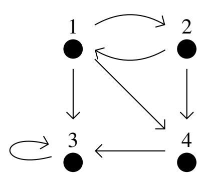
MA 253 Exam 2 - Study Guide

Included in this study guide are the "Basic Facts about Bases" study guide and the "Basic Facts about Eigenstuff" study guide. We begin, however, with some sample problems. These are not intended to be all-inclusive of what you should study, just to give you some additional problems to think about.

1. Sample Problems

(1) Consider the network below and let A be its transition matrix.



(a) What is the transition matrix for this system?
(b) What are its eigenvalues?
(c) Is there a single equilibrium vector or more than one?
(d) Do the eigenvectors form a basis for $\mathbb{R}^4$?

(2) Suppose that $T\colon \mathbb{R}^n \to \mathbb{R}^n$ is a linear transformation with the eigenvalue $\lambda = 1$ having multiplicity n.

(a) Might T be invertible?
(b) Must T be diagonalizable?

(3) Show that if A and B are $n \times n$ matrices, and if k is a real number, then

(a) $\mathrm{tr}(A+B) = \mathrm{tr}(A) + \mathrm{tr}(B)$
(b) $\mathrm{tr}(kA) = k\,\mathrm{tr}(A)$.

(4) Show that if λ is an eigenvalue for a matrix A, then λ^k is an eigenvalue for A^k. What can you say about the associated eigenvectors?

(5) Explain why if n is odd, then every $n \times n$ matrix A must have at least one real eigenvalue.

(6) Is it possible for an $n \times n$ matrix with entries in $\mathbb{R}$ to have exactly one eigenvalue which has a non-zero imaginary part?

(7) Explain why a matrix whose columns are a basis for $\mathbb{R}^n$ must be the matrix for a one-to-one linear transformation. Must such a matrix be invertible? What if it is square?

(8) Let $T\colon \mathbb{R}^n \to \mathbb{R}^n$ be reflection across the subspace $x_1 + x_2 + \ldots + x_n = 0$.

(a) Find all the eigenvalues of T and their algebraic multiplicities without doing any matrix computations.

(b) Find a basis $\mathscr{B}$ for $\mathbb{R}^n$ in which $[T]_{\mathscr{B}}$ is diagonal.

(9) If A is an $n \times n$ matrix such that there is an invertible matrix S and an upper triangular matrix U such that

$$A = SUS^{-1}$$

what is the relationship, if any between the eigenvalues of A and those of U? Are the eigenvalues of A easy to find? Why or why not?

(10) Suppose that $A = XY$ where A, X, Y are $n \times n$ matrices and X and Y are an upper triangular and lower triangular matrices. Explain why 0 is *not* an eigenvalue of A if and only if neither X nor Y has a 0 on the diagonal.

2. Bases

Here are some basic facts

(1) If $V \subset \mathbb{R}^n$ is a subspace, then a **basis** for $\mathbb{R}^n$ is a list of vectors $\mathbf{b}_1, \ldots, \mathbf{b}_k$ which are linearly independent and which span V.

(2) If $\mathbf{b}_1, \ldots, \mathbf{b}_k$ is a basis for a subspace V then each vector $\mathbf{x} \in \mathbb{R}^n$ can be uniquely written as a linear combination of $\mathbf{b}_1, \ldots, \mathbf{b}_k$.

(3) All bases for a subspace V have the same number of elements.

(4) Every subspace other than $\mathbf{0}$ has a basis.

3. Linear transformations $T: \mathbb{R}^n \to \mathbb{R}^m$

Here are some important facts:

(1) If $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ and if $k, \ell \in \mathbb{R}$, then

$$T(k\mathbf{v} + \ell\mathbf{w}) = kT(\mathbf{v}) + \ell T(\mathbf{w})$$

(2) There is an $m \times n$ matrix A such that $T(\mathbf{x}) = A\mathbf{x}$ for every $\mathbf{x} \in \mathbb{R}$. Furthermore, the columns of this matrix are

$$T(\mathbf{e}_1) \quad T(\mathbf{e}_2) \quad \ldots \quad T(\mathbf{e}_n)$$

where $\mathbf{e}_1, \ldots, \mathbf{e}_n$ are the standard basis vectors of $\mathbb{R}^n$. We often denote this matrix by the symbol $[T]$.

(3) If $\mathscr{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ is any basis for $\mathbb{R}^n$ and if $\mathscr{C} = \{\mathbf{c}_1, \ldots, \mathbf{c}_m\}$ is any basis for $\mathbb{R}^m$, then there is a matrix $[T]_{\mathscr{B}\mathscr{C}}$ such that for every $\mathbf{x} \in \mathbb{R}^n$ we have

$$T(\mathbf{x})_{\mathscr{C}} = [T]_{\mathscr{B}\mathscr{C}} \mathbf{x}_{\mathscr{B}}$$

where the subscripts mean that we write the vectors in coordinates from $\mathscr{B}$ or $\mathscr{C}$. The matrix $[T]_{\mathscr{B}\mathscr{C}}$ has columns:

$$T((\mathbf{b}_1)_{\mathscr{B}})_{\mathscr{C}} \quad T((\mathbf{b}_2)_{\mathscr{B}})_{\mathscr{C}} \quad \ldots \quad T((\mathbf{b}_n)_{\mathscr{B}})_{\mathscr{C}}$$

(4) The following are equivalent for T:

- The transformation T is one-to-one
- The columns of the matrix $[T]$ are linearly independent
- $\ker(T) = \{\mathbf{0}\}$
- The reduced row echelon form of $[T]$ has a leading one in every column

(5) The following are equivalent for T:

- The transformation T is onto
- The rows of T are linearly independent
- $\mathrm{im}(T) = \mathbb{R}^m$
- The reduced row echelon form of $[T]$ has a leading one in every row (i.e. no row of all zeros).

(6) If $n \neq m$, then T is not invertible.

4. Linear transformations $T: \mathbb{R}^n \to \mathbb{R}^n$

In the case when $n = m$, all of the previous facts remain true, but we can say even more:

(1) The following are equivalent:

- T is invertible
- T is one-to-one
- T is onto
- $\det[T] \neq 0$
- 0 is not an eigenvalue for T

(2) If the eigenvalues of T form a basis $\mathscr{B}$ for $\mathbb{R}^n$, then $[T]_{\mathscr{B}}$ is diagonal.

(3) If T has n distinct eigenvalues, then the eigenvalues for T form a basis for $\mathbb{R}^n$.

MA 253 Notes on Eigenvalues and Eigenvectors

Throughout these notes, let A be an $n \times n$ matrix and let $T: \mathbb{R}^n \to \mathbb{R}^n$ be the linear transformation $T(\mathbf{x}) = A\mathbf{x}$.

A number λ and a *non-zero* vector $\mathbf{v} \in \mathbb{R}^n$ are an **eigenvalue**, **eigenvector** pair if

$$A\mathbf{v} = \lambda\mathbf{v}.$$

In otherwords, an eigenvector is a vector which is stretched by the transformation and λ is the amount it is stretched by.

Here are some useful facts about eigenvalues and eigenvectors. Some of these were explained in class and some were not.

- If A has eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ which form a basis $\mathscr{B}$ for $\mathbb{R}^n$ then the matrix $[T]_{\mathscr{B}}$ for T in that basis is diagonal. Put another way, there is an invertible matrix B and a diagonal matrix D so that

  $$A = BDB^{-1}$$

  Remark: The columns of B are the eigenvectors (in any order you choose) and the diagonal entries of D are the eigenvalues with the kth entry on the diagonal the eigenvalue corresponding to the kth column of B. Different choices of eigenvectors or a different choice of eigenvalue will make B and D different matrices.

- The eigenvalues of A can be found by solving the equation

  $$\det(A - \lambda I) = 0$$

  for λ.

  The polynomial $\det(A - \lambda I)$ is a polynomial of degree n. It is called the **characteristic polynomial** of A.

- If λ_0 is an eigenvalue for A, the **eigenspace** for λ_0 is the set of eigenvectors for A corresponding to λ_0 together with the vector $\mathbf{0}$ (which never counts as an eigenvector). A basis for the eigenspace for λ_0 can be found by finding a basis for

  $$\ker(A - \lambda_0 I).$$

- If A has n distinct eigenvalues then the eigenvectors for A form a basis for $\mathbb{R}^n$ and so A can be diagonalized, as above.

- You should be able to instantly (i.e. no computations) find eigenvectors for projections, reflections, rotations (if they exist), and certain transition matrices.

- The characteristic polynomial $f(\lambda)$ for A has constant term equal to the determinant of A and the coefficient of the $(n-1)$st term is equal to $\pm$ the trace of A. These facts follow from the more interesting fact that the determinant of A is the product of the eigenvalues and the trace of A is the sum of the eigenvalues.
- The matrix A is invertible if and only if 0 is *not* an eigenvalue of A.
- The Perron-Frobenius Theorem (see below)
- If z is a complex number which is a root of the characteristic polynomial $f(\lambda)$ of A then its complex conjugate $\bar{z}$ is also a root of $f(t)$. There is a certain sense in which these complex eigenvalues, if they have non-zero imaginary part, correspond to a certain kind of rotating action of A on $\mathbb{R}^n$.
- If A is upper or lower triangular, the eigenvalues for A appear as the diagonal entries of A (with their correct multiplicities)
- The eigenvalues of A and of the transpose A^T of A are the same, although the eigenvalues may be different.
- If A and X are similar matrices (that is, there is a change of basis which takes one to the other) then they have the same eigenvalues. The coordinates of their eigenvectors may differ.

5. A partial proof of the Perron-Frobenius Theorem

An $n \times n$ matrix A is a **transition** matrix if the entries of A are non-negative real numbers and if each column of A sums to 1. A **distribution vector** $\mathbf{x} \in \mathbb{R}^n$ is a vector with non-negative entries summing to 1. Notice that if A is a transition matrix and if $\mathbf{x}$ is a distribution vector then $A\mathbf{x}$ is also a distribution vector. (You should verify this!)

A distribution vector $\mathbf{w}$ is an **equilibrium** for a transition matrix A if $A\mathbf{w} = \mathbf{w}$. Observe that this is equivalent to saying that 1 is an eigenvalue for A with associated eigenvector $\mathbf{w}$ which is also a distribution vector.

Theorem (Perron-Frobenius)**.** Suppose that A is a transition matrix. Then the following hold:

(1) there is an equilibrium vector $\mathbf{w}$ for A
(2) If λ is an eigenvalue for A then $|\lambda| \leq 1$
(3) If the eigenvectors of A form a basis, then for any initial distribution vector $\mathbf{x}_0$, there is an equilibrium vector $\mathbf{w}$ for A such that

$$\lim A^m \mathbf{x}_0 = \mathbf{w}.$$

Proof. It is easier to work with the transpose A^T of A rather than A itself, so recall that A^T and A have the same eigenvalues. Suppose, first that λ is an eigenvalue of

A^T (and hence of A) let $\mathbf{v}$ be an associated eigenvector. Since $\mathbf{v} \neq \mathbf{0}$, we may scale $\mathbf{v}$ so that one of the entries $v_j = 1$ and all other entries are at most 1 in absolute value. Let

$$\mathbf{r}_j = \begin{pmatrix} r_{j1} & r_{j2} & \dots & r_{jn} \end{pmatrix}$$

be the jth row of A^T. By the definition of matrix multiplication, the jth entry of the vector $A\mathbf{v}$ is equal to

$$\begin{pmatrix} r_{j1} & r_{j2} & \dots & r_{jn} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = r_{j1}v_1 + r_{j2}v_2 + \ldots + r_{jj}v_j + \ldots + r_{jn}v_n.$$

But since $A\mathbf{v} = \lambda\mathbf{v}$, the jth entry of $A\mathbf{v}$ is just $\lambda v_j = \lambda$. Thus,

$$\lambda = r_{j1}v_1 + r_{j2}v_2 + \ldots + r_{jj}v_j + \ldots + r_{jn}v_n.$$

Taking absolute values and using the triangle inequality, we get:

$$\begin{aligned} |\lambda| &= |r_{j1}v_1 + r_{j2}v_2 + \ldots + r_{jj}v_j + \ldots + r_{jn}v_n| \\ &\leq |r_{j1}v_1| + |r_{j2}v_2| + \ldots + |r_{jj}v_j| + \ldots + |r_{jn}v_n| \end{aligned}$$

Each r_{ji} is non-negative, so

$$|\lambda| \leq r_{j1}|v_1| + r_{j2}|v_2| + \ldots + r_{jj}|v_j| + \ldots + r_{jn}|v_n|$$

Each $|v_i| \leq 1$, so

$$|\lambda| \leq r_{j1} + r_{j2} + \ldots + r_{jj} + \ldots + r_{jn}$$

Since A^T is the transpose of a transition matrix, each of its rows sums to 1. Hence, $\lambda \leq 1$ as desired. Notice also in the previous work each inequality $\leq$ is actually an $=$ exactly when all the entries of $\mathbf{v}$ are equal to 1. Thus, the vector $\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$ is an eigenvector for $\lambda = 1$. Thus, $\lambda = 1$ is an eigenvalue for both A^T and A.

Suppose now that the eigenvectors for A form a basis $\mathbf{b}_1, \ldots, \mathbf{b}_n$ of $\mathbb{R}^n$. Let $\lambda_1, \ldots, \lambda_n$ be the associated eigenvalues. We may choose the ordering so that $\lambda_1, \ldots, \lambda_k$ are exactly the eigenvalues equal to 1. Let $\mathbf{x}_0$ be an initial distribution vector. We can write $\mathbf{x}_0$ uniquely in terms of the basis as

$$\mathbf{x}_0 = \alpha_1\mathbf{b}_1 + \alpha_2\mathbf{b}_2 + \ldots + \alpha_n\mathbf{b_n}.$$

Applying A^m to both sides, we get

$$A^m\mathbf{x}_0 = \alpha_1 A^m\mathbf{b}_1 + \alpha_2 A^m\mathbf{b}_2 + \ldots + \alpha_n A^m\mathbf{b_n}.$$

Using the fact that the basis vectors are all eigenvectors we get:

$$A^m\mathbf{x}_0 = \alpha_1\lambda_1^m\mathbf{b}_1 + \alpha_2\lambda^m\alpha_2 + \ldots + \alpha_k\lambda^m\mathbf{b}_k + \alpha_{k+1}\lambda^m\mathbf{b}_{k+1} + \ldots + \alpha_n\lambda_n\mathbf{b}_n.$$

Recall that $\lambda_1 = \lambda_2 = \ldots = \lambda_k = 1$ and that for each $i \geq k+1$, we have $|\lambda_i| \leq 1$. Thus, taking a limit as $m \to \infty$ we get

$$\lim A^m \mathbf{x}_0 = \alpha_1 \mathbf{b}_1 + \alpha_2 \mathbf{b}_2 + \ldots + \alpha_k \mathbf{b}_k.$$

Each vector $A^m \mathbf{x}_0$ is a vector with non-negative entries summing to 1, so the limit of $A^m \mathbf{x}_0$ is also such a vector. Thus,

$$\mathbf{w} = \alpha_1 \mathbf{b}_1 + \alpha_2 \mathbf{b}_2 + \ldots + \alpha_k \mathbf{b}_k$$

is an equilibrium vector (since it is a distribution vector in the eigenspace of the eigenvalue 1) and any initial distribution vector will limit to it.

$\square$