4 Lecture 4

4.1 Affinity

We are now going to move on to a discussion of affinity. This is going to be closely related to the concept of linearity, but is more general. What I mean by this can be demonstrated in the following example. Consider the set

$$\{x \in \mathbb{R}^n | x_1 + x_2 = 1\}$$

Is this a linear subspace? No - because it doesn't contain the zero element. However it *looks* like a lot like a linear subspace. In fact, it is the linear subspace $\{x \in \mathbb{R}^n | x_1 + x_2 = 0\}$ translated (or shifted in a parallel way) by adding a vector (say $(0, 1)$) to each of its elements. This is how we are going to define a concept called an **affine manifold.**

In order to define this concept formally, we are going to first need to define the notation of the addition of subsets

Definition 17 *For any two subsets S, T of a linear space V, we define $S + T$ in the following way:*

$$S + T = \{s + t | s \in S, \ t \in T\}$$

if $S = \{s\}$, we use the shorthand $s + T$ rather than $\{s\} + T$

We can now define an affine manifold, and the related concept of a **hyperplane**

Definition 18 *A subset S of a linear space V is an affine manifold of V if $S = Z + x^*$ for some linear subspace Z of V and some $x^* \in V$. If Z is a $\supseteq -$maximal proper subspace of V then we call S a hyperplane*

So what I mean by affinity being more general than linearity is something like the following:

$$\text{affinity=linearity+translation}$$

So what do manifolds and hyperplanes look like in $\mathbb{R}^2$? Well, firstly, remember that there are two classes of linear subspace in $\mathbb{R}^2$: 0, and any line through the origin. Thus, we have two classes of affine manifold: any point, and any line. The latter is clearly the class of hyperplanes in $\mathbb{R}^2$. Similarly, in $\mathbb{R}^3$ we have three classes of affine manifolds: points, lines and planes, with planes being the hyperplanes.

This means that in $\mathbb{R}^2$, any hyperplane can be defined by a function of the form $n_1x_1+n_2x_2 = b$ (as any line can be written in this form). In other words, they are described by the set

$$\{x| < n.x >= b\}$$

for some $n \in \mathbb{R}^2$ and $b \in \mathbb{R}$. In fact, it turns out that this is a general characterization of hyperplanes in $\mathbb{R}^n$

Proposition 2 *A set $S \in \mathbb{R}^n$ is a hyperplane if and only if $S = \{x \in \mathbb{R}^n| < n.x >= b\}$ for some $n \in \mathbb{R}^n/\emptyset$, $b \in \mathbb{R}$.*

Proof. *Beyond the scope of the course. See Ok F.2.4. if you are keen.* ■

In $\mathbb{R}^2$ this proposition formalizes the idea that a hyperplane is a line, and so defined by its slope (n) and intercept (b).

This definition allows us to think of dividing $\mathbb{R}^n$ into two sets that we call **half-spaces:**

Definition 19 *Any hyperplane in $\mathbb{R}^n$ generates two closed half-spaces*

$$\begin{aligned} H_{\geq}(n, \beta) &= \{x| < n, x > \geq b\} \\ H_{\leq}(n, \beta) &= \{x| < n, x > \leq b\} \end{aligned}$$

Open half-spaces are defined analogously.

Hyperplanes and half spaces are going to turn out to be very important, because it turns out that hyperplanes can be used to separate and support convex sets. This important property is going to be the basis of the optimization results that we will prove later on.

What are some of the properties of affine manifolds and hyperplanes? Well first of all, note that all linear subspaces are affine manifolds (yes?), though the converse is clearly not true. However, the following is true

Remark 5 *An affine manifold S of a linear space V is a linear subspace of V if and only if $\emptyset \in S$*

Proof. *If S is a linear space, then clearly $\emptyset$ must be in S. To go the other way, start with the fact that $S = Z + x^*$ for some linear subspace Z of V and some $x^* \in V$. If $\emptyset \in S$ then it must be the case that $\emptyset = y + x^*$ for some $y \in Z$, and so $-x^* \in Z$. As Z is closed under addition, this implies that $x^* \in Z$, and so (again by the fact that Z is closed under addition) $S = Z + x^* = Z$* ■

If we translate an affine manifold S by subtracting any $x \in S$ we will get a linear subspace. Moreover, we will get the same linear subspace regardless of which element in S we use for the translation. Thus, any affine manifold is closely linked to a particular subspace.

Remark 6 *Let S be an affine manifold of a linear space V and $x, y \in S$. Then $S - y = S - x$ and both are linear subspaces.*

Proof. *To show that $S - x$ is a linear space, note first that it is an affine manifold, as for some subspace Z and $x^* \in V .S = Z + x^*$, then*

$$
\begin{aligned}
& S - x \\
= \; & \{s + (-x) | s \in S\} \\
= \; & \{s + (-x) | s \in \{z + x^* | z \in Z\}\} \\
= \; & \{z + (x^* + (-x)) | z \in Z\}
\end{aligned}
$$

Next note that it contains $\emptyset$ as $x \in S$ so $\emptyset = (x + (-x)) \in S - x$, and so by remark 5, $S - x$ is a linear subspace

To show that any two elements produce the same subspace, take some element $z \in S - x$ and we will show that it is also in $S - y$. First note that $y - x \in S - x$, and as $S - x$ is a linear subspace, this means that $z + y - x \in S - x$. This means that there is some $s \in S$ such that

$$
\begin{aligned}
z + y - x &= s - x \\
z + y &= s \\
z &= s - y
\end{aligned}
$$

■

We can extend the concept of a dimension to affine manifolds by using the dimension of the linear subspace that generated the affine manifold.

Definition 20 *Let V be a linear space and S be an affine manifold in V then we define the dimension of S as*

$$\dim(S) = \dim(S - x) \ x \in S$$

Note that this definition is independent of the choice of x by remark 6. Note also that this implies that, for a linear space V of dimension n, the dimension of any hyperplane is $n - 1$.

One extremely nice property of affine manifolds is the following.

Proposition 3 *Let V be a linear space, and $\emptyset \neq S \subset V$. Then S is an affine manifold if and only if*

$$\lambda x + (1 - \lambda) y \in S \text{ for all } x, y \in S, \ \lambda \in \mathbb{R}$$

Proof. *Exercise* ■

This is not the same as saying that affine manifolds are closed under addition and scalar multiplication (the above results tell us that, right?), but it has a similar flavor.

5 Lecture 5

5.1 Linear Operators and Linear Functionals

We are now going to move on to talk about an extremely important class of operators. These are **linear operators** (or **linear transformation**) - operators that map one linear space into another in a way that preserves linearity

Definition 21 *Let V and W be two linear spaces. A function $L : V \to W$ is called a linear operator (or linear transformation) if*

$$L(\alpha x + x') = \alpha L(x) + L(x') \ \forall \ x, x' \in X, \ \alpha \in \mathbb{R}$$

A real valued linear operation is called a linear functional

Note that the addition and scalar multiplication on the left hand side of the above equation are not necessarily the same as those on the right.

Definition 22 *$\mathcal{L}(V, W)$ is the set of all linear functionals that map V to W*

As usual, we can gain intuition by thinking about what linear operators look like in $\mathbb{R}^n$. It turns out that any linear operator that maps $\mathbb{R}^n$ to $\mathbb{R}^m$ can be written as a an $m \times n$ matrix. (This, finally, may answer the question as to how the 'linear algebra' that we have been studying relates to the 'linear algebra' that you have done in high school or college

Remark 7 *The function $L : \mathbb{R}^n \to \mathbb{R}^m$ is a linear operator if and only if $\exists \ A \in \mathbb{R}^{m \times n}$ such that $L(x) = Ax \ \forall \ x \in \mathbb{R}^n$, defined as*

$$Ax = \left(\sum_{j=1}^{n} a_{1j} x_j, ..., \sum_{j=1}^{n} a_{mj} x_j \right)$$

Proof. *The fact that such a function is a linear operator is easy to check. To show that any $L \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ can be written this way, take such a function, and define $L_i : \mathbb{R}^n \to \mathbb{R}$ as the function that maps each $x \in \mathbb{R}^n$ to the ith component of $L(x)$. Check that you agree that L_i has* *to be a linear functional on $\mathbb{R}^n$. Now, let $\{e^1, .., e^n\}$ be the standard basis for $\mathbb{R}^n$. Then, for any $x \in \mathbb{R}^n$*

$$x = \sum_{j=1}^{n} x_j e^j$$

meaning that

$$\begin{aligned} L_i(x) &= L_i\left(\sum_{j=1}^{n} x_j e^j\right) \\ &= \sum_{j=1}^{n} L_i(e^j) x_j \end{aligned}$$

Thus, we can obtain the desired result by setting $A := [L_i(e^j)]_{m \times n}$ ■

Note that this implies we can write any linear functional on $\mathbb{R}^n$ as $x =< v.y >$ for some $v \in \mathbb{R}^n$

Of course, we have other examples of linear operators

Example 4 *Let $C^1[0,1]$ be the linear space of all continuously differentiable functions on $[0,1]$ (with addition and scalar multiplication defined in the standard way), and define $D : C^1[0,1] \to C[0,1]$ as $D(f) = f'$. Then $D(\alpha f + g) = af' + g'$ (as, if $h(x) = \alpha f(x) + g(x)$ then $\frac{\partial h(x)}{\partial x} = \alpha \frac{\partial f(x)}{\partial x} + \frac{\partial g(x)}{\partial x}$), and so D is a linear operator.*

5.2 Null Space

Definition 23 *The null space of a linear operator L that maps V to W is the subset of V that maps to the zero element in W*

$$null(L) := \{v \in V | L(v) = \emptyset\}$$

Some properties of linear operators and the null-space which are going to come in useful are as follows

Remark 8 *The following hold true for any arbitrary linear spaces V and W and any $L \in \mathcal{L}(V, W)$*

1. $\emptyset \in null(L)$

2. *L is injective if and only if $null(L) = \{\emptyset\}$*

3. *$null(L)$ is a linear subspace of V*

4. *$L(V)$ is a linear subspace of W*

Proof. *The proofs for these claims are as follows*

1. *By the definition of a linear operator, $L(v) = L(v + \emptyset) = L(v) + L(\emptyset)$. for any $v \in V$. Add $w + (-L(v))$ to either side of this equation gives $w = w + L(\emptyset)$ for any arbitrary $w \in V$*

2. *Say $\emptyset \neq v \in null(L)$. Then $L(s + v) = L(s) + L(v) = L(s)$ for any $s \in V$, so the function is not injective. Now, say that $null(L) = \{\emptyset\}$. Then, for $s, v \in V$*

$$\begin{aligned} L(s) &= L(v) \\ &\Rightarrow L(s) - L(v) = \emptyset \\ &\Rightarrow L(s - v) = \emptyset \\ &\Rightarrow s = v \end{aligned}$$

3. *Here we have to show that $null(L)$ is closed under addition and scalar multiplication (we know it is non-empty by point 1 above). Say that $s, v \in null(L)$ Then, for any $\alpha \in \mathbb{R}$*

$$\begin{aligned} & L(\alpha s + v) \\ =\ & \alpha L(s) + L(v) \\ =\ & \emptyset \end{aligned}$$

   so $\alpha s + v \in null(L)$

4. *Again, we have to show that $L(V)$ is closed under addition and scalar multiplication. Say that $t, w \in L(V)$ Then there exists $s, v \in V$ such that $L(s) = t$ and $L(v) = W$. Then, for any $\alpha \in \mathbb{R}$*

$$\begin{aligned} & L(\alpha s + v) \\ =\ & \alpha L(s) + L(v) \\ =\ & \alpha t + w \end{aligned}$$

   so $\alpha t + w \in L(V)$

$\blacksquare$

5.3 The Fundamental Theorem of Linear Algebra

We are next going to prove a result so important that it goes by the name of the Fundamental Theorem of Linear Algebra (actually this is just one version of it - we will come back to it - and explain why it is so important - later on).

Theorem 13 *Given any two linear spaces V and W*

$$\begin{aligned} \dim(null(L)) + \dim(L(V)) &= \dim(V) \\ \forall\ L &\in \mathcal{L}(V, W) \end{aligned}$$

Proof. *This proof is going to rely on the result that any linearly independent subset of a linear space can be extended to a basis for that space. You have proved this for the finite dimension case for homework, but it is in fact also true for the infinite dimension case. Thus, let A be a basis for $null(L)$. Now extend A to be a basis for V, B. I claim that $L(B/A)$ is a basis for $L(V)$. First, we need to show that $L(B/A)$ spans $L(V)$. To see this, say that $w \in L(V)$. Then there must a $v \in V$ such that $w = L(v)$. As B is a basis for V, we know that there is some finite collection $x_1, ...x_n \in B$ and $\alpha_1, ..., \alpha_n \in \mathbb{R}$ such that*

$$\sum_{i=1}^{n} \alpha_i x_i = v$$

WLOG, assume $x_i, .., x_k \in A$ and $x_{k+1}, .., x_n \in B/A$. Then

$$\begin{aligned} w &= L(v) \\ &= L\left(\sum_{i=1}^{n} \alpha_i x_i\right) \\ &= \sum_{i=1}^{n} \alpha_i L(x_i) \\ &= \sum_{i=1}^{k} \alpha_i L(x_i) + \sum_{i=k+1}^{n} \alpha_i L(x_i) \end{aligned}$$

But by definition of A (and $null(L)$), $L(x_i) = \emptyset\ \forall\ i \leq k$, so we have

$$w = \sum_{i=k+1}^{n} \alpha_i L(x_i)$$

and so $L(B/A)$ spans $L(X)$. All that remains to be shown is that $L(B/A)$ is a linearly independent set. This is left as an exercise

Thus, we have

$$\begin{aligned} |B| &= |A| + |A/B| \\ &\Rightarrow \dim(V) = \dim(null(L)) + \dim(L(V)) \end{aligned}$$

■

6 Lecture 6

6.1 Linear Algebra on $\mathbb{R}^n$

One thing that might be puzzling you (it certainly puzzled me for a long time) is what the concepts of linear algebra that we have discussed here have to do with the concepts that are taught in standard undergraduate classes, which are generally to do with matrices, and finding solutions to systems of equations. We have already had one answer: that linear operators mapping $\mathbb{R}^m$ to $\mathbb{R}^n$ can be thought of as matrices. However, there is a deeper link as we will see.

This section is going to deal with $n \times m$ matrices

$$A = \begin{matrix} a_{1,1} & ... & a_{1,m} \\ \vdots & & \vdots \\ a_{n,1} & ... & a_{n,m} \end{matrix}$$

Remember, we can think of A as a linear function mapping $\mathbb{R}^m$ to $\mathbb{R}^n$, which we write as Ax.

We can think of this matrix as consisting of a set of n vectors in $\mathbb{R}^m$ which we denote as $\{a_1, ..a_n\}$. i.e. we can think of

$$A = \left\{ \begin{matrix} a_1 \\ \vdots \\ a_n \end{matrix} \right\}$$

Where $a_i = \{a_{i,1}, ..., a_{i,m}\}$. By doing so, we can define the **row space** of A as the subspace of $\mathbb{R}^m$ spanned by $\{a_1, ..a_n\}$.

$$\text{row}(A) = span\{a_1, ..a_n\}$$

Now, remember that an **elementary row operation** is one of the following

1. interchange two rows of a matrix
2. change a row by adding to it a multiple of another row
3. multiplying each element in a row by a non-zero number

Also remember that a matrix A_r is in **row echelon form of** A if A_r can be obtained from

A using elementary row operations, and every row of A_r has more leading zeros than the one proceeding it.

The following facts should now be self evident (if they are not, make sure you prove them)

Remark 9 *Let A_r be a row echelon form of A then*

1. $\text{row}(A_r) = \text{row}(A)$
2. *The non-zero row vectors of A_r are a basis for* $\text{row}(A)$
3. $\dim(\text{row}(A))$ *is the rank of A*

Next, we move on to the **column space** attached to A. This is entirely analogous to the row space: the columns $\{b_1, .., b_m\}$ of A are a collection of vectors in $\mathbb{R}^n$. i.e. we can rewrite A as

$$A = \begin{Bmatrix} b_1 & \cdots & b_m \end{Bmatrix}$$

where $b_i = \begin{Bmatrix} b_{1,i} \\ \vdots \\ b_{n,i} \end{Bmatrix}$. So we can define the column space of A as the subspace of $\mathbb{R}^n$ spanned by $\{b_1, .., b_m\}$:

$$\text{col}(A) = span(b_1, .., b_m)$$

Now, remember that a **pivot** is the first non-zero entry of any row in the row echelon form of a matrix. We can define a column of A as **basic** if the corresponding column of a row echelon form of the matrix contains a pivot. It is less obvious but true (you can read it in Simon and Blume pp 775) that the basic columns of A form a basis of $\text{col}(A)$. This gives the following result:

Remark 10 $\dim(\text{row}(A)) = \dim(\text{col}(A)) = rank(A)$

The column space of a matrix is important for solving systems of equations. Think of the system represented by

$$Ax = c$$

or in other words

$$\begin{aligned}
a_{11}x_1 + ... + a_{1m}x_m &= c_1 \\
a_{21}x_1 + ... + a_{2m}x_m &= c_2 \\
&\vdots \\
a_{n1}x_1 + ... + a_{nm}x_m &= c_n
\end{aligned}$$

or, but another way

$$b_1x_1 + ... + b_mx_m = c$$

In other words, denoting Ax as $L_A(x)$, we can see that $L_A(\mathbb{R}^m) = \mathrm{col}(A)$

This gives us the following theorem

Theorem 14 *Let A be an $n \times m$ matrix*

1. *A system of equations $Ax = c$ has a solution if and only if $c \in \mathrm{col}(A)$*
2. *A system of equations only has a solution for any c if $\mathrm{col}(A)$ spans $\mathbb{R}^n$, and so $\dim(\mathrm{col}(A)) = n = rank(A)$*
3. *If this is the case, then $n = rank(A) \leq m$*
4. *The space of solutions to any system of equations is an affine manifold*

1. The dimention of that set of solutions is equal to $m - rank(A)$

Finally, note that, for the linear function represented by A, the nullspace is the set of solutions $\{x \in \mathbb{R}^m | Ax = 0\}$

6.2 Isomorphisms

We are now going to move on to talk about isomorphisms, a class of linear functionals that are important in determining the relationship between two linear spaces

Definition 24 *For two linear spaces V and W, a function $L \in \mathcal{L}(V,W)$ is an isomorphism if it is bijective - i.e. it is one-to one and onto - i.e. for every element $w \in W$ there is one and only one $v \in V$ such that $L(v) = w$. If there exists an isomorphism between V and W we call the two spaces isomorphic.*

It should be clear that there cannot be an isomorphism between a finite dimension and an infinite dimensional space, so here we are going to focus on finite dimensional space. Here, we get a nice result

Proposition 4 *Two finite dimensional spaces are isomorphic if and only if they have the same dimension*

Proof. *First we are going to prove that the dimension of any two isomorphic linear spaces X and Y, must be the same. The case where $\dim(X) = 0$ is trivial, so let $dim(X) \in \mathbb{N}$. Let $\{x_1, ..x_n\}$ be a basis for X and let $L \in \mathcal{L}(X,Y)$ be an isomorphism We are going to show that $\{L(x_1), .., L(x_n)\}$ is a basis for Y. First, we show that they are linearly independent. Assume not, then $\exists$ $\{\alpha_1, ...\alpha_n\} \in \mathbb{R}^n/\emptyset$ such that*

$$\sum_{i=1}^{n} \alpha_i L(x_i) = \emptyset$$

implying

$$L\left(\sum_{i=1}^{n} \alpha_i x_i\right) = \emptyset$$

But, as for injective functions $null(L) = \emptyset$ (yes?), this implies

$$\sum_{i=1}^{n} \alpha_i x_i = \emptyset$$

a contradiction, as $\{x_1, .., x_n\}$ was a basis for X

Furthermore, we need to show that $\{L(x_1), .., L(x_n)\}$ spans Y. But this is clear, as we know that for every $y \in Y$ there exists an $x \in X$ such that $y = L(x)$. As $\{x_1, ..x_n\}$ is a basis for X then *we can find* $\{\alpha_1, ...\alpha_n\} \in \mathbb{R}^n$ *such that* $x = \sum_{i=1}^n \alpha_i x_i$*. Thus*

$$\begin{aligned} y &= L(x) \\ &= L\left(\sum_{i=1}^n \alpha_i x_i\right) \\ &= \sum_{i=1}^n \alpha_i L(x_i) \end{aligned}$$

Showing that $\{L(x_1), .., L(x_n)\}$ *is indeed a basis for* Y

The proof of the other direction we will omit for now $\blacksquare$

This is a phenomenally useful result. One main reason is an obvious corollary is that any finite dimensional space is isomorphic to $\mathbb{R}^n$ for some n. Thus, for many purposes, we can show things to hold true in $\mathbb{R}^n$, then use the isomorphic result to show that they hold true in all finite dimensional linear spaces. Hurrah!