# WEEK 2: REVIEW OF (FINITE-DIMENSIONAL) LINEAR ALGEBRA. PART II.

This section of notes introduces the basics of linear transformations on finite-dimensional vector spaces. The primary topics are

- linear transformations,
- changes of basis,
- diagonalization, and
- orthonormal bases in inner product spaces.

All of these results should be familiar from a course in elementary linear algebra, although the context might be slightly more abstract. In order to save time and space, we will mostly list results and omit proofs; the notation is adapted from the first chapter of M. Frazier's book *An Introduction to Wavelets Through Linear Algebra*, where all proofs can be found. (This book is available electronically from ASU library.)

## 1. Finite-Dimensional Linear Algebra in General Spaces

Throughout this section of notes, $\mathbb{F}$ represents a field, either $\mathbb{C}$ or $\mathbb{R}$.

### 1.1. Matrices.

**Definition 1.1.** Let $\mathbb{F}$ be a field. An $m \times n$ *matrix* over $\mathbb{F}$ is an array of the form

$$A = \left( \begin{array}{cccc} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn}, \end{array} \right)$$

where $a_{ij}$, called the $ij$-th entry of $A$, is in the field $\mathbb{F}$ for $i = 1, ..., m$ and $j = 1, ..., n$. $A$ is sometimes denoted by $(a_{ij})$ or $[a_{ij}]$.

The operations of addition and scalar multiplication on matrices are defined pointwise as on $\mathbb{F}^n$. We can also define matrix multiplication for matrices of appropriate dimensions.

**Definition 1.2.** Let $A = [a_{ij}]$ and $B = [b_{ij}]$ be $m \times n$ and $n \times r$ matrices over a field $\mathbb{F}$. Define the $m \times r$ matrix $C$ by

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}.$$

The matrix $C$ is called the *product* of $A$ and $B$ and is denoted $C = A * B$. As with multiplication of scalars in $\mathbb{F}$, the $*$ is often omitted.

Recall that matrix multiplication is *not* commutative, but is associative, and recall that, although we often write ordered $n$-tuples as row vectors, we always treat them as column vectors—that is, matrices with one column—when discussing matrix algebra.

1.2. **Linear Maps.**

Next we define a linear map on vector spaces and representation of them by matrices.

**Definition 1.3.** Suppose $V$ and $W$ are vector spaces over $\mathbb{F}$ and $T$ is a function from $V$ to $W$ such that for every $u$, $v \in V$ and $a \in \mathbb{F}$, we have that

- $T(av) = aT(v)$ and
- $T(u+v) = T(u) + T(v)$.

Then we say that $T$ is a *linear map* from $V$ to $W$.

It is common to omit the parentheses around the arguments of linear functions, i.e., to write $Tv$ for $T(v)$. Linear maps are also referred to as linear functions or linear transformations.

Next we discuss how a representation of a linear map changes from one space to another.

**Theorem 1.4.** *Let $V$ and $W$ be vector spaces over $\mathbb{F}$ with bases $R = \{v_1, ..., v_n\}$ and and let $S = \{w_1, ..., w_m\}$, respectively. Then $T : V \to W$ is a linear map if and only if there is an $m \times n$ matrix $A$ representing $T$ in the bases $R$ and $S$ in the sense that*

$$[T(v)]_S = A[v]_R,$$

*i.e.,*

$$T(v_j) = a_{1j}w_1 + a_{2j}w_2 + .... + a_{mj}w_m$$

*for all $j = 1, ..., n$. Moreover, $T$ and $A$ uniquely determine one another. We say that $A$* represents $T$ with respect to $R$ and $S$.

Let $L(V, W)$ be the space of linear functions from $V$ to $W$ as above, and $M_{m\times n}$ be the space of $n \times m$ matrices over $\mathbb{F}$. The above lemma shows that $L(V, W)$ and $M_{m\times n}$ are in one-to -one correspondence. When the choice of basis is obvious, it is common to identify $T$ and $A$, although technically they are quite different mathematical objects except in the case when $V = \mathbb{F}^n$ and $W = \mathbb{F}^m$ and we are using the standard Euclidean basis. It is worth emphasizing that the matrix $A$ depends heavily on the basis, as we will see when we look at diagonalization of linear maps.

**Proposition 1.5.** *Let $V$ be an $n$-dimensional vector space over $\mathbb{F}$., and let $R = \{v_1, ..., v_n\}$ and $S = \{w_1, ..., w_n\}$ be two bases for $V$. By the definition of a basis, there are unique scalars $a_{ij}$, $i, j = 1, .., n$, such that for each $j$, we have*

$$u_j = a_{ij}v_1 + a_{2j}v_2 + ... + a_{nj}v_n.$$

*Let the matrix $A$ be defined by $A = (a_{ij})$. Then $A$ is the unique matrix such that for each $v \in V$, $[v]_S = A[v]_R$. We call $A$ the $R$ to $S$* change of basis *matrix.*

If we know $R$ to $S$ *change of basis* matrix, how can one find the $S$ to $R$ *change of basis* matrix? For that we need to recall a few definitions: injections, or one-to-one maps; surjections, or onto maps; invertible matrices.

**Definition 1.6.** Let $V$ and $W$ be vector spaces over $\mathbb{F}$ and let $T$ be a linear map $V \to W$. The *range* of $T$, sometimes called the *image* of $T$, is the set

$$\{w \in W : \text{ there exists } v \in V \text{ such that } \quad w = T(v)\}.$$

The *kernel* of $T$, denoted by $KerT$, is the set $\{v \in V : T(v) = 0\}$.

**Example 1.7.** If $P_{xy}$ is a projection in $\mathbb{R}^3$ onto $xy$ plane, then $Ker(P_{xy}) =$ $z$-axis.

*Remark* 1.8. Note that by linearity, the map $T$ is injective if and only if $\ker(T) = \{0\}$. Also note that the range of $T$ is always a subspace of $W$. The map $T : V \to W$ is surjective if the range of $T$ is the entire vector space $W$.

Recall the definition of $I_{n\times n}$, the square matrix with ones along the diagonal and zeros off the diagonal, which acts as an identity when pre- or post-multiplied by any matrix of suitable dimension.

**Definition 1.9.** Let $A$ be an $n \times n$ matrix over a field $\mathbb{F}$. If there is a matrix $A^{-1}$ such that $AA^{-1} = A^{-1}A = I_{n\times n}$, then we say that $A$ is *invertible* and $A^{-1}$ is its *inverse*.

From Theorem 1.4 comes the following:

**Proposition 1.10.** *Let $V$, $W$, $R$, $S$ and $A$ be as in Theorem 1.4. Then $T$ is invertible if and only if $A$ is.*

**Definition 1.11.** Let $A$ be an $m \times n$ matrix over $\mathbb{F}$. By Theorem 1.4, $A$ defines a linear transformation $T : \mathbb{F}^m \to \mathbb{F}^n$. Define the *rank* of $A$ to be the dimension of the range of $T$.

**Proposition 1.12.** *Let $A$ be an $n \times n$ matrix over $\mathbb{F}$. Then $A$ is invertible if and only if rank $A = n$.*

**Proposition 1.13.** *If $V$ is an $n$-dimensional vector space over $\mathbb{F}$, and $R = \{v_1, ..., v_n\}$ and $S = \{w_1, ..., w_n\}$ are two bases for $V$, then the $R$ to $S$ change of basis matrix is invertible, and its inverse is the $S$ to $R$ change of basis matrix.*

**Definition 1.14.** Let $V$ be a finite dimensional vector space over $\mathbb{F}$ with basis $R$, and let $T : V \to V$ be linear. Let the matrix $A_{T,R}$ be as in Theorem 1.4, i.e.,

$$[T(v)]_R = A_{T,R}[v]_R$$

for all $v \in V$. We call $A_{T,R}$ the *matrix that represents $T$ with respect to $R$.*

**Proposition 1.15.** *Let $V$ be a finite dimensional vector space over $\mathbb{F}$, and let $R$ and $S$ be two bases for $F$. Let $T : V \to V$ be linear, and let $A_{T,R}$ and $A_{T,S}$ represent $T$ with respect to $R$ and $S$, respectively. Let $C$ be the $R$ to $S$ change of basis matrix. Then*

$$A_{T,S} = CA_{T,R}C^{-1}, \tag{1.1}$$

*or equivalently,*

$$A_{T,R} = C^{-1}A_{T,S}C. \tag{1.2}$$

The relation between $A_{T,R}$ and $A_{T,S}$ in (1.1) and (1.2) look alike to another concept from Linear Algebra, namely, similarity between matrices. We recall this next.

**Definition 1.16.** Let $A$ and $B$ be square matrices over $\mathbb{F}$. We say $A$ and $B$ are *similar* if there is an invertible matrix $P$ such that

$$A = P^{-1}AP.$$

It is easy to show that similarity is an equivalence relation.

**Proposition 1.17.** *Let $R$ be a basis for an $n$-dimensional vector space $V$ over $\mathbb{F}$, and let $P$ be any invertible $n\times n$ matrix over $\mathbb{F}$. Then $V$ has another basis $S$ such that $P$ is the $R$ to $S$ change of basis matrix.*

**Corollary 1.18.** *Suppose $T$ is a linear map from a finite dimensional vector space $V$ over $\mathbb{F}$ to itself. Let $R$ be a basis for $V$, and let $A$ represent $T$ with respect to $R$. Then if $B$ is any matrix similar to $A$, there is a basis $S$ for $V$ such that $B$ represents $T$ with respect to $S$.*

Suppose now that there is some linear transformation that we would like to make repeatedly on some vector space. Given any basis for that space, we know there is a matrix $A$ that represents this transformation with respect to that basis. Then, if we can find the matrix $B$ similar to $A$ which allows for the fastest, easiest computation, and we find the change-of-basis matrix necessary to convert $A$ to $B$, and we find a fast algorithm to make the change of basis, then we can considerably improve our ability to quickly compute the transformation. This sounds like a lot of ifs, but it turns out that the Fourier basis meets all these requirements for a large class of transformations, called shift-invariant linear transformations. This is why Fourier analysis is so useful.

### 1.3. Some basic facts from Spectral Theory.

**Definition 1.19.** Let $V$ be a vector space over $\mathbb{F}$, and $T$ be a linear map from $V$ to itself.

A nonzero vector $v \in V$ is called an *eigenvector* of $T$ if there is a scalar $\lambda \in \mathbb{F}$ such that

$$T(v) = \lambda v.$$

We call $\lambda$ an *eigenvalue* of $T$ corresponding to $v$.

Given a particular eigenvalue $\lambda$ of $T$, the set $E_\lambda(T)$ of all eigenvectors of $T$ with eigenvalue $\lambda$, clearly a subspace by linearity, is called the *eigenspace* of $T$ corresponding to $\lambda$. The *geometric multiplicity* of $\lambda$ is the dimension of $E_\lambda$.

We can also define the eigenvectors, eigenvalues, and eigenspaces of an $n \times n$ matrix $A$ over $\mathbb{F}$, simply by viewing $A$ as a linear operator on $\mathbb{F}^n$ in the obvious way.

**Proposition 1.20.** *Let $V$ be a finite dimensional vector space over $\mathbb{F}$ with basis $R$. Let $T : V \to V$ be linear and $A$ represent $T$ with respect to $R$. Then $A$ and $T$ have the same eigenvalues,*

$$v \in E_\lambda(T) \quad \textit{if and only if} \quad [v]_R \in E_\lambda(A).$$

*Also, that the eigenvalues have the same geometric multiplicities, i.e.,* $\dim E_\lambda(T) = \dim E_\lambda(A)$.

**Corollary 1.21.** *Let $A$ and $B$ be similar matrices. Then $A$ and $B$ have the same eigenvalues of the same geometric multiplicities.*

**Proposition 1.22.** *Let $V$ be an $n$-dimensional vector space over $\mathbb{F}$ and let $T : V \to V$ be linear with eigenvalues $\lambda_1, ..., \lambda_K$. For each $k = 1, ..., k$, let $R_k$ be a basis for $E_{\lambda_k}$. Let $R$ be the union of the $R_k$'s. Then $R$ is linearly independent.*

**Corollary 1.23.** *Let $V$ and $T$ be as above. The sums of the geometric multiplicities of all eigenvalues of $T$ is at most $n$. This also shows that $T$ cannot have more than $n$ distinct eigenvalues.*

Note that, by Proposition 1.20, the proposition and corollary above also hold for any $n \times n$ matrix $A$.

**Definition 1.24.** We say that a linear function $T$ from a finite-dimensional vector space to itself is diagonalizable if $V$ has a basis consisting of eigenvectors of $T$.

**Definition 1.25.** A square matrix is *diagonal* if all the entries not on the main diagonal are zero, i.e., if $a_{ij} = 0$ whenever $i \neq j$. A square matrix is *diagonalizable* if it is similar to a diagonal matrix.

**Proposition 1.26.** *Let $V$ be a finite dimensional space, and $T$ a linear transformation on $V$. Then*

- *$T$ is diagonalizable if and only if there is a basis $R$ for $V$ such that the matrix $A_{T,R}$ representing $T$ with respect to $R$ is diagonal, and*

- *if $S$ is any basis for $V$, then $T$ is diagonalizable if and only if the matrix $A_{T,S}$ representing $T$ with respect to $S$ is diagonalizable.*

Diagonal matrices are by far the easies matrices to work with. One can very rapidly compute matrix products and powers of a diagonal matrix, whereas for a non-diagonal matrix of any significant size such computations require a prohibitively large number of individual calculations. Next, we look at how to actually go about diagonalizing a linear map:

**Proposition 1.27.** *Let $A$ be an $n \times n$ diagonalizable matrix.*

(a) *Suppose $A = P^{-1}DP$, where $D$ is diagonal. Then the columns of $P$ are the (linearly independent) eigenvectors of $A$, and their eigenvalues are the corresponding entries in $D$.*
(b) *Conversely, if the $n$ linearly independent eigenvectors of $A$ are $v_1, ..., v_n$, with corresponding eigenvalues $\lambda_1, .., \lambda_n$ (not necessarily distinct), and we let $P$ be the square matrix with columns $v_1, ..., v_n$ and $D$ be the diagonal matrix with entries $\lambda_1, ..., \lambda_n$, in the same order, then we have $A = P^{-1}DP$.*

**Definition 1.28.** Let

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

be a $2 \times 2$ matrix over $\mathbb{F}$. The *determinant* of $A$ is defined by

$$\det A = ad - bc.$$

We define determinants for larger matrices inductively. Define the $M_{ij}$ minor of an $n \times n$ matrix $A$ to be the $(n-1) \times (n-1)$ matrix gotten be deleting the $i$th row and $j$th column of $A$. Then the determinant of $A$ is defined to be

$$\det A = \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \det M_{1j}.$$

**Proposition 1.29.** *Let $A$ and $B$ be an $n \times n$ matrices.*

$$\det AB = \det A \det B.$$

**Proposition 1.30.** *Let $A$ be an $n \times n$ matrix. Then $A$ is invertible if and only if* $\det A \neq 0$.

**Definition 1.31.** Let $A$ be an $n \times n$ matrix. The *characteristic polynomial* of $A$ is the polynomial defined by

$$\det(\lambda I - A),$$

regarded as a function of $\lambda$.

**Proposition 1.32.** *Let $A$ be an $n \times n$ matrix. The eigenvalues of $A$ are the roots of the characteristic polynomial of $A$.*

It follows from the fact that polynomials in $\mathbb{C}$ factor completely that if $A$ is a square matrix over $\mathbb{C}$, then the characteristic polynomial of $A$ can be written as a product of terms $(\lambda - \lambda_k)^{m_k}$, where the $\lambda_k$'s are the distinct eigenvalues of $A$. The number $m_k$ is called the *algebraic multiplicity* of $\lambda_k$. It can be shown that an eigenvalue's geometric multiplicity is less than or equal to its algebraic multiplicity. It follows that a matrix is diagonalizable over $\mathbb{C}$ if and only if the geometric multiplicity of all its eigenvalues is equal to the algebraic multiplicity.

**Proposition 1.33.** *Let $A$ and $B$ be similar matrices. Then*

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

*for all $\lambda$. It follows that the algebraic multiplicities of eigenvalues of $A$ and $B$ are the same.*