Math S-21b – Lecture #6 Notes

General Linear Spaces (Vector Spaces)

Though we have dealt exclusively so far with $\mathbf{R}^n$ and its subspaces, almost everything that we have developed so far will work the same way in any space where we can add elements and scale elements in a manner analogous to the way we add and scale vectors.

Definition: A **vector space** *V* is a set, with addition and scaling of element defined for all elements of the set, that is closed under addition and scaling, contains a zero element (0), and satisfies the following axioms:
For all $f, g, h \in V$ and scalars c, c_1, c_2

(1) $(f+g)+h = f+(g+h)$	(5) $c(f+g) = cf+cg$
(2) $f+g = g+f$	(6) $(c_1+c_2)f = c_1 f + c_2 f$
(3) $f+0 = f$	(7) $c_1(c_2 f) = (c_1 c_2) f$
(4) $f+(-f) = 0$	(8) $1f = f$

We'll deal primarily with the case where the scalars are *real numbers*. Such a vector space is called a **real vector space**. We could also use complex scalars in which case we'd call this a **complex vector space**.

Though we can give definitions and prove theorems about vector spaces in general, it's helpful to develop a library of examples to which we can refer.

1. $\mathbf{R}^n$ is a vector space. Indeed, the motivation for our definition and axioms is to define vector spaces to be spaces which a fundamentally like $\mathbf{R}^n$. All of the required axioms are familiar facts about vectors in $\mathbf{R}^n$.
2. Any subspace of $\mathbf{R}^n$ is a vector space. All of the axioms are inherited and every subspace contains the zero vector, and the definition of subspace ensures that a subspace is closed under addition and scaling.
3. The **complex numbers** $\mathbb{C} = \{a+bi : a, b \text{ are real numbers, } i^2 = -1\}$ can be viewed as a real vector space where addition is defined by $(a_1+b_1 i)+(a_2+b_2 i) = (a_1+a_2)+(b_1+b_2)i$ and scaling by a real number c is defined by $c(a+bi) = ca+cbi$. Note that <u>we do not define multiplication of complex numbers</u> in this context. The complex numbers contain the zero element $0 = 0+0i$ and all of the axioms follow from corresponding facts about real numbers.
4. $\mathbf{R}^{m\times n} = M(m,n) = \{m\times n \text{ matrices with real entries}\}$ is a real vector space with addition and scaling of matrices defined component-wise. The $m\times n$ zero matrix is the zero element and the axioms are all known properties of matrix algebra. Note that in this context <u>we do not define the product of matrices</u>.
5. $F(\mathbf{R},\mathbf{R}) = F(\mathbf{R}) = \{\text{functions } f:\mathbf{R}\to\mathbf{R} \text{ with domain } \mathbf{R}\}$ is a real vector space where addition of functions and scaling of functions are defined by pointwise by $(f+g)(x) = f(x)+g(x)$ and $(cf)(x) = c\,f(x)$. The zero element in this case is the function that is identically zero for all *x*. (This is quite different than just the real number 0.) Once again, the axioms all follow from familiar facts about real numbers.
6. $P_n = \{\text{real polynomials of degree} \le n\} = \{a_0 + a_1 x + \cdots + a_n x^n : a_0, a_1, \cdots, a_n \in \mathbf{R}\}$ is a real vector space. Note that we must allow all polynomials less than or equal to *n* because we might add two polynomials (or scale by 0) and get a polynomial of lesser degree.
7. $C^0(\mathbf{R},\mathbf{R}) = C^0(\mathbf{R}) = \{\text{continuous functions } f:\mathbf{R}\to\mathbf{R}\}$ is a real vector space. Closure follows from theorems of Calculus that the sum of continuous functions is continuous and a scalar multiple of a continuous function is also continuous. The zero function is clearly continuous, and the axioms are all easily verified.

Definition: A **subspace** *W* of a vector space *V* is a subset that is closed under addition and scaling of elements. That is, for any vectors $\mathbf{v}_1, \mathbf{v}_2 \in W$ and scalars c_1, c_2, it must be the case that $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 \in W$. We write $W \subseteq V$.

8. $C^1(\mathbf{R},\mathbf{R}) = C^1(\mathbf{R}) = \{\text{differentiable functions } f:\mathbf{R}\to\mathbf{R}\}$ is a real vector space. Closure follows from the theorems of Calculus that $(f+g)' = f' + g'$ and $(cf)' = c\,f'$. The zero function is clearly differentiable.

Note: All polynomials are differentiable, and a (hopefully) familiar theorem of Calculus tells us that differentiable functions must be continuous, so for any *n*, $P_n \subseteq C^1(\mathbf{R}) \subseteq C^0(\mathbf{R}) \subseteq F(\mathbf{R})$.

9. $C^k(\mathbf{R},\mathbf{R}) = C^k(\mathbf{R}) = \{\text{functions } f:\mathbf{R}\to\mathbf{R} \text{ that are at least } k \text{ times differentiable}\}$ is a real vector space. This follows similarly from Calculus theorems. The zero function is differentiable to all orders.

10. $C^\infty(\mathbf{R},\mathbf{R}) = C^\infty(\mathbf{R}) = \{\text{functions } f:\mathbf{R}\to\mathbf{R} \text{ that are differentiable to all orders}\}$ is a real vector space. This also follows from the Calculus theorems above. The zero function is differentiable to all orders.

Note: For any *n*, $P_0 \subseteq P_1 \subseteq P_2 \subseteq \cdots \subseteq P_{n-1} \subseteq P_n \subseteq C^\infty(\mathbf{R}) \subseteq \cdots \subseteq C^k(\mathbf{R}) \subseteq \cdots \subseteq C^1(\mathbf{R}) \subseteq C^0(\mathbf{R}) \subseteq F(\mathbf{R})$. It's also important to note that when dealing with spaces of functions, these spaces are much "larger" that $\mathbf{R}^n$. To better understand this, we'll need some more definitions, but some of the important details will have to wait until you take course in analysis and topology.

Many of the definitions when working with vectors spaces are essentially the same as those in $\mathbf{R}^n$.

Definition: Given a collection of elements $\{f_1, f_2, \ldots, f_k\} \in V$, we define the **span** of these elements as:

$$\text{span}\{f_1,\ldots,f_k\} = \{c_1 f_1 + \ldots + c_k f_k \text{ where } c_1,\ldots c_k \text{ are scalars}\}.$$

Definition: A set of elements $\{f_1, f_2, \ldots, f_k\} \in V$ is called **linearly independent** if given any linear combination of the form $c_1 f_1 + \cdots + c_k f_k = 0$, this implies that $c_1 = \cdots = c_k = 0$. That is, there is no nontrivial way to combine these vectors to yield the zero element.

Definition: Given a subspace $W \subseteq V$, a collection of elements $\{f_1, f_2, \ldots, f_k\} \in W$ is called a **basis** of *W* if $\text{Span}\{f_1, f_2, \ldots, f_k\} = W$ and $\{f_1, f_2, \ldots, f_k\}$ are linearly independent.

A basis is a minimal spanning set and, as was the case in $\mathbf{R}^n$, if a basis for *W* consists of finitely many elements then any other basis will have the same number of elements, the **dimension** of *W*. It is important to note, however, that it will often be the case, especially in the case of function spaces, that a subspace might not be spanned by finitely many elements.

Coordinates relative to a basis

Definition: If $\mathcal{B} = \{f_1, \cdots, f_k\}$ is a basis for a finite dimensional vector space *V* (or a subspace of *V*), and if $f \in V$, then *f* can be expressed uniquely as $f = c_1 f_1 + \cdots + c_k f_k$ for scalars $\{c_1, \cdots, c_k\}$. These uniquely determined scalars are called the **coordinates of *f* relative to this basis**. If we express these coordinates as a column vector (effectively a vector in $\mathbf{R}^k$), we denote this coordinate vector by $[f]_{\mathcal{B}}$.

Example: $P_n = \{a_0 + a_1 x + \cdots + a_n x^n : a_0, a_1, \cdots, a_n \in \mathbf{R}\}$ is an $(n+1)$-dimensional vector space with basis $\{1, x, x^2, \cdots, x^n\}$. The coefficients of a polynomial are then the coordinates relative to this (standard) basis. The argument regarding linear independence is a little subtle. To show linear independence of $\{1, x, x^2, \cdots, x^n\}$, we would set $a_0 + a_1 x + \cdots + a_n x^n = 0$ and try to then show that all the coefficients must be 0. It's important to understand that this is an equation in P_n, not in the space of real numbers, so we're not seeking roots. The 0 on the right-hand-side is not the number 0, but rather the function that is identically 0 for all *x*. Thus the equation

$a_0 + a_1 x + \cdots + a_n x^n = 0$ for all x. In particular, when $x = 0$ this gives $a_0 = 0$. You can reason similarly (for example, by substituting a variety of other values for x) that all the other coefficients must also be 0.

If we look specifically at $P_2 = \left\{ a + bx + cx^2 : a,b,c \in \mathbf{R} \right\}$, the standard basis would be $\mathcal{E} = \left\{ 1, x, x^2 \right\}$ but we might also choose to express these polynomials in terms of powers of $(x-2)$, i.e. $\mathcal{B} = \left\{ 1, x-2, (x-2)^2 \right\}$. Given a polynomial of the form $p(x) = a + b(x-2) + c(x-2)^2$, we would say that $\left[p \right]_{\mathcal{B}} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ and we can multiply out to get $p(x) = a + b(x-2) + c(x-2)^2 = a + bx - 2b + cx^2 - 4cx + 4c = (a - 2b + 4c) + (b - 4c)x + cx^2$. So $\left[p \right]_{\mathcal{E}} = \begin{bmatrix} a - 2b + 4c \\ b - 4c \\ c \end{bmatrix} = \begin{bmatrix} 1 & -2 & 4 \\ 0 & 1 & -4 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \mathbf{S} \left[p \right]_{\mathcal{B}}$. We would then, quite naturally, refer to this matrix $\mathbf{S}$ as a "change of basis" matrix and the relation $\left[p \right]_{\mathcal{E}} = \mathbf{S} \left[p \right]_{\mathcal{B}}$ is reminiscent of the way we changed coordinates in $\mathbf{R}^3$.

We can also do this the other way around by starting with a polynomial of the form $p(x) = a + bx + cx^2$ with standard coordinates $\left[p \right]_{\mathcal{E}} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$, and then figure out its coordinates relative to the basis $\mathcal{B} = \left\{ 1, x-2, (x-2)^2 \right\}$. we might write:

$$\begin{aligned} p(x) &= a + bx + cx^2 = a + b(x-2+2) + c(x-2+2)^2 = a + b(x-2) + 2b + c(x-2)^2 + 4c(x-2) + 4c \\ &= (a + 2b + 4c) + (b + 4c)(x-2) + c(x-2)^2 \end{aligned}$$

So $\left[p \right]_{\mathcal{B}} = \begin{bmatrix} a + 2b + 4c \\ b + 4c \\ c \end{bmatrix} = \begin{bmatrix} 1 & 2 & 4 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \mathbf{S}^{-1} \left[p \right]_{\mathcal{E}}$. You can verify that $\mathbf{S}^{-1}\mathbf{S} = \begin{bmatrix} 1 & 2 & 4 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -2 & 4 \\ 0 & 1 & -4 \\ 0 & 0 & 1 \end{bmatrix} = \mathbf{I}$.

Definition: Given two vectors spaces V and W, a function $T : V \to W$ is called a **linear transformation** if for all elements $f_1, f_2 \in V$ and for all scalars c_1, c_2, T satisfies $T(c_1 f_1 + c_2 f_2) = c_1 T(f_1) + c_2 T(f_2)$. This can also be expressed by saying that <u>T preserves addition and scalar multiplication</u>. We call the input space V the **domain** of T and we call the output space W the **codomain**.

Definition: Suppose $T : V \to W$ is a linear transformation. We define:

$$\text{image}(T) = im(T) = \left\{ T(f) : f \in V \right\} \subseteq W \qquad \text{and} \qquad \text{kernel}(T) = \ker(T) = \left\{ f \in V : T(f) = 0 \right\} \subseteq V$$

These are both subspaces. The argument is the same as we've seen before.

Matrix of a linear transformation relative to bases for the domain and codomain

Definition: If V is a finite-dimensional vector space with basis $\mathcal{B} = \left\{ f_1, \cdots, f_n \right\}$ and W is a vector space with basis $\mathcal{C} = \left\{ g_1, \cdots, g_m \right\}$, and if $T : V \to W$ is a linear transformation, we define the **matrix of *T* relative to these bases** as: $\left[T \right]_{\mathcal{B},\mathcal{C}} = \begin{bmatrix} \uparrow & & \uparrow \\ \left[T(f_1) \right]_{\mathcal{C}} & \cdots & \left[T(f_n) \right]_{\mathcal{C}} \\ \downarrow & & \downarrow \end{bmatrix}$. This will be an $m \times n$ matrix.

This is simpler in the case where the domain and codomain coincide. In that case we might use the same basis for both and define $[T]_{\mathcal{B}} = \begin{bmatrix} \uparrow & & \uparrow \\ [T(f_1)]_{\mathcal{B}} & \cdots & [T(f_n)]_{\mathcal{B}} \\ \downarrow & & \downarrow \end{bmatrix}$. In this case, $[T]_{\mathcal{B}}$ would be an $n \times n$ (square) matrix.

Note: It is sometimes desirable to find the kernel and image of a linear transformation by choosing a basis or bases to express everything in coordinates and then finding the kernel and image of the matrix for the linear transformation relative to this basis or bases.

Definition: Given a linear transformation $T : V \to W$, we define: $\operatorname{rank}(T) = \dim(\operatorname{im} T)$ and $\operatorname{nullity}(T) = \dim(\ker T)$ when these subspaces have finite dimension.

We can also state (without proof) the corresponding fact regarding the relationship between rank and nullity.

Rank-Nullity Theorem: If $T : V \to W$ is a linear transformation and V has finite dimension, then $\operatorname{rank}(T) + \operatorname{nullity}(T) = \dim(V)$.

Definition: A linear transformation $T : V \to W$ is called an **isomorphism** if it is one-to-one and onto its codomain. That is, for every $g \in W$ there is a unique $f \in V$ such that $T(f) = g$.

For example, the correspondence between a polynomial $p(x) = a + bx + cx^2$ in P_2 and its coordinates $[p]_{\mathcal{E}} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ in $\mathbf{R}^3$ is clearly linear, one-to-one, and onto; so we can say that P_2 and $\mathbf{R}^3$ are isomorphic.

Example: In any space consisting of differentiable functions, the differentiation operator $D(f) = f'$ is a linear transformation. This follows from the Calculus facts that $D(f+g) = (f+g)' = f' + g' = D(f) + D(g)$ and $D(cf) = (cf)' = c\, f' = c\, D(f)$. If we restrict our attention to a finite dimensional space such as P_2, we note that $\left\{ \begin{array}{r} 1 \xrightarrow{D} 0 \\ x \xrightarrow{D} 1 \\ x^2 \xrightarrow{D} 2x \end{array} \right\}$, so relative to the basis $\mathcal{E} = \{1, x, x^2\}$, we have the matrix $[D]_{\mathcal{E}} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$. The kernel of this matrix is $\operatorname{span}\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\} \leftrightarrow \{\text{constant functions}\}$, and the image is $\operatorname{span}\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\} \leftrightarrow \{a + bx : a, b \in \mathbf{R}\} = P_1$. Said differently, the set of all such functions whose derivative is 0 consists of the constant functions, and if you differentiate a quadratic function you'll get a polynomial of degree 1. It should be clear that this linear transformation is definitely NOT an isomorphism.

Example: Let $V = \mathbf{R}^{2\times 2}$ and let $T(\mathbf{A}) = \mathbf{A}\begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}$ for any 2×2 matrix **A**. It's easy to see that this is linear. Find bases for its kernel and its image and the matrix of T relative to the basis $\mathcal{E} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$.
(It's easy to see that $\mathcal{E} = \{E_1, E_2, E_3, E_4\}$ is a basis.) Is T an isomorphism?

Solution: Let $\mathbf{A}=\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and calculate $T(\mathbf{A})=\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}=\begin{bmatrix} a+b & 2(a+b) \\ c+d & 2(c+d) \end{bmatrix}$. If $T(\mathbf{A})=\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, then $a+b=0$ and $c+d=0$, so $b=-a$ and $d=-c$. So if $\mathbf{A}\in\ker(T)$, then $\mathbf{A}=\begin{bmatrix} a & -a \\ c & -c \end{bmatrix}=a\begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}+c\begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix}$.

So $\left\{\begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix},\begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix}\right\}$ is a basis for $\ker(T)$ since these are clearly linearly independent.

Therefore $\text{nullity}(T)=2$.

Anything in the image is of the form $T(\mathbf{A})=\begin{bmatrix} a+b & 2(a+b) \\ c+d & 2(c+d) \end{bmatrix}=(a+b)\begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}+(c+d)\begin{bmatrix} 0 & 0 \\ 1 & 2 \end{bmatrix}$, so $\left\{\begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix},\begin{bmatrix} 0 & 0 \\ 1 & 2 \end{bmatrix}\right\}$ is a basis for $\text{im}(T)$ and $\text{rank}(T)=2$. Note that $\text{rank}(T)+\text{nullity}(T)=4=\dim(V)$.

This is NOT an isomorphism (since $\ker(T)\neq\{0\}$). There are a couple of good ways to find the matrix of *T*.

By the columns: We could calculate what the linear transformation does to each of the basis elements in order to determine the columns of the matrix. This is a bit tedious, but it goes like this:

$$\left\{\begin{aligned} T(E_1)&=\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}=\begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}=E_1+2E_2 \\ T(E_2)&=\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}=\begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}=E_1+2E_2 \\ T(E_3)&=\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}=\begin{bmatrix} 0 & 0 \\ 1 & 2 \end{bmatrix}=E_3+2E_4 \\ T(E_4)&=\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}=\begin{bmatrix} 0 & 0 \\ 1 & 2 \end{bmatrix}=E_3+2E_4 \end{aligned}\right\} \Rightarrow [T]_{\mathcal{E}}=\begin{bmatrix} 1 & 1 & 0 & 0 \\ 2 & 2 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 2 & 2 \end{bmatrix}$$

By the rows: If the input element is $\mathbf{A}=\begin{bmatrix} a & b \\ c & d \end{bmatrix}$, its coordinates are $[\mathbf{A}]_{\mathcal{E}}=\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$. The output element is

$T(\mathbf{A})=\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}=\begin{bmatrix} a+b & 2(a+b) \\ c+d & 2(c+d) \end{bmatrix}$, and its coordinates are $[T(\mathbf{A})]_{\mathcal{E}}=\begin{bmatrix} a+b \\ 2a+2b \\ c+d \\ 2c+2d \end{bmatrix}=\begin{bmatrix} 1 & 1 & 0 & 0 \\ 2 & 2 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 2 & 2 \end{bmatrix}\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$.

Since $[T(\mathbf{A})]_{\mathcal{E}}=\begin{bmatrix} 1 & 1 & 0 & 0 \\ 2 & 2 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 2 & 2 \end{bmatrix}\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}=[T]_{\mathcal{E}}[\mathbf{A}]_{\mathcal{E}}$, it follows that $[T]_{\mathcal{E}}=\begin{bmatrix} 1 & 1 & 0 & 0 \\ 2 & 2 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 2 & 2 \end{bmatrix}$.

It is worth noting that the rank of this matrix is 2, and we could alternatively have used this matrix to determine the kernel and image of the linear transformation that it represents.

Note: A linear transformation $T:V\to W$ where V and W are finite-dimensional will be an **isomorphism** if and only if its matrix relative to suitable bases for V and W is an invertible matrix.

Notes by Robert Winters