity one.

To prove (6), let $A\mathbf{v} = \lambda\mathbf{v}$ be any eigenpair of A, and normalize $\mathbf{v}$ so that some component of $\mathbf{v}$ is a positive real. Again, set $\mathbf{w}$ to be the vector with $w_i = |v_i|$. The same argument used to prove (2) shows that $A\mathbf{w} \geq |\lambda|\mathbf{w}$, and strict inequality holds at each component unless all non-zero components of $\mathbf{v}$ are positive. If strict equality held at each component, then $A\mathbf{w} \geq (|\lambda| + \epsilon)\mathbf{w}$ for sufficiently small $\epsilon > 0$, and hence $|\lambda| < \lambda_{\mathrm{PF}}$. Otherwise $\mathbf{v}$ has all its components being non-negative reals; but then λ is necessarily a positive real. So $A\mathbf{w} \geq |\lambda|\mathbf{w} = \lambda\mathbf{w}$ implies that $\lambda \leq \lambda_{\mathrm{PF}}$ and hence either $|\lambda| \leq \lambda_{\mathrm{PF}}$ unless $\lambda = \lambda_{\mathrm{PF}}$. □

Corollary 6.9. *Let $A \in \mathcal{M}_n(\mathbb{R})$ have non-negative entries and be irreducible and aperiodic. Then all the conclusions of Theorem 6.8 hold for A.*

Proof. Let $m \in \mathbb{N}$ be such that A^m has all positive entries; then A^{m+1} does as well since its entries are the dot products of the rows of A with the columns of A^m, and each row of A has at least one positive component. It follows that the conclusions of Theorem 6.8 hold for both A^m and A^{m+1}. But the eigenvalues of these matrices are those of A raised to the powers m and $m+1$. Hence A^m has a unique largest eigenvalue in absolute value, $\lambda_{\mathrm{PF}}(A^m)$, which equals λ_0^m for some eigenvalue of A which is therefore the unique eigenvalue of A of largest absolute value. The same reasoning shows that $\lambda_{\mathrm{PF}}(A^{m+1}) = \lambda_0^{m+1}$ for this same largest eigenvalue in absolute value of A, λ_0. But λ_0^m and λ_0^{m+1} are both positive reals, and hence so is λ_0. If $\mathbf{w}$ is eigenvector of λ_0 for A, then it also one of A^m of eigenvalue

λ_0^m. By Theorem 6.8, after scaling $\mathbf{w}$ has all positive entries. Since $A\mathbf{w} = \lambda_0\mathbf{w}$, we have $\lambda_0 \le \lambda_{\mathrm{PF}}$.

It remains to show that $\lambda_0 = \lambda_{\mathrm{PF}}(A)$ and any pseudo-eigenvector for A is propotional to $\mathbf{w}$. But if $A\mathbf{v} \ge \lambda_{\mathrm{PF}}(A)\mathbf{v}$ with $\mathbf{v} \ge 0$, then $A^m\mathbf{v} \ge (\lambda_{\mathrm{PF}}(A))^m\mathbf{v}$. It then follows that $(\lambda_{\mathrm{PF}}(A))^m \le \lambda_{\mathrm{PF}}(A^m) = \lambda_0^m and hence \lambda_{\mathrm{PF}}(A) = \lambda_0$. It then follows that $\mathbf{v}$ is a pseudo-eigenvector for A^m, and hence $\mathbf{v}$ is proportional to $\mathbf{v}$. □

6.6. **The Full Perron-Frobenius Theorem.**

Theorem 6.10. *Let $A \in \mathcal{M}_n(\mathbb{R})$ have non-negative entries. If A is irreducible, then*

1. *any Perron-Frobenius pseudo-eigenvector of A, $\mathbf{v}$, is an actual eigenvector of A (with eigenvalue λ_{PF});*
2. *if $\mathbf{v}$ is any eigenvector of A with eigenvalue λ_{PF}, then if some component of $\mathbf{v}$ is a positive real, then every component of $\mathbf{v}$ is a positive real;*
3. *if A has period p and λ is any eigenvector of A, then $|\lambda| \le \lambda_{\mathrm{PF}}$ and equality holds iff $\lambda = \zeta\lambda_{\mathrm{PF}}$ for some ζ with $\zeta^p = 1$;*
4. *the eigenvectors $\mathbf{v}$ with eigenvalue λ_{PF} are unique up to scaling;*
5. *there is a $C > 0$ such that for any $k \in \mathbb{N}$, all entries of A^k are at least 0 and at most $C\lambda_{\mathrm{PF}}^k$; and*
6. *λ_{PF} has (algebraic) multiplicity one as an eigenvalue of A.*

Finally, if A has period p, then for each ζ with $\zeta^p = 1$, $\zeta\lambda_{\mathrm{PF}}$ is an eigenvalue with multiplicity one, and these are the only eigenvalues of absolute value (at least) λ_{PF}.

One way to prove this theorem is to infer that Theorem 6.8 also holds if A is aperiodic, and then if A has period p, then we note that A^p is aperiodic and irreducible as it acts on any of the p parts into which $[n]$ is partitioned by the periodicity of A. Below we give a direct proof based on repeating the same type of argument.

Proof. The proofs of (1), (2), (3), and (4) we give involve very similar arguments: we show that if any of these fail to hold, then for some $I \subset [n]$ there is no $i \in I$ and $j \notin I$ with $a_{ji} > 0$, and hence A is not irreducible. It can be helpful to keep Example 6.6 in mind.

To prove (1), we will show that if $A\mathbf{v}$ has one entry strictly greater $\lambda_{\mathrm{PF}}\mathbf{v}$, then A is not irreducible, as in Example 6.6. Indeed, in this case consider an $I \subset [n]$ such that I is as large as possible and for some pseudo-eigenvector we have $(A\mathbf{v})_i > \lambda_{\mathrm{PF}}v_i$ for all $i \in I$. We cannot have $I = [n]$, for then $(A\mathbf{v})_i > \lambda_{\mathrm{PF}}v_i$ for all $i \in [n]$ and hence $(A\mathbf{v})_i > (\lambda_{\mathrm{PF}} + \epsilon)v_i$ for $\epsilon > 0$ sufficiently small. We claim that there is no positive a_{ji} with $i \in I$ and $j \notin I$; otherwise for small ϵ we have $(A\mathbf{v})_i > (\lambda_{\mathrm{PF}}(v_i + \epsilon)$ for some sufficiently small $\epsilon > 0$, and hence if $\mathbf{w}$ is given by $w_i = v_i + \epsilon$ for $i \in I$ and $w_{i'} = v_{i'}$ for $i' \notin I$ we have

$$(A\mathbf{w})_{i'} \ge \lambda_{\mathrm{PF}}v_{i'} = \lambda_{\mathrm{PF}}\mathbf{w}_{i'}.$$

Furthermore, we have

$$(A\mathbf{w})_j \ge (A\mathbf{v})_j + a_{ji}\epsilon = \lambda_{\mathrm{PF}}w_j + a_{ij}\epsilon,$$

and so $(A\mathbf{w})_i > \lambda_{\mathrm{PF}}w_i$ if $i \in I$ or $i = j$, which contradicts the maximality of I.

Hence any pseudo-eigenvector is a genuine eigenvector. Now we prove (2), (3) and (4) with similar arguments.

To prove (2), let $\mathbf{w}$ be the vector given by $w_i = |v_i|$. Then since

$$(A\mathbf{v})_j = \sum_{i=1}^{n} a_{ji} v_i,$$

for any j, we have

$$\lambda_{\mathrm{PF}} w_j = \lambda_{\mathrm{PF}} |v_j| = \left| \sum_{i=1}^{n} a_{ji} v_i \right| \leq \sum_{i=1}^{n} a_{ji} |v_i| = (A\mathbf{w})_j.$$

Hence $\mathbf{w}$ is a pseudo-eigenvector, and hence an eigenvector; and hence for each j,

$$\lambda_{\mathrm{PF}} w_j = \left| \sum_{i=1}^{n} a_{ji} v_i \right| \leq \sum_{i=1}^{n} a_{ji} |v_i|$$

holds with equality. It follows that if $I \subset [n]$ is the set of $i \in [n]$ such that v_i is nonzero, then $I = [n]$, for otherwise we have $a_{ji} = 0$ whenever $j \notin I$ and $i \in I$, since

$$\lambda_{\mathrm{PF}} |v_j| = \sum_{i=1}^{n} a_{ji} |v_i|.$$

To finish the proof of (2), we want to show that if $I \subset [n]$ is the set of $i \in [n]$ such that v_i is a positive real, then $I = [n]$; we argue similarly: the equality

$$\left| \sum_{i=1}^{n} a_{ji} v_i \right| = \sum_{i=1}^{n} a_{ji} |v_i|$$

for all j shows that for all j, whenever $a_{ji}, a_{ji'}$ are nonzero, then $v_i, v_{i'} \in \mathbb{C}$ have the same argument (i.e., point in the same direction). The fact that

$$\lambda_{\mathrm{PF}} v_j = \sum_{i=1}^{n} a_{ji} v_i$$

then implies that $a_{ji} = 0$ when $i \in I$ and $j \notin I$, for otherwise right-hand-side above would involve a positive real $a_{ji} v_i$, and hence would consist entirely of positive reals, and hence v_j would be a positive real. It follows that $I = [n]$.

To prove (3), we let $A\mathbf{v} = \lambda v$ and similarly define $\mathbf{w}$ via $w_i = |v_i|$. We similarly see that if $|\lambda| \geq \lambda_{\mathrm{PF}}$, then $|\lambda| = \lambda_{\mathrm{PF}}$ and for any j, for all i with $a_{ji} > 0$ all the corresponding v_i have the same argument. So scaling $\mathbf{v}$ to have v_1 real, and letting $V_0, \ldots, V_{p-1}$ be the partition of $[n]$ as in Proposition 6.3, we see that the argument of λ is precisely that of $v_j / v_{j'}$ for any $j' \in V_\ell$ and a $j \in V_{\ell+1}$. Since there are paths from any vertex to itself of any sufficiently large length divisible by p, we have the argument of λ is a p-th root of unity. The converse holds by Proposition 6.3.

To prove (4), first we show that the space of eigenvectors with eigenvalue λ_{PF} is one dimensional: indeed, if $\mathbf{v}, \mathbf{v}'$ are any such eigenvectors, then after scaling we may assume they have all positive components. Then after scaling we may assume that $\mathbf{v} \leq \mathbf{v}'$, and that $\mathbf{v}, \mathbf{v}'$ are equal at one of their components. So let I be the subset of i such that $v_i = v_i'$. We similarly show that if $i \in I$ and $j \notin J$, then $a_{ji} = 0$. Hence, since I is nonempty, either $I = [n]$ or A is reducible.

To prove (5), we need to find a C such that

$$0 \leq (A^k)_{ij} \leq C \lambda_{\mathrm{PF}}^k. \tag{45}$$

The bound $0 \leq (A^k)_{ij}$ is immediate from the fact that A has non-negative entries. To prove the other bound, let $\mathbf{v}$ be a Perron-Frobenius eigenvector, scaled so that

$\mathbf{v} \geq \mathbf{1}$; then the non-negativity of A implies that for any $k \in \mathbb{N}$ we have $A^k(\mathbf{v}-\mathbf{1}) \geq 0$, and hence

$$A^k\mathbf{1} \leq A^k\mathbf{v} = \lambda_{\mathrm{PF}}^k\mathbf{v},$$

and hence for any $i \in [n]$ we have

$$\lambda_{\mathrm{PF}}^k v_i \geq (A^k\mathbf{1})_i = \sum_{j=1}^n (A^k)_{ij}.$$

This establishes (45) for $C = \max_i v_i$.

To prove (6), recall that the Jordan canonical form writes A as $M^{-1}JM$ where $M, J \in \mathcal{M}_n(\mathbb{R})$, M is invertible, and J is a block diagonal matrix of Jordan blocks; (4) implies that the eigenspace of λ_{PF} is one-dimensional, hence J has a single Jordan block associated to λ_{PF}; to show that λ_{PF} has algebraic multiplicity one, we need to show that this Jordan block is of size one. But (5) implies that there is a C' such that we have that the entries of $(MAM^{-1})^k = MA^kM^{-1}$ are bounded by $C'\lambda_{\mathrm{PF}}^k$. It follows that A cannot have a Jordan blocks, J, of size 2×2 or more corresponding to the eigenvalue λ_{PF}, since these blocks have J^k equal to $k\lambda_{\mathrm{PF}}^{k-1}$ in their immediate off-diagonal entries.

If A is p-periodic, then one can similarly prove that for $\zeta^p = 1$, the eigenvalues $\zeta\lambda_{\mathrm{PF}}$ have algebraic multiplicity one. Indeed, Proposition 6.3 implies that these eigenvalues have geometric multiplicity one. Hence corresponding to $\zeta\lambda_{\mathrm{PF}}$ there is a single Jordan block of A, and if this block were of size 2×2 or greater, then its k power would have entries of size $k\lambda_{\mathrm{PF}}^{k-1}$ in absolute value which is again impossible in view of (45).

Finally, if $A\mathbf{v} = \lambda\mathbf{v}$ is any eigenvector, then taking $\mathbf{w}$ to be the vector with $w_i = |v_i|$ for all i, the triangle inequality as used in part (2) shows that $A\mathbf{w} \geq |\lambda|\mathbf{w}$. Hence $|\lambda| \leq \lambda_{\mathrm{PF}}$, and hence either $|\lambda| < \lambda_{\mathrm{PF}}$ or $\lambda = \nu\lambda_{\mathrm{PF}}$ with $|\nu| = 1$; it remains to show that $\nu^p = 1$. Since $\mathbf{v} \neq \mathbf{0}$, we may scale $\mathbf{v}$ so that for some i, v_i is a positive real. Now let $V_0, \dots, V_{p-1}$ be the partition of $[n]$ given by the periodicity of A. We have $i \in V_j$ for some j, and $A\mathbf{v} = \lambda\mathbf{v}$ implies that all the i' with $a_{i'i} > 0$ $v_{i'} \neq 0$ have $v_{i'}$ having the same complex argument, which is exactly that of λ. Since A is irreducible and p-periodic, for some k there is a path from i to itself of length kp, and another of length $(k+1)p$. It follows that λ^{kp} and $\lambda^{(k+1)p}$ have the same argument, which are those of ν^{kp} and $\nu^{(k+1)p}$, and hence ν^p is a positive real. Hence $\nu^p = 1$. □

6.7. **Application to Information Theory and Run-Length Constrained Data.** On April 6 and 8 we discussed (d, k) run-length constrained data, meaning strings on $\{0, 1\}$ such that between any two successive 1's there are at least d 0's and at most k 0's. For certain values of d and k such strings are more reliable stored and read from magnetic storage devices. Coding from general $\{0, 1\}$ data into run-length constrained data is a large topic, and has (at the time) surprising connections to the field of Symbolic Dynamics.

In class we modeled $(0, 1)$ run-length data by the Fibonacci graph, and similarly modeled (d, k) run-length data by a graph with vertices $v_1, \dots, v_{k+1}$, with an edge labeled 0 from v_i to v_{i+1} for $i = 1, \dots, k$, and an edge labeled 1 from v_i to v_1 for each $i = d + 1, \dots, k + 1$.

If G is any digraph, we define the *capacity* of G to be $\log_2 \lambda_{\mathrm{PF}}$, where λ_{PF} is the Perron-Frobenius eigenvalue of A_G, the adjacency matrix of G; we view this

capacity as measured in *bits* (binary digits), given that the logarithm is taken in base 2. For example, if G consists of a single vertex with d self-loops, then its capacity is $\log_2 d$. If $G[m]$ is the graph whose vertex set is that of G, and whose edges are walks of length m in G, then $A_{G[m]} = A_G^m$, and hence the capacity of $G[m]$ is m times that of G.

EXERCISES.

Exercise 6.1. Let $A \in \mathcal{M}_n(\mathbb{R})$ have non-negative entries and be irreducible. The point of this exercise is to give a short proof that any Perron-Frobenius pseudo-eigenvector of A is an actual eigenvector.

6.1(a) Show that each entry of $B = I + A + A^2 + \ldots + A^{n-1}$ is positive.

6.1(b) Show that if $\mathbf{v}$ is a pseudo-eigenvector of A, then if $A\mathbf{v} \geq \lambda\mathbf{v}$ where strict inequality holds at some component, then setting $\mathbf{w} = B\mathbf{v}$, we have $A\mathbf{w} \geq \lambda\mathbf{w}$ where strict inequality holds at every component.

6.1(c) Explain why this implies that every pseudo-eigenvector is an eigenvector.

Exercise 6.2. Let $A \in \mathcal{M}_n(\mathbb{R})$ have non-negative entries and be irreducible. Let $B \in \mathcal{M}_n(\mathbb{R})$ satisfy $b_{ij} \geq a_{ij}$ for all $i, j \in [n]$, and assume that $b_{ij} > a_{ij}$ for at least one value of i and one value of j. Show that $\lambda_{\mathrm{PF}}(B)$ is strictly larger than $\lambda_{\mathrm{PF}}(A)$.

Exercise 6.3. Let G, G' be graphs be strongly connected graphs such that G is a proper subgraph of G' (i.e., $G \neq G'$). Show that the capacity of G' is strictly larger than that of G. (You may use Exercise 6.2.)

Exercise 6.4. Consider the graph, G, associated to (d, k) run-length constrained data with vertex set $v_1, \ldots, v_{k+1}$ as described in class and the previous subsection. The point of this exercise is to show that $\lambda_{\mathrm{PF}}(G)$ is $1/z_0$ where $z_0 > 0$ is the smallest positive root of the equation

$$1 = z^{d+1} + \cdots + z^{k+1}.$$

6.4(a) For each $m = 0, 1, \ldots$, let a_m denote the number of walks of length m that begin and end at v_1. Show that there is an equality of formal power series

$$\sum_{m=0}^{\infty} a_m z^m = 1 + p + p^2 + \cdots, \quad \text{where} \quad p = p(z) = z^{d+1} + \cdots + z^{k+1}.$$

6.4(b) Conclude that if

$$f(z) = \sum_{m=0}^{\infty} a_m z^m,$$

then $f(z)$ converges for $z < z_0$ above, and that $f(z) \to \infty$ as z tends to z_0 from below.

6.4(c) Conclude that

$$\limsup_{m \to \infty} (a_m)^{1/m} = 1/z_0.$$

6.4(d) Use (45) to conclude that $\lambda_{\mathrm{PF}}(G) = 1/z_0$.

6.4(e) Let G be any digraph such that there is a vertex, v_1, of G such that any walk of sufficiently long length passes through v_1. Describe an analog of the the above computation of λ_{PF} to this more general situation.

Exercise 6.5. Let G be the graph associated to $(1,3)$ run-length constrained data.

6.5(a) Use Exercise 6.4 to show that the capacity of this graph is larger than $1/2$.

6.5(b) Show that one can give an encoding/decoding scheme to convert a general $\{0,1\}$ strings to a string that is roughly twice as long that satisfies the $(1,3)$ run-length constraint. (Base this scheme on the one given on April 6 to convert a general string to a string of roughly $3/2$ as long that satisfies the $(0,1)$ run-length constraint.

Appendix A. Reference for Various Useful Facts

Here we gather some important tools that we will constantly use in some proofs of our theorems and to gain intuition regarding matrices.

In CPSC 531F this year, we will only use this section as a reference, to be consulted when needed.

[See the Table of Contents for the contents of this section.]

A.1. **Norms on $\mathbb{R}^n$.** For any $\mathbf{u} \in \mathbb{R}^n$, we set

$$\|\mathbf{u}\|_1 = |u_1| + \cdots + |u_n|$$

(the L^1-*norm* of $\mathbf{u}$). Note that if $\mathbf{u} \neq \mathbf{0}$, then $\|\mathbf{u}\|_1 > 0$ and $\mathbf{u}' = \mathbf{u}/\|\mathbf{u}\|$ is a (the only) positive scalar multiple of $\mathbf{u}$ such that is an L^1-*unit vector*, i.e., such that $\|\mathbf{u}'\|_1 =$.

The reader has likely seen the idea of unit vectors with respect to the usual norm

$$\|\mathbf{u}\|_2 = \left(|u_1|^2 + \cdots + |u_n|^2\right)^{1/2}.$$

Other useful norms include

$$\|\mathbf{u}\|_p = \left(|u_1|^p + \cdots + |u_n|^p\right)^{1/p}$$

for $p \geq 1$, and the $p \to \infty$ limit of $\|\mathbf{u}\|_p$, which is just

$$\|\mathbf{u}\|_\infty = \max_i |u_i|.$$

All of these functions are *norms*, meaning functions $\mathbb{R}^n \to \mathbb{R}$, written $\mathbf{u} \mapsto \|\mathbf{u}\|$, such that for any $\mathbf{u}, \mathbf{w} \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$:

(1) $\|\mathbf{u}\| \geq 0$, with equality iff $\mathbf{u} = \mathbf{0}$;
(2) $\|\alpha\mathbf{u}\| = |\alpha|\,\|\mathbf{u}\|$;
(3) $\|\mathbf{u} + \mathbf{w}\| \leq \|\mathbf{u}\| + \|\mathbf{w}\|$.

A.2. **Dot Products and Inner Products on $\mathbb{R}^n$.** For $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, as usual, we set

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + \cdots + u_n v_n = \mathbf{u}^{\mathrm{T}}\mathbf{v},$$

where the last term, $\mathbf{u}^{\mathrm{T}}\mathbf{v}$, is really a 1×1 matrix, but we regard it as a scalar (i.e., a real). If $A \in \mathcal{M}_n(\mathbb{R})$, then the equality

$$(A\mathbf{u})^{\mathrm{T}}\mathbf{v} = \mathbf{u}^{\mathrm{T}}A^{\mathrm{T}}\mathbf{v} = \mathbf{u}^{\mathrm{T}}(A^{\mathrm{T}}\mathbf{v})$$

is a convenient way to show $(A\mathbf{u}) \cdot \mathbf{v} = \mathbf{u} \cdot (A^{\mathrm{T}})$.

More generally, an *inner product* on $\mathbb{R}^n$ is any map $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, denoted $\langle \mathbf{u}, \mathbf{v} \rangle$ in [HJ85, HJ13](sometimes $(\mathbf{u}, \mathbf{v})$ in the literature, that is (1) bilinear, (2) symmetric, and (3) positive definite (see an equivalent set of terms and definitions in Section 5.1 of [HJ85, HJ13]). The dot product is the special case $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u} \cdot \mathbf{v}$.

[Notice that in physics, i.e., relativity, one may want to work with an inner product that is (non-degenerate but) not positive definite, such as the famous inner product on tangent vectors expressed by $dx^2 + dy^2 + dz^2 - (c\,dt)^2$.]

If $A \in \mathcal{M}_n(\mathbb{R})$, and $\langle\ ,\ \rangle$ is any inner product then there is a unique $A^* \in \mathcal{M}_n(\mathbb{R})$, called the *adjoint of A (with respect to $\langle\ ,\ \rangle$)* such that

$$\langle A\mathbf{x}, \mathbf{y}\rangle = \langle \mathbf{x}, A^*\mathbf{y}\rangle$$

for all $\mathbf{x}, \mathbf{y}$.

Example A.1. If $P \in \mathcal{M}_n(\mathbb{R})$ is an irreducible Markov matrix with stationary distribution $\boldsymbol{\pi}$, then P is *reversible* if for all i, j, $\pi_i p_{ij} = \pi_j p_{ji}$, which is equivalent to saying that the *invariant measure* that P induces on $[n]^{\mathbb{Z}}$ is the same as the reverse measure, i.e., the "time reversal" of this Markov chain is the same as the original chain (see elsewhere in these notes and/or [LP17] for details). In this case we have

$$\langle P\mathbf{x}, \mathbf{y}\rangle_{\boldsymbol{\pi}} = \langle \mathbf{x}, P\mathbf{y}\rangle_{\boldsymbol{\pi}},$$

where

$$\langle \mathbf{x}, \mathbf{y}\rangle_{\boldsymbol{\pi}} = \sum_{i=1}^{n} x_i y_i \pi_i$$

(which we easily see an inner product) and similarly

$$\langle \mathbf{x}^{\mathrm{T}} P, \mathbf{y}^{\mathrm{T}}\rangle_{1/\boldsymbol{\pi}} = \langle \mathbf{x}^{\mathrm{T}}, \mathbf{y}^{\mathrm{T}} P\rangle_{1/\boldsymbol{\pi}},$$

where (we view $\mathbb{R}^n$ as row vectors, which is what P really operates on...), and

$$\langle \mathbf{x}^{\mathrm{T}}, \mathbf{y}^{\mathrm{T}}\rangle_{1/\boldsymbol{\pi}} = \sum_{i=1}^{n} x_i y_i (1/\pi_i).$$

It follows from the theory of symmetric operators—which is a straightforward generalization of the theory of symmetric matrices—that P has real eigenvalues with a (real) orthonormal eigenbasis of eigenvectors with respect to $\langle\ ,\ \rangle_{\boldsymbol{\pi}}$ and of left eigenvectors with respect to $\langle\ ,\ \rangle_{1/\boldsymbol{\pi}}$.

A.3. **Complex Dot Products and Inner Products.** We follow the conventions of [HJ85, HJ13], e.g., Section 0.6. Here the standard dot product on $\mathbb{C}^n$ is written $\langle \mathbf{x}, \mathbf{y}\rangle = \mathbf{y}^*\mathbf{x}$, where * is the conjugate transpose; it is more common to see $^{\mathrm{H}}$ written for the conjuage transpose elsewhere in the literature.

We say that A is *Hermitian* if $A^H = A$; this is the usual way to extend the notion of a symmetric matrix to $\mathcal{M}_n(\mathbb{C})$; indeed, it turns out that any Hermitian matrix has real eigenvalues and a (generally complex) orthonormal eigenbasis with respect to the complex dot product.

Similarly the usual way of generalizing the notion of an *orthogonal matrix* to $\mathcal{M}_n(\mathbb{C})$ is a *unitary matrix*, meaning a matrix $U \in \mathcal{M}_n(\mathbb{C})$ for which $U^* = U$ or $U^{\mathrm{H}} = U$, or, equivalently, a matrix whose rows (and therefore all its columns) form an orthonormal complex eigenbasis of $\mathbb{C}^n$.

A.4. **Matrix Norms.** (See [HJ85, HJ13], Section 5.6.)

By a *matrix norm* one means a norm on $\mathcal{M}_{m,n}(\mathbb{R})$ that—at the very least—is a norm when we identify $\mathcal{M}_{m,n}(\mathbb{R})$ with $\mathbb{R}^{mn}$ (or the same with $\mathbb{C}$). When we are working on a context where it makes sense to multiply matrices, we want $\|AB\| \leq \|A\|\,\|B\|$.

For example, when n is fixed and $\mathcal{M}_n(\mathbb{R})$ is viewed as operating on $\mathbb{R}^n$, then likely want this; in this way, for example, $\|A^k\| \le \|A\|^k$.

Example A.2. Let V_1, V_2 be vector spaces with respective norms $\|\ \|_1, \|\ \|_2$. Then if $L\colon V_1 \to V_2$, then

$$\|L\| = \max_{\mathbf{v}_1 \neq \mathbf{0}} \frac{\|L\mathbf{v}_1\|_2}{\|\mathbf{v}_1\|_1} = \max_{\|\mathbf{v}_1\|_1 = 1} \|L\mathbf{v}_1\|_2$$

measures the maximum amount that L "stretches the first norm with respect to the second;" this is often called the *induced operator norm*. We easily see that if $L'\colon V_2 \to V_3$ is a map to a vector space, V_3, with norm $\|\ \|_3$, then $\|L'L\| \le \|L'\|\,\|L\|$.

Example A.3. Concretely, if $p \in [1,\infty]$, and $A \in \mathcal{M}_n(\mathbb{R})$, then we write $\|A\|_p$ for the above operator norm induced from $\|\ \|_p$ on $\mathbb{R}^n$. It turns out that $\|A\|_\infty$ and $\|A\|_1$ are quite simple to compute directly, as they are the maximum row sum of absolute values and, respectively, column sum.

A.5. **The Resultant and Discriminant.** Our approach to the next subsection will use the notion of the discriminant which we now review.

Recall that two polynomials $r(t), s(t)$ with real or complex coefficients[7] have a common factor $t-\alpha$ with $\alpha \in \mathbb{C}$,[8] iff α is a common root of $r(t), s(t)$, iff $r(t)\tilde{s}(t) + s(t)\tilde{r}(t) = 0$ where $\deg(\tilde{s}) < \deg(s)$ and $\deg(\tilde{r}) < \deg(r)$; moreover, one can use Euclid's algorithm to find the greatest common factor of r and s. The condition $r(t)\tilde{s}(t) + s(t)\tilde{r}(t) = 0$ is equivalent to the equation

(46)

$$\begin{bmatrix} \tilde{s}_0 & \cdots & \tilde{s}_{k-1} & \tilde{r}_0 & \cdots & \tilde{r}_{k'-1} \end{bmatrix} \begin{bmatrix} r_0 & r_1 & \cdots & r_k & & & \\ & \ddots & \ddots & \ddots & \ddots & & \\ & & r_0 & r_1 & \cdots & r_k & \\ s_0 & s_1 & \cdots & \cdots & s_{k'} & & \\ & \ddots & \ddots & \ddots & \ddots & \ddots & \\ & & s_0 & s_1 & \cdots & \cdots & s_{k'} \end{bmatrix} = \mathbf{0}^{\mathrm{T}}$$

in the coefficients of $r, s, \tilde{r}, \tilde{s}$ (i.e., $r(t) = r_0 + \cdots + r_k t^k$, etc.) [In (46) we depict the case $k < k'$, where $\deg(r) = k$, $\deg(s) = k'$; however, we may well have $k = k'$ or $k > k'$.] The determinant of the square matrix in (46) is called the *resultant* of r, s.

If $p(t)$ is any polynomial, then $p(t)$ has a repeated root (i.e., at least one root with multiplicity at least two) iff $p(t)$ and its derivative $p'(t)$ have a common factor. Hence, setting $r(t) = p(t)$ and $s(t) = p'(t)$ (46), $p(t)$ has a repeated root iff

(47)

$$\mathrm{Det} \begin{bmatrix} p_1 & 2p_2 & \cdots & (k-1)p_k & & & \\ & \ddots & \ddots & \ddots & \ddots & & \\ & & p_1 & 2p_2 & \cdots & (k-1)p_k & \\ p_0 & p_1 & \cdots & p_{k-1} & p_k & & \\ & \ddots & \ddots & \ddots & \ddots & \ddots & \\ & & p_0 & p_1 & \cdots & p_{k-1} & p_k \end{bmatrix}$$

[7] All this discussion works with coefficients in any field, provided that $\mathbb{C}$ is replaced by the algebraic closure of the field.

[8] which for polynomials over $\mathbb{R}$ implies that either $\alpha \in \mathbb{R}$ or they have a common real polynomial factor $(t-\alpha)(t-\overline{\alpha})$

vanishes. For example, if $p(t) = at^2 + bt + c$, then

$$\det \begin{bmatrix} b & 2a & 0 \\ 0 & b & 2a \\ c & b & a \end{bmatrix} = a(b^2 - 4ac),$$

which expresses the fact that if p is truly of degree 2, i.e., $a \neq 0$, then p has repeated roots iff $b^2 - 4ac$ (the *discriminant* of p) vanishes.

A.6. **The Approximation of Matrices by Diagonalizable Matrices.** In this section we prove the following theorem very useful theorem.

Theorem A.4. *Let $A \in \mathcal{M}_n(\mathbb{R})$. Then there exists $A_1, A_2, \ldots$ with distinct eigenvalues such that $A_m \to A$ as $m \to \infty$. Similarly if A is symmetric. Similarly with $\mathbb{C}$ replacing $\mathbb{R}$, and Hermitian replacing symmetric.*

Here $A_m \to A$ as $m \to \infty$ can be taken to mean that each entry of A_m tends to that of A; it is equivalent to say that $\|A_m - A\| \to 0$ in any matrix norm.

There are a number of ways to prove this theorem. We will give a proof based on the following fact.

Lemma A.5. *For any $n \in \mathbb{R}$, there is a (nonzero) polynomial q_n in the variables a_{ij} with i, j ranging over $[n]$, such that $A = (a_{ij})$ has a multiple eigenvalue iff $q_n(a_{ij}) = 0$.*

Proof. We have

$$p_A(t) = \det(tI - A) = t^n + r_1(a_{ij})t^{n-1} + \cdots + r_n(a_{ij})$$

where $r_k(a_{ij})$ is a polynomial of degree k in the variables a_{ij} (i.e., given by $(-1)^k$ times the sum of determinants of the $k \times k$ principal minors of A). Since $p_A(t)$ has leading coefficient 1, $p_A(t)$ is of degree n, and hence we take $q_n(a_{ij})$ to be as in (47) with $p_i(a_{ij}) = r_{n-i}(a_{ij})$. Since there exist matrices in $\mathcal{M}_n(\mathbb{R}, \mathbb{C})$ with distinct eigenvalues, q_n is not the zero polynomial. □

Proof of Theorem A.4. Let D diagonal matrix with distinct diagonal elements. Consider for $\epsilon \in \mathbb{R}$, $A_\epsilon = A + \epsilon(D - A)$. Then if q_n is as in Lemma A.5, we have $r(\epsilon) = q_n(A_\epsilon)$ is a nonzero polynomial, since $A_1 = D$ so $r(1) \neq 0$. Hence $r(\epsilon)$ has at most finitely many zeros, so for ϵ sufficiently close to 0 we have A_ϵ has distinct eigenvalues.

Moreover, if A is symmetric or Hermitian, then so is A_ϵ. □

A.7. **The Cayley-Hamilton Theorem.** This theorem says that if $A \in \mathcal{M}_n(\mathbb{R}, \mathbb{C})$, then $p_A(A) = 0$, i.e., the characteristic polynomial of A applied to A is the all 0's matrix (understood in the context of $n \times n$ matrices, where a constant represents this constant times the identity matrix)

There are two proofs that one typically sees first, in that they rely on important general principles; however, there is a third proof that uses the Schur decomposition, which is a nice way to view unitary matrices (e.g., from Quantum mechanics)—i.e., as a special case of *normal matrices*—which we will do in Subsection **??**.

Proof 1: For any eigenvector, $\mathbf{v}$, of A, $p_A(A)\mathbf{v} = \mathbf{0}$ since $p_A(t)$ has a factor of $t - \lambda$ for each eigenvalue λ. Hence $p_A(A) = 0$ for any diagonalizable matrix, A, since there is a basis of eigenvectors. In general there is a sequence of diagonalizable matrices A_m tending to A, and by continuity of the characteristic polynomial we have $p_{A_m}(A_m) = 0$ implies $p_A(A) = 0$.

Proof 2: It suffices to show that $p(A^{\mathrm{T}}) = 0$. Let I denote the $n \times n$ identity matrix, and let M be the block $n \times n$ matrix

$$M = \begin{bmatrix} a_{11}I - A^{\mathrm{T}} & a_{12}I & \cdots & a_{1n}I \\ a_{21}I & a_{22}I - A^{\mathrm{T}} & \cdots & a_{2n}I \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}I & a_{n2}I & \cdots & a_{nn}I - A^{\mathrm{T}} \end{bmatrix}.$$

We see that

$$M \begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \vdots \\ \mathbf{e}_n \end{bmatrix} = \mathbf{0}, \tag{48}$$

which is a bit more visible in an example, such as

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

For any $B \in \mathcal{M}_n(\mathbb{R}, \mathbb{C})$ we have

$$\mathrm{adj}(B)B = \begin{bmatrix} \det B & & \\ & \ddots & \\ & & \det B \end{bmatrix},$$

where $\mathrm{adj}(B)$ is the the *adjugate* (see [HJ85, HJ13], Section 0.8, also called the *classical adjoint*). Notice that this formula holds in any context where $B = [b_{ij}]$ and the b_{ij} lie in any setting with a commutative addition and multiplication (with certain standard properties, such as having additive inverses so it makes sense to speak of $-r$ for each element r, i.e., a *commutative ring*), we can use polynomials in A as the entries of B (where the constant 1 polynomial is taken to mean $A^0 = I$), and multiply (48) by the adjugate of M and conclude that

$$\begin{bmatrix} p_A(A^{\mathrm{T}}) & 0 & \cdots & 0 \\ 0 & p_A(A^{\mathrm{T}}) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & p_A(A^{\mathrm{T}}) \end{bmatrix} \begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \vdots \\ \mathbf{e}_n \end{bmatrix} = \mathbf{0}.$$

Hence $p_A(A^{\mathrm{T}})\mathbf{e}_i$ vanishes for all $i \in [n]$, and hence $p_A(A^{\mathrm{T}})$ is the zero matrix.

Proof 3: The proof in [HJ85, HJ13]uses the *Schur decomposition*, see Section 2.4 there, or Subsection **??** of these notes). This technique is extremely useful when we discuss *normal matrices*.

A.8. **The Supremum and Compactness.** When you prove Rolle's Theorem in calculus (used to prove the Mean-Value Theorem and Taylor's Theorem), you need to use the fact that any continuous function $f\colon [a,b] \to \mathbb{R}$ attains its maximum somewhere on $[a,b]$. It is easy to see that this is not true if $[a,b]$ is replaced by an open interval (a,b) or an unbounded interval such as $(-\infty, b)$. But how do you prove this? In this subsection we will discuss the main ingredients of a standard proof: supremum and compactness.

[This is why you take a course like UBC's Math 320, which is often called "real variables," which in the past was called "advanced calculus."] In this class we will see quite a few applications of these ideas; one can often circumvent the need for these ideas, but they will provide valuable insight, especially when we consider "perturbations" of matrices.

A.8.1. *The Supremum or Least Upper Bound (and Infimum or Greateset Lower Bound).*

Lemma A.6. *Let $S \subset \mathbb{R}$ be nonempty and have an upper bound $B \in \mathbb{R}$, i.e, we have $s \leq B$ for all $s \in S$. Then S has an upper bound M such that if $M' < M$, then M' is not an upper bound on M. Moreover, there is a sequence $s_1, s_2, \ldots$ of elements of S (which are not necessarily distinct) such that $s_i \to M$ as $i \to \infty$.*

The value M above is called the *least upper bound* or *supremum* of S. One can prove this using bisection, repeatedly taking an element $s \in S$ and the upper bound B, and considering whether or not $(s + B)/2$ is an upper bound for S (see Exercise **??**.

Behind the scenes of any proof of Lemma A.6 you essentially need to know that if $s_1, s_2, \ldots$ is a *Cauchy sequence* of real numbers, i.e., such that as $i, j \to \infty$, $|s_i - s_j| \to 0$, then $s_1, s_2, \ldots$ has a limit $s \in \mathbb{R}$; at the very least—when using a bisection argument—you need to know this in the case $|s_{i+1} - s_i| \leq C/2^i$ for a constant C. This is not true if $\mathbb{R}$ is replaced with the set of rational numbers.

If S is unbounded from above, i.e., does not have an upper bound, then one defines it supremum to be $+\infty$; one defines the supremum of the empty set to be $-\infty$ (if one really needs to). One similarly defines the *infimum* or *greatest lower bound* of a set S, which equals $-\sup(-S)$.

A.8.2. *Compactness.* The other idea we need to prove Rolle's theorem is *compactness*: any closed interval $[a, b] \subset \mathbb{R}$ is *compact* in the sense that any sequence, $s_1, s_2, \ldots$ in $[a, b]$ has a convergent subsequence. One can prove this using bisection: let s_1 be the first element of the subsequence; consider $m = (a + b)/2$ is the interval's midpoint: infinitely many of $s_1, s_2, \ldots$ must lie in at least one of $[a, m]$ and $[m, b]$; if, say, $[a, m]$ has infinitely many, then choose a new element of $s_1, s_2, \ldots$ in $[a, m]$ and consider the midpoint of this interval; etc.

From this it easily follows (EXERCISE) that the unit ball in any L^p-norm (including $p = \infty$) is compact. More generally, a subset of $\mathbb{R}^n$ is compact iff it is bounded and *closed*. We won't explain this in detail, but if follows that if S is any bounded set that is defined as the zero set of a number of continuous function, then S is compact; this is another way to see that the unit ball in the norms above—or any norm on $\mathbb{R}^n$ or $\mathbb{C}^n$ (defined similarly)—is compact.

From this it follows that for each m we have an orthonormal basis, $\mathbf{x}_1^m, \ldots, \mathbf{x}_n^m$ of $\mathbb{R}^n$ or $\mathbb{C}^n$, then by passing to a subsequence of $m = 1, 2, \ldots$ we may assume that $\mathbf{x}_i^m$ has limit, $\mathbf{x}_i$, for $i \in [n]$, and then by continuity of the dot product we have that $\mathbf{x}_1, \ldots, \mathbf{x}_m$ is orthonormal basis (EXERCISE).

[Now we can outline a proof of the above ingredient of Rolle's Theorem, or leave it for exercises.]

A.9. **The Brouwer Fixed-Point Theorem.** We are unlikely to need the Brouwer fixed-point theorem in this course, but sometimes it can simplify proofs (at the risk

of obscuring a more efficient algorithm for finding these fixed points than a general algorithm for finding Brouwer fixed-points).

Theorem A.7. *Let $f\colon B^n \to B^n$ be a continuous map from the closed unit ball in $\mathbb{R}^n$ to itself. Then f has a fixed point.*

Hence the same theorem holds with B^n replaced with any topological space homeomorphic to B^n.

Proof. If not, consider the map $g\colon B^n \to S^{n-1}$, where S^{n-1} is the unit sphere in $\mathbb{R}^n$, and $g(x)$ is given by the point on the ray from $f(x)$ in the direction of x (that intersects S^{n-1}, which is well-defined if f has no fixed point). Then g takes $S^{n-1} \subset B^n$ to itself. Hence if $h\colon S^{n-1} \to B^n$ is the inclusion, then $g \circ h$ is the identity, which means that any homology, cohomology, and homotopy group on S^{n-1} factors through that on B^n. But for $n \geq 2$, S^{n-1} has non-trivial homology, cohomology, and homotopy groups of degree greater than 0 (which B^n does not), and for $n = 1$ S^{n-1} has two connected components (and B^n has only one). □

[More generally, a map g from a topological space, X, to a subspace, Y, is called a *retraction* if g is the identity when restricted to Y. Hence, as above, if $h\colon Y \to X$ is the inclusion, then for any topological functor (or "functorial topological invariant"), either covariant or contravariant, its value on Y factors through that on X via $g \circ h$.]

Replacing B^n with an n-simplex, and taking successively finer triangulations of B^n, one can alternatively use *Sperner's lemma* to find the fixed point; this seems analogous to the above proof where one computes (co)homology via successively finer triangulations, except that Sperner's lemma identifies a particular triangle of the triangulation (which can be arbitrarily fine) that must contain a fixed point.

References

[AC88] N. Alon and F. R. K. Chung, *Explicit construction of linear sized tolerant networks*, Proceedings of the First Japan Conference on Graph Theory and Applications (Hakone, 1986), vol. 72, 1988, pp. 15–19. MR 975519

[ACH83] Roy L. Adler, Don Coppersmith, and Martin Hassner, *Algorithms for sliding block codes. An application of symbolic dynamics to information theory*, IEEE Trans. Inform. Theory **29** (1983), no. 1, 5–22. MR 711274

[AFKM86] R. Adler, J. Friedman, B. Kitchens, and B. Marcus, *State splitting for variable length graphs*, IEEE Transactions on Information Theory **32** (1986), no. 1, 108–115.

[AM84] Noga Alon and V. D. Milman, *Eigenvalues, expanders and superconcentrators (extended abstract)*, FOCS, 1984, pp. 320–322.

[AM85] N. Alon and V. D. Milman, λ_1, *isoperimetric inequalities for graphs, and superconcentrators*, J. Combin. Theory Ser. B **38** (1985), no. 1, 73–88. MR 782626 (87b:05092)

[BBM+10] Marie-Pierre Béal, Jean Berstel, Brian H. Marcus, Dominique Perrin, Christophe Reutenauer, and Paul H. Siegel, *Variable-length codes and finite automata*, Selected topics in information and coding theory, Ser. Coding Theory Cryptol., vol. 7, World Sci. Publ., Hackensack, NJ, 2010, pp. 505–584. MR 3290069

[BL06] Yonatan Bilu and Nathan Linial, *Lifts, discrepancy and nearly optimal spectral gap*, Combinatorica **26** (2006), no. 5, 495–519. MR 2279667 (2008a:05160)

[Buc86] Marshall W. Buck, *Expanders and diffusers*, SIAM J. Algebraic Discrete Methods **7** (1986), no. 2, 282–304. MR 87e:68088

[Hae78] Willem Haemers, *A generalization of the Higman-Sims technique*, Nederl. Akad. Wetensch. Indag. Math. **40** (1978), no. 4, 445–447. MR 515603

[Hae80] Wilhelmus Hubertus Haemers, *Eigenvalue techniques in design and graph theory*, Mathematical Centre Tracts, vol. 121, Mathematisch Centrum, Amsterdam, 1980, Dissertation, Technische Hogeschool Eindhoven, Eindhoven, 1979. MR 568704

[HH71] M. D. Hestenes and D. G. Higman, *Rank* 3 *groups and strongly regular graphs*, Computers in algebra and number theory (Proc. SIAM-AMS Sympos. Appl. Math., New York, 1970), 1971, pp. 141–159. SIAM–AMS Proc., Vol. IV. MR 0340088

[HJ85] Roger A. Horn and Charles R. Johnson, *Matrix analysis*, Cambridge University Press, Cambridge, 1985. MR 832183

[HJ90] ———, *Matrix analysis*, Cambridge University Press, Cambridge, 1990, Corrected reprint of the 1985 original. MR 1084815

[HJ13] ———, *Matrix analysis*, second ed., Cambridge University Press, Cambridge, 2013. MR 2978290

[HLW06] Shlomo Hoory, Nathan Linial, and Avi Wigderson, *Expander graphs and their applications*, Bull. Amer. Math. Soc. (N.S.) **43** (2006), no. 4, 439–561. MR 2247919

[HMS91] Chris D. Heegard, Brian H. Marcus, and Paul H. Siegel, *Variable-length state splitting with applications to average runlength-constrained (ARC) codes*, IEEE Trans. Inform. Theory **37** (1991), no. 3, part 2, 759–777. MR 1102783

[LP17] David A. Levin and Yuval Peres, *Markov chains and mixing times*, American Mathematical Society, Providence, RI, 2017, Second edition of [MR2466937], With contributions by Elizabeth L. Wilmer, With a chapter on "Coupling from the past" by James G. Propp and David B. Wilson. MR 3726904

[Tan84] R. Michael Tanner, *Explicit concentrators from generalized N-gons*, SIAM J. Algebraic Discrete Methods **5** (1984), no. 3, 287–293. MR 752035 (85k:68080)

Department of Computer Science, University of British Columbia, Vancouver, BC V6T 1Z4, CANADA.

E-mail address: jf@cs.ubc.ca

URL: http://www.cs.ubc.ca/~jf