Lorem Ipsum

Chase P. Ross
Yale School of Management

May 7, 2016

Contents

1 Sets **2**

2 Symmetric Matrices **2**
2.1 Prove eigenvalues of a real symmetric matrix are real. 3
2.2 Prove eigenvectors of a real symmetric matrix corresponding to district eigenvalues are orthogonal to each other. 4
2.3 Prove a real symmetric matrix is always diagonizable. 5

3 Positive Definite Matrices **5**
3.1 Positive Definite Tests 6

4 Similar Matrices **7**
4.1 Facts about similar matrices 7

5 Markov Matrices **8**

6 What About Non-Diagonizable Matrices? Jordan Blocks **9**

1 Sets

Definition. Set. Any well specified collection of elements. It may contain a finite or infinite number of elements.

Example. $\sqrt{2}$ is irrational (cannot be written as a fraction). That is, $\sqrt{2} \neq \frac{a}{b}, a, b \in \mathbb{N}$.

Proof. We use proof by contradiction.

1. Suppose not. Suppose we can write $\sqrt{2} = \frac{a}{b}$ where $a, b \in \mathbb{Z}$ (integers).
2. We can assume either a or b is not divisible by 2, because otherwise you could simplify further until one of the two numbers is not divisible by 2.
3. Notice if $n = \{1, 2, 3, 4, 5...\}$ then $n^2 = \{1, 4, 9, 16, 25...\}$. So a must be even because only even numbers squared result in even numbers, which $2b^2$ must be.
4. From (2), we know that b must be odd, because if either a or b is not divisible by 2 *and* a must be even, then b must be odd.
5. Since a is even, we can say $a = 2m$.
6. Then $2b^2 = (2m)^2 \rightarrow 2b^2 = 4m^2 \rightarrow b^2 = 2m^2$.
7. b must be an even number because $2m^2$ must be an even number, and so b must also be an even number which contradicts (2) because both b and a cannot be even. Therefore there is a contradiction.

$\square$

Example. All numbers are definable in under 11 words.

Proof. Suppose not. Then there exists some number that cannot be defined in under 11 words. $n =$ *smallest number not definable with less than 11 words.* $\square$

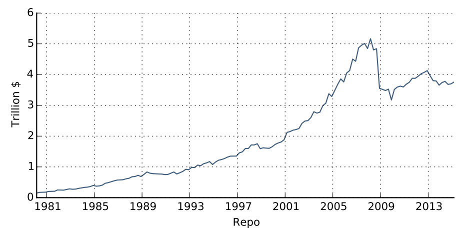
2 Symmetric Matrices

Recall matrix A is symmetric if $A = A^\intercal$ and orthogonal if $A^\intercal = A^{-1}$.

Definition. Real Symmetric Matrix: If A is a real symmetric matrix:

1. Eigenvalues of A are real.
2. Eigenvectors of A corresponding to district eigenvalues are orthogonal to each other.
3. A is always diagonizable; in fact, $Q^{-1}AQ = \Lambda$ is diagonal for some orthogonal matrix Q so $Q^{-1}AQ = \Lambda$ or $A = Q\Lambda Q^\intercal$. That is, $A = Q\Lambda Q^{-1} = Q\Lambda Q^\intercal$

Figure 1: Sample Figure



Example. We can see these properties hold for real symmetric matrices. For example: $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$. Going through the normal process to find eigenvalues and eigenvectors we find:

$$\lambda_1 = 1 \rightarrow x_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \lambda_2 = 3 \rightarrow x_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Thus the $\lambda' s$ are real and the eigenvectors are orthogonal to each other. We can make them unit vectors to produce Q which can diagonalize A:

$$Q = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix} \rightarrow A : A = Q\Lambda Q^{-1} = Q\Lambda Q^{\intercal} \text{ with } \Lambda = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$$

In other examples, the biggest problem will be that the eigenvectors won't be orthogonal to each other, e.g. $x_1 \perp x_2$ but not $x_1, x_2 \perp x_3$, which tends to happen with repeated eigenvalues. This is fixable with the Gram-Schmidt process, though.

We will now prove the 3 key facts of real symmetric matrices (mentioned above):

1. Eigenvalues of A are real.

2. Eigenvectors of A corresponding to district eigenvalues are orthogonal to each other.

3. A is always diagonizable; in fact, $Q^{-1}AQ = \Lambda$ is diagonal for some orthogonal matrix Q so $Q^{-1}AQ = \Lambda$ or $A = Q\Lambda Q^{\intercal}$. That is, $A = Q\Lambda Q^{-1} = Q\Lambda Q^{\intercal}$

2.1 Prove eigenvalues of a real symmetric matrix are real.

Proof. Eigenvalues of a real symmetric matrix are real.

First we need to mention complex numbers. Recall complex numbers take the form $x + iy$ with x, y real.

Definition. Complex Conjugate: if $z = x + iy$ then its complex conjugate is $\overline{z} = x - iy$. Some useful properties of complex conjugates:

- $\overline{z + w} = \overline{z} + \overline{w}$
- $\overline{zw} = \bar{z}\bar{w}$
- $z\overline{z} = |z|^2$
- z is real iff $z = \overline{z}$ (i.e. $0i$)
- You can "complex conjugate" a matrix by replacing each value with its complex conjugate
- Then, $\bar{A}\bar{B} = \overline{AB}$

Also note that for vector $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \overline{x}^{\mathsf{T}} x = [\overline{x_1} \dots \overline{x_n}] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \overline{x_1} x_1 + \dots \overline{x_n} x_n = |x_1|^+ \dots |x_n|^2 = ||x||^2$ which is greater than 0 unless each entry is also 0.

Goal: Show $\lambda = \bar{\lambda}$

Let A be a real, symmetric $n \times n$ matrix, thus $A = A^T$ and $A = \bar{A}$.

First: $Ax = \lambda x \to \bar{A}\bar{x} = \bar{\lambda}\bar{x} = A\bar{x} = \bar{\lambda}\bar{x} \to \bar{x}^T A = \bar{\lambda}\bar{x}^T \to \bar{x}^T A x = \bar{\lambda}\bar{x}^T x$

Second: $Ax = \lambda x \to \bar{x}^T A x = \bar{x}^T \lambda x$

Thus we see the left hand sides are equal, therefore the right sides are equal. They multiple $\bar{x}^T x =$ length squared with is greater than zero. Thus $\bar{\lambda} = \lambda$ and $a + ib = a - ib$ with imaginary part $b = 0$. Can do a similar process for Hermitian matrix remembering that $(Ax)^T = x^T A^T$ □

2.2 Prove eigenvectors of a real symmetric matrix corresponding to district eigenvalues are orthogonal to each other.

Proof. Eigenvectors of a real symmetric matrix corresponding to district eigenvalues are orthogonal to each other.

Let A be a real, symmetric matrix so its eigenvalues are real. If λ, μ are eigenvalues for $A, \lambda \neq \mu$ with real eigenvectors x, y. We want to show that $x \perp y$.

$$Ax = \lambda x \to x^{\mathsf{T}} A^{\mathsf{T}} = \lambda x^{\mathsf{T}} \to x^{\mathsf{T}} A = \lambda x^{\mathsf{T}} (A^{\mathsf{T}} = A) \to \overbrace{x^{\mathsf{T}} A y = \lambda x^{\mathsf{T}} y}^{1}$$

$$Ay = \mu y \to \overbrace{x^{\mathsf{T}} A y = x^{\mathsf{T}} \mu y}^{2}$$

Now combining the highlighted relationships:

$$\lambda x^{\mathsf{T}} y = \mu x^{\mathsf{T}} y \to (\lambda - \mu) x^{\mathsf{T}} y = 0 \text{ and we know } \lambda - \mu \neq 0 \to x^{\mathsf{T}} y = 0 \therefore x \perp y$$

□

2.3 Prove a real symmetric matrix is always diagonizable.

Proof. A real symmetric matrix is always diagonizable.

This proof comes from **Schur's Theorem**: If A is a real $n \times n$ matrix with real eigenvalues then $A = QTQ^{-1}$ with Q orthogonal $[(QQ^\intercal = I) \to (Q^{-1} = Q^\intercal)]$, and T is an upper triangular matrix. Thus this is a triangularizing theorem, not diagonalizing.[1] $T = Q^{-1}AQ \to T = Q^\intercal AQ \to T^\intercal = (Q^{-1}AQ)^\intercal = Q^\intercal A^\intercal Q \to Q^\intercal AQ = T$, thus $T = T^\intercal \to T$ is an upper triangular matrix, meaning it has 0's below diagonal and is also symmetric (so below diagonal = above diagonal =0's) $\to \therefore T$ is diagonal. □

Example. We've already seen A is not diagonizable: $A = \begin{bmatrix} 2 & 1 \\ -1 & 0 \end{bmatrix}$ because $\lambda_1 = \lambda_2 = 1$. But we can use Shur's theorem to help. The process is to find the first eigenvector, then find a second orthogonal vector to the first (this vector probably will not be an eigenvector of A, but that's okay), then make orthogonal matrix Q with these two vectors as usual and then solve for T using $T = Q^{-1}AQ = Q^\intercal AQ$. Thus we get:

$$T = Q^{-1}AQ = Q^\intercal AQ \to \frac{1}{2}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$$

which is clearly triangular. For a 3×3 matrix we'd get an orthogonal Q such that the bottom right 2×2 minor which would then need to go through the process again. Main point: this is laborious and requires doing it twice.

Definition. Hermitian Matrix: a matrix in which $A = \bar{A}^\intercal$ is a Hermitian matrix.

- If $A = \bar{A}^\intercal \to A$ is diagonizable and $A = Q\Lambda Q^{-1}$ with Λ a real diagonal matrix with eigenvalues (so same as before!), and $Q^{-1} = \bar{Q}^\intercal$ where Q is a **unitary matrix**.
- This is the true general form of what we've seen, as the conjugate has no impact if A is real. (!)

3 Positive Definite Matrices

Recall that a real symmetric matrix A has 3 important properties

- A has real eigenvalues.
- Eigenvectors for distinct eigenvalues of A are orthogonal.
- A is diagonizable.

Definition. Positive Definite Matrix: A real symmetric matrix A is positive definite iff $x^\intercal Ax > 0$ for every nonzero x.

[1] There is a more general treatment which eliminates the awkward real statement we're requiring now.

Example. Why do we care? Let's see an example.

1 variable: $f(x)$ has a local minimum if:

1. $f'(x) = 0$, and
2. $f''(x) = \frac{d^2 f}{dx^2} > 0$.

2 variables: $f(x, y)$ has local minimum if:

1. $\frac{\delta f}{\delta x} = \frac{\delta f}{\delta y} = 0$, and
2. $\begin{bmatrix} \frac{d^2 f}{dx^2} & \frac{d^2 f}{dxdy} \\ \frac{d^2 f}{dxdy} & \frac{d^2 f}{dy^2} \end{bmatrix}$ is positive definite.

3.1 Positive Definite Tests

How do we test whether a given matrix A is positive definite? 4 ways:

Theorem. A real symmetric matrix A is positive definite if any of the following holds:

1. $x^{\intercal} A x > 0$ for all nonzero vectors x.
2. All eigenvalues of A are > 0.
3. All pivots of A are > 0 with no row swaps allowed.
4. If the top left $k \times k$ det > 0 for $k = 1, 2, \ldots n$.

   For example; $\begin{bmatrix} a & b \\ b & c \end{bmatrix}$ is positive definite iff: $a > 0$ (first top left determinant) and $ac - b^2 > 0$ (second top left determinant).

A few examples are provided below:

Example. Is A positive definite?

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$$

Yes. Using rule (4) above; $2 > 0$ and $(2)(1) - (1) > 0 \rightarrow \therefore A$ is positive definite.

Example. For which values of $c4$ is A positive definite?

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & c \end{bmatrix}$$

Eliminate to:

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

By rule (3) above, if $c > \frac{2}{3}$ then A is positive definite.

Example. Is A positive definite?

$$A = \begin{bmatrix} 2 & 1 & & & \\ 1 & 2 & 1 & & \\ & 1 & 2 & 1 & \\ & \cdots & \cdots & \cdots & \cdots \\ & \cdots & 1 & 2 & 1 \end{bmatrix}$$

Eliminate without row swaps and use rule (3) above. The matrix after elimination is an upper triangle, and the bottom right most element is $\frac{n+1}{n}$. Therefore, by rule (3) the matrix is positive definite becuase all the pivots, including $\frac{n+1}{n}, > 0$.

It is not difficult to show that these four rules are equivalent, and are discussed in the handnotes in further detail. I'm skipping them now though.

4 Similar Matrices

Definition. Similar matrix: Let A, B be $n \times n$ matrices. A is similar to B if we can find an invertible matrix M so that $B = M^{-1}AM$. If A is similar to B:

1. A, B have the same eigenvalues (including reptitions)
2. A, B usually don't have the same eigenvectors

Proof. Show if A, B similar then they have the same eigenvalues.

Suppose $B = M^{-1}AM : M^{-1}(A - \lambda I)M = M^{-1}AM - \lambda M^{-1}IM = B - \lambda I$. Take determinants: $|B - \lambda I| = |M^{-1}(A - \lambda I)M| \rightarrow$ recall $|AB| = |A||B| \rightarrow |B - \lambda I| = |M^{-1}||A - \lambda I||M| = |A - \lambda I|$. Thus if A, B similar then $|B - \lambda I| = |A - \lambda I|$ so their characteristics polynomials are the same, therefore they have the same eigenvalues with the same algebraic multiplicity (so includes repetitions). □

Proof. What about eigenvectors?

Suppose x is an eigenvector for A with eigenvalue λ, and suppose $B = M^{-1}AM$. We know λ is an eigenvalue for B, but what about x?

First rewrite $B = M^{-1}AM \rightarrow A = MBM^{-1} \rightarrow Ax = \lambda x = MBM^{-1}x = \lambda x \rightarrow B(M^{-1}x) = M^{-1}\lambda x \rightarrow B(M^{-1}x) = \lambda(M^{-1}x) \rightarrow \therefore y = M^{-1}x$ is an eigenvector of B.

Note: y is nonzero ($y = M^{-1}x \neq 0$) since x is nonzero and M^{-1} is invertible.

Therefore, similar matrices share the same eigenvalues but not necessarily the same eigenvectors.

□

4.1 Facts about similar matrices

The following are key facts about similar matrices. If A similar to B:

1. If A similar to B, and B similar to C, then A similar to C.

2. If A similar to B (using M), then B similar to A (using M^{-1}).
3. A is always similar to A, (use $M = I$).
4. If A is diagonizable with eigenvalues $\lambda_1, \ldots \lambda_n$ (including repteitions) then A is similar to Λ, a diagonal matrix (like we've already seen, with eigenvalues on the diagonal). Then $A = S\Lambda S^{-1}$ where S is an eigenvector matrix. Just like we've seen before.
5. If A, B similar and A diagonizable then B is diagonizable. (B similar to A, A similar to Λ, then B similar to Λ and therefore diagonizable).
6. If A, B both diagonizable with the same eigenvalues then A, B are similar. So check if they have the same eigenvalues to see if similar.

Remark. What about nondiagonizable matrices? You can use the next best thing, a **Jordan Block**. But I won't cover this now.

5 Markov Matrices

Definition. A is a **Markov matrix** if

1. all entries of $A \geq 0$ and
2. the sum of any column of $A = 1$.

Theorem. If A is a positive Markov matrix (i.e. all entries > 0, each column adds to 1), then $\lambda_1 = 1$ is larger than any other eigenvalue. The eigenvectors x_1 corresponding to λ_1 is the steady state.

An example will clarify.

Example. For example, if you have two cable companies, x, y; suppose each month 20 percent of x's customers switch to y, and 5 percent of y's customers switch to x (assuming nobody entering or exiting). What happens in the long run. The beginning fraction with company x is 2%.

If we say that u_0 is the starting vector, we can see how this changes over time after multiplying by itself many times:

$$A = \begin{bmatrix} 0.8 & 0.05 \\ 0.2 & 0.95 \end{bmatrix}, u_1 = Au_0 \to u_2 = Au_1 = A^2 u_0$$

After k steps we will have $A^k u_0$. Thus these vectors, $u_1, u_2 \ldots$ approach a steady state. We want to find a steady state such that:

$$\begin{bmatrix} 0.8 & 0.05 \\ 0.2 & 0.95 \end{bmatrix} x = x$$

Now we will solve for the steady state:

1. Find the eigenvalues. A Markov matrix always has an eigenvalue of 1, and the other can be derived from the trace (the sum of the diagonals) as we know the sum of the eigenvalues = the trace. Thus $\lambda_1 = 1, \lambda_2 = 0.75$.

2. Find the eigenvectors. This is the normal process.

$$\lambda_1 = 1 : x_1 = \begin{bmatrix} 0.2 \\ 0.8 \end{bmatrix}, \lambda_2 = 0.75 : x_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

3. Now diagonalize and solve:

$$u_k = A^n u_0 \text{ and } A^n = S\Lambda^n S^{-1} \to u_k = S\Lambda^n S^{-1} u_0$$

$$u_k = \begin{bmatrix} 0.2 & 1 \\ 0.8 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0.75 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0.8 & -0.2 \end{bmatrix} \begin{bmatrix} 0.02 \\ 0.98 \end{bmatrix} = \begin{bmatrix} 0.2 \\ 0.8 \end{bmatrix}$$

4. So the eigenvector with $\lambda = 1$ is the steady state.

Theorem. If A is Markov and all entries > 0 then

1. $\lambda_1 = 1$ is an eigenvalue of A (also true of all entries are ≥ 0.)
2. If λ_1 is any other value, then $|\lambda| < 1$.
3. For eigenvalue $\lambda_1 = 1$ there's a unique eigenvector v_1 (normalized so sum of components is 1).
4. All components of v_1 are positive.
5. v_1 is unique steady state: for any vector $u_k, A^k u_0 \to cv_1$ where c is sum of components of vector u_0.

So in Markov matrices, if eigenvalues are distinct $\to A$ is diagonalizable $\to$ eigenvectors are linearly independent and importantly form a basis in R^n. Then you can find the eigenvector for $\lambda = 1$.

6 What About Non-Diagonizable Matrices? Jordan Blocks

What's the best that can be done?

Definition. A $k \times k$ **Jordan Block** with eigenvalue λ is $J \begin{bmatrix} \lambda_1 & 1 & \dots & & & \\ & \lambda_2 & 1 & \dots & & \\ & & \lambda_3 & 1 & \dots & \\ \vdots & & & & & \end{bmatrix}$

where λ is the only eigenvalue of J.

Therefore, $J - \lambda I = \begin{bmatrix} 0 & 1 & \dots & & & \\ & 0 & 1 & \dots & & \\ & & 0 & 1 & \dots & \\ \vdots & & & & & \end{bmatrix}$

Then: $(J - \lambda I)x = \begin{bmatrix} 0 & 1 & \dots & & & \\ & 0 & 1 & \dots & & \\ & & 0 & 1 & \dots & \\ \vdots & & & & & \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ \vdots \\ x_k \end{bmatrix} = 0$. So all eigenvectors of J are multiples of $\begin{bmatrix} 1 \\ 0 \\ \\ \vdots \\ 0 \end{bmatrix}$

J is not diagonizable if $k > 1$, eigenvalue λ has geometric multiplicity 1 (and only 1 linear independent eigenvector).

Theorem. Every $n \times n$ matrix A is similar to a **Jordan Form** $\begin{bmatrix} J_1 & \dots & & & \\ & J_2 & \dots & & \\ & & J_3 & 1 & \dots \\ \vdots & & & & \end{bmatrix}$ where $J_1 \dots J_2$ are Jordan Blocks for eigenvalues $\lambda_1, \dots, \lambda_s$ of A.

Note: Not going to learn how to reduce A to a Jordan Block.

Jordan Form for A is unique (except reordering blocks).

For example, every 3×3 matrix with eigenvalues 3,3,3 is similar to one and only one of:

1. $\begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}$ If the matrix is diagonizable, with geometric multiplicity 3.
2. $\begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{bmatrix}$ with a 2×2 Jordan Block, J, in the bottom right corner, with geometric multiplicity 2.
3. $\begin{bmatrix} 3 & 1 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{bmatrix}$ with a 3×3 Jordan Block, J with geometric multiplicity 1.
4. Note that none of these are similar to each other.