Ordinary Differential Equations (noc 24 ma 78)

Dr Kaushik Bal
Department of Mathematics and Statistics
Indian Institute of Technology, Kanpur

Week-01

Lecture-02

Linear Transformation

Welcome students to this video and in this video we are going to talk about linear transformations.

So, what is the linear transformation?

Essentially what we are doing is we are starting out with two vector spaces.

So, let x and y are two vector spaces.

And we are going to define a map and define a map okay and let us just call it a map, A, okay A:X-Y so this is called a linear map this is called a linear map if, A$(x \quad + x \quad) = Ax \quad + Ax$

So, this should hold for all x and x in X. Of course, and also we omit this A(Cx)=C(Ax)

Again, this should hold for all x in X and C is in ℝ, whatever the field you are choosing.

So, for us, this is ℝ.

here so you see we often write so essentially what happens is see we in this course we are always going to write Ax okay so in this is essentially nothing but a acting at the element Ax≅ $A(x)$ okay so this is a small notation notation let me just put it this way okay we are going to write this a x okay of course you can see that if x and x and x is what you have is A(x) so let's say note

that if x and x is , then you can see that A*=*A, acting at .

So, which will give you $A_x = \quad y$

Now, so once that is clear, you see what happens is basically why we are interested in linear transformation because we are going to work with a lot of matrices.

And in this case, what is happening is this a linear transformation is basically nothing but matrices.

So, how do you put it?

You see what happens is a linear transformation.

So, let me put it as a remark, let us say.

that at a linear transformation, a linear transformation, transformation A from x to y, okay, is completely determined.

So, you can actually exactly know what it does, okay, by completely determined, determined

by its action on the basis, by its action on any basis, right, any basis.

So, what do I mean by this?

What I mean is, let us say that you start with a basis, let us say basis: $\{ x \;, \ldots, x_n\}$ be the basis of be the basis of X, okay.

Then you see any x in X can be written as, X=$\sum_{i=}^{n} c_i\, x_i$

We can write it like this.

And if we look at the action of A, Ax=A($\sum_{i=}^{n} c_i\, x_i$)

and since this is linear from the second property combining the first and the second property we know that you can actually write it as, Ax=A($\sum_{i=}^{n} c_i\, x_i$)= $\sum_{i=}^{n} c_i\, Ax_i$

We can write it like this.

So you see what is happening is this that the linearity of A will allow us to compute A of x from the vectors Ax , Ax ,…., Ax_n.

So this is what we want to say that it is completely determined by the action of its basis.

And you see another small remark A, B, let me put it this way, small remark B, that the linear transformations, linear transformations from x to X, let us say, transformations from x to X, transformations.

from, from x to X, x to X, we will call that, are called, are called linear operators on x, okay, are called linear operators, operators on X. I think, I hope this is clear, operators on X. This is just a name, nothing special, okay.

Preliminaries :- (Linear Transformation)

Let X and Y are two vector space and define $A : X \to Y$ is linear map if

$$A(x_1 + x_2) = Ax_1 + Ax_2 \quad \forall\, x_1, x_2 \in X$$

and, $A(cx) = cA(x) \quad \forall\, x \in X$ and $c \in \mathbb{R}$.

Notation $Ax \simeq A(x)$

Note, $A0_x = 0_Y$.

Remark :- (a) A linear transformation A from X to Y is completely determined by its it action on any basis: $\{x_1, \ldots x_n\}$ be the basis of X, $x \in X \Rightarrow x = \sum_{i=1}^{n} c_i x_i$

and, $Ax = A\left(\sum_{i=1}^{n} c_i x_i\right) = \sum_{i=1}^{n} c_i A x_i$

(b) Linear transformations from X to X are called linear operators on X.

So, you see, what is happening is this, okay,

So now what is happening with this?

So basically we want to find out some properties.

What are the properties that can be satisfied by a linear operator?

So let us look at a small theorem.

So let us say that a linear operator, let us say you are given a linear operator on a finite dimensional vector space X, right?

So a linear operator.

a linear operator, let us say, and let us just call it a L now, operator L. Operator, you remember, it is from x to X. A linear operator, let us say A:x -X is -.

So, when can we say it is -?

This is a if and only if condition, if and only if the range of A, if the range of A, of A,

is X. So, basically it is all of X, okay.

So, if the range is all of X, then basically you have that the linear operator a is -, right.

So, what is the proof of this?

Let us look at the proof of this thing.

Very simple proof, but let us look at the proof.

So, let us say we start with a basis, right, of $\{x$,…., $x_n\}$ be the basis of X, be the basis of X, basis of X,

Okay, now the linearity since a is linear, right?

Okay, what does this mean?

It means that the range of a, the range of a, let us just write it like this.

This is the $\mathcal{R}$(A)=Range of A. Okay, how do you, how can you write the range of A?

So, this is nothing but the span of this element, right?

You see, this can be written as $\mathcal{R}$(A)={Ax ,...,Ax_n} $=$ the span of{Ax $,\ldots,$ Ax_n}, right?

span of this.

This is called the span.

This or we write it as the span of these vectors, Ax ,...,Ax_n.

You remember because you see from the last slide, you can see that any Ax can be written like this.

So basically, what happens is if you start with the element of range of A, which is Ax, then you can write it as a linear combination of Ax_i.

So you can write it like this.

Now, you see that, when is, you can easily see that $\mathcal{R}$(A)=x

If and only, $\{\mathrm{A}x\ ,\ldots,\mathrm{A}x_n\}$, this, this set, okay, this is independent.

Clear?

this you can easily see that $\mathcal{R}(A)$has to be x if and only if and only this is independent okay so basically we have to show that A is - okay so basically oh sorry sorry i am writing a here so let's write it as a because i i started our linear operator as a right so we should write it as a okay so

We should do it like this.

So, $\mathcal{R}(A)$ is x if and only if this is independent.

So, we have to show that this can happen if and only if A is -.

So, how do you show that?

So, suppose, now let us say, suppose A is -.

If that is the case and let us say that $\sum_{i=}^{n} c_i\, Ax_i =$. So, basically it is independent, I am assuming this thing.

Then what is happening is that we'll imply A($\sum_{i=}^{n} c_i\, x_i) =$, right, because of the linearity and then one can write down $\sum_{i=}^{n} c_i\, x_i =$, right.

And then, since a, so a is -, right?

So, basically, you see, that will actually imply that $c \quad = c \quad = c \quad = \cdots = c_n =$.

that should be equals to , right?

Why?

Because since a is - and A()=, so it will give you that $\sum_{i=}^{n} c_i\, x_i$=, okay?

And then you always have that c_i= because $x \quad , x \quad ,\ldots, x_n$ are the basis of X, okay?

So, that is there.

So, that will imply that the set {A$x \quad ,\ldots,$ Ax_n}, this set is

is linearly independent, right, is linearly independent, linearly independent, okay.

Now, the thing is, let us look at the converse, okay.

So, let us say that the set {A$x \quad ,\ldots, Ax_n$}

let us say this set is independent, linearly independent, that is linearly independent, linearly independent, dependent, and A($\sum_{i=}^{n} c_i\, x_i) =$.

So, that will actually imply that $\sum_{i=}^{n} c_i\, Ax_i =$.

It is very easy.

Of course, because of the linearity, we can write it.

And that will imply that c_i= for all i, right?

i=,…..,n. This is true because a of x , x ,…., x_n, I have assumed that this is going to be linearly independent, right?

So, this is true.

So, this will imply that A(x)= only if x=, okay?

So, you see, now let us say, if now Ax=Ay, okay, that will imply A(x-y)= , linearity and that will imply x=y, okay.

And that will imply that A is -.

So basically you see you started with the linearly independent set, you showed it is - and the converse also.

Theorem :- A linear operator $A : X \to X$ is 1-1 iff the range of A is X.

Proof :- $\{x_1, \dots x_n\}$ be the basis of X.

$\mathcal{R}(A)$ = Range of A = $\overline{\{Ax_1 \dots, Ax_n\}}$ = span $\{Ax_1 \dots Ax_n\}$

$\mathcal{R}(A) = X$ iff $\{Ax_1, \dots, Ax_n\}$ is independent.

Suppose A is 1-1 and $\sum_{i=1}^{n} c_i Ax_i = 0 \Rightarrow A(\sum c_i x_i) = 0 \Rightarrow \sum_{i=1}^{n} c_i x_i = 0$

$\Rightarrow c_1 = c_2 = \dots = c_n = 0$.

$\Rightarrow \{Ax_1 \dots Ax_n\}$ is linearly independent.

Converse, $\{Ax_1 \dots Ax_n\}$ is linearly independent and $A(\sum c_i x_i) = 0$

$\Rightarrow \sum c_i Ax_i = 0 \Rightarrow c_i = 0$ for $i = 1, 2, \dots, n$.

$\Rightarrow Ax = 0$ only if $x = 0$.

If, $Ax = Ay \Rightarrow A(x-y) = 0 \Rightarrow x = y \Rightarrow A$ is 1-1

So we can actually say that the linear operator A is - if and only if the range of A is going to be whole of X. Now we need some definitions.

Let us do some definitions here.

definitions.

So, the first definition which we are going to do is this.

You see, we are going to write down L(X,Y)=This is the set of all linear transformations, set of all linear transformations, linear transformations, transformations or linear maps, whatever you want to call it, linear transformations from X to Y and from X to Y and of course, we are assuming X and Y are

vector space X to Y, clear?

And you see, if there is a linear operator on X, it means that we just write L(X,Y). So, and L(X,X)=L(X). Clear?

Okay, so this is just a notation which we are going to use, right?

Okay, now the thing is this, what we are going to do is we are going to define a very important notion.

So, we are going to define the norm, right?

So, let us say for A$\in$ $L(\mathbb{R}^n, \mathbb{R}^m)$ clear?

What happens is we define the norm, the norm.

||A|| of A, okay and how do we define it we define it to be the sup oh sorry we define the norm as ||A||:=sup$\{|\text{Ax}|_m\text{: } |\text{x}|_n \leq\}$

Okay so basically we are looking at where A is acting so basically we are choosing a from this unit ball less than equal one okay and we are looking at how a is acting on that in every element of that ball okay

And we take the supremum of such a set.

So how do you know such a supremum exists?

You see A is a linear map from $\mathbb{R}^n$ to $\mathbb{R}^m$.

We know the linear map.

So we know that linear maps are going to be continuous on finite dimensional spaces.

And since it is a continuous function and this set, this set is

compact set we know that the maximum or supremum whatever you call it in this case maximum but let us just stick with supremum here but the supremum exists okay and that we are going to call it as a norm of A okay right and we can also observe okay observe this observe that,
$|\text{Ax}|_m \leq \left||A|\right| * |\text{x}|_n$

I hope you can see why this is true.

Okay.

So please check this part.

Check this part.

Just keep this in mind that any element y. Okay.

Let us say you are starting with any element in Y in $\mathbb{R}^n$.

Then $\frac{Y}{|Y|_n}$ =. Okay?

Now you can use this fact to prove something like this.

So this is very easy to show.

So please do that.

Okay.

Right.

So, now there are two theorems which I need to know proof theorems.

So, the first theorem is this.

If a linear transformation from $A \in L(\mathbb{R}^n, \mathbb{R}^m)$, then $||A|| << \infty$, is finite, okay.

And, and A

is uniformly continuous, okay, uniformly continuous, continuous.

Continuity part is very easy, okay, that is then of course true, but it is not only continuous, it is more than that, it is uniformly continuous.

Now, B, okay, so this is a small, another, let me put a small note here, check that $A \in L(\mathbb{R}^n,\mathbb{R}^m)$

is continuous.

So maybe we can, of course, there is nothing special to check because we are going to do it here itself.

But so basically, it is not only continuous, this is uniformly continuous.

So, and of course, what we have is this.

So if, let us say, if $A, B \in L(\mathbb{R}^n,\mathbb{R}^m)$, we are choosing two elements from $\mathbb{R}^n$to $\mathbb{R}^m$, okay?

So two linearly, sorry, linear transformation from $\mathbb{R}^n$to $\mathbb{R}^m$, and you are choosing a $C \in \mathbb{R}$

scalar which is $\mathbb{R}$ in our case then you can actually show that, $||A+B|| \leq ||A|| + ||B||$, so it satisfies the triangle inequality, okay and $|cA|=|c|\ ||A||$

Clear?

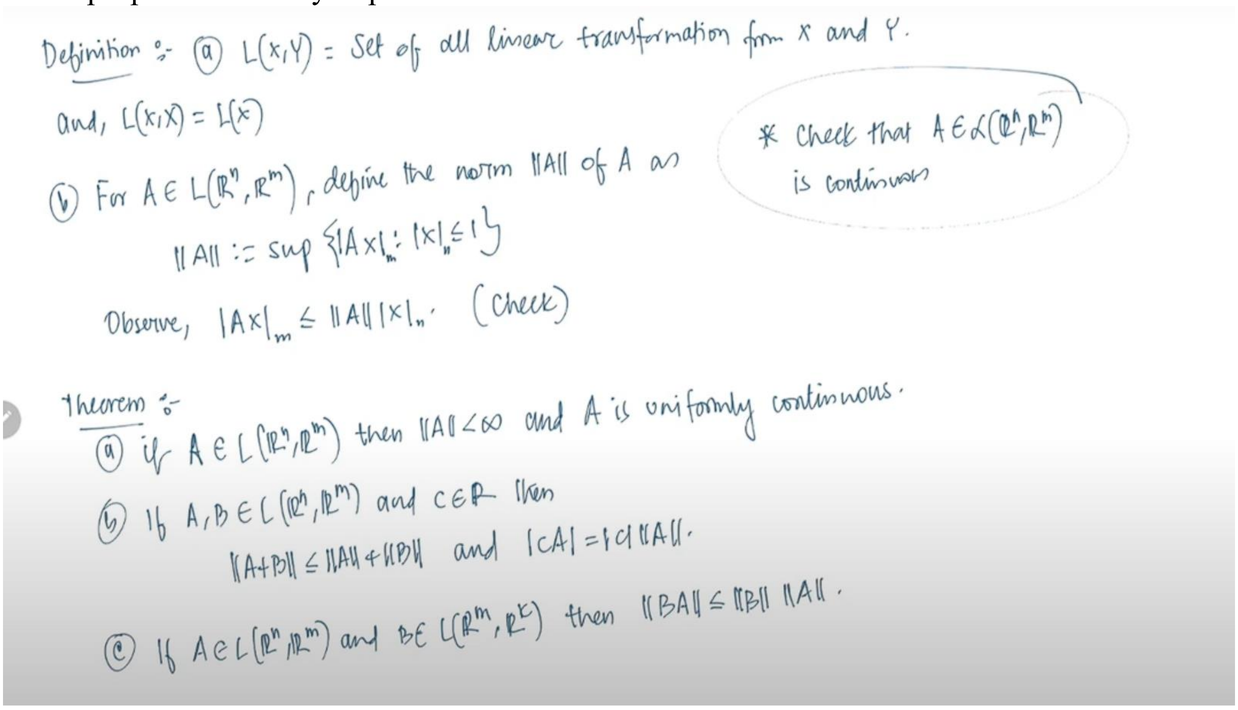
So, this is there.

And of course, what another important thing which we need to do is this.

Let us say if $A \in L(\mathbb{R}^n,\mathbb{R}^m)$ and $B \in L(\mathbb{R}^m,\mathbb{R}^k)$.

And we can define the $||BA|| \leq ||B||\ ||A||$.

These properties are very important.



We are going to use these properties when we define something called a matrix exponential.

So you should know these things for a fact, okay?

So let us look at the proof of this.

Proof.

So the first thing is this.

You have to show that the ||A|| is finite and A is uniformly continuous, right?

That is what we need to show, okay?

So, what we are going to do is we are going to start with a standard basis of $\mathbb{R}^n$.

So, $\{e\ \ , e\ \ , \ldots, e_n\}$, let us say be the standard basis of $\mathbb{R}^n$, okay?

and let us start with an element x=$\sum_{i=}^{m} e_i\, x_i$.

So, once we do that and we also choose this x in such a way that $|x|_n \leq$.

Now, you see why we are choosing it?

Because you see we have to show that ||A|| is finite, right?

And how is ||A|| chosen?

You see we are not looking at the whole of $\mathbb{R}^n$, but only in the unique ball, which is less than equal .

||x|| is less than equal .

So, I am choosing my x accordingly.

Clear?

Once we do that, this will actually imply, so you see, this will actually imply that $|x_i|$≤, okay?

This, of course, you can see, right, this is very easy to see for i=,…,n.

Clear? okay so once that is true then one can write down you see |Ax| how does |Ax| looks like it looks like, |Ax|=$|\sum_{i=}^{n} x_i\, Ae_i| \leq \sum_{i=}^{n} |x_i|\, |Ae_i| \leq \sum_{i=}^{n} |Ae_i|$

Therefore, ||A||≤ $\sum_{i=}^{n} |Ae_i| \leq \infty$.

Which is finite yes so this that is clear because this is just the um n numbers which is basically uh the sum of n numbers which is going to be finite so norm of a is always going to be finite okay and you see and |Ax-Ay|=≤ $||A|||X-Y|$ if x,y∈ $\mathbb{R}^n$, okay

|Ax|$_m \leq \left||A|\right| *$ |x|$_n$

This is the inequality what, I am using here.

So, this is whatever we are using here.

Since this is true, A is uniformly, A is uniformly continuous, uniformly continuous.

Clear?

Okay.

Now, what about the second property?

So, the second property C. Sorry, B. Okay.

So, first of all, I am starting out with

$|(A+B)x|=|Ax + Bx|_m \leq |Ax|_m+|Bx|_m \leq (\ ||A|| + ||B||) * ||x||$

I hope what I am doing is clear to you.

I am using triangle inequality here.

Clear?

Since this vector is in $\mathbb{R}^m$, I can use the triangle inequality in $\mathbb{R}^m$.

Once this is true, of course, you can take the supremum now and you can see that the result follows.

So, taking supremum, the result follows.

And the second part is, I want to, so please check, check.

that this this property that's |cA|=|c||A|

I want you to check this part.

So, please check this part.

This is accordingly done in the same exact same way.

So, if you want to look at |cA|, you start with |cA| acting at an element A and then you do accordingly.

You use triangle integral or whatever appropriate inequality is there, you can actually get it.

Proof :- (a) $\{e_1, e_2, \ldots, e_n\}$ be the standard basis of $\mathbb{R}^n$ and $x = \sum_{i=1}^{n} x_i e_i$, $|x|_n \le 1$

$\Rightarrow |x_i| \le 1$ for $i = 1, 2, \ldots, n$.

Then, $|Ax| = \left|\sum_{i=1}^{n} x_i A e_i\right| \le \sum_{i=1}^{n} |x_i||Ae_i| \le \sum_{i=1}^{m} |Ae_i|$

$\therefore \|A\| \le \sum_{i=1}^{n} |Ae_i| < \infty$

And, $|Ax - Ay| \le \|A\| \, |x - y|$ if $x, y \in \mathbb{R}^n$

$\Rightarrow$ A is uniformly continuous.

(b) $|(A+B)x| = |Ax + Bx|_m \le |Ax|_m + |Bx|_m \le (\|A\| + \|B\|)\|x\|$

Taking sup the result follows.

Check :- $|cA| = |c|\,|A|$

Now, let us look at the third property what you have, C, the third property.

Third property is this, that norm of BA is dominated by norm of B times norm of A. So, again we want to look at what BA does to an arbitrary element X, the norm of that. |(BA)(x)|.

So, this is nothing but |B(Ax)|, that is the definition.

And this is, we can write it like this, Ax, you see Ax is an element here.

Ax is an element of $\mathbb{R}^n$

B acting at element of $\mathbb{R}^m$ $from$ $\mathbb{R}^k$, right?

Okay.

So, you see, this is an element of $\mathbb{R}^n$.

Okay.

And the whole element is in $\mathbb{R}^k$.

So, this is the norm in k. Okay.
$|(BA)(x)| = |B(Ax)|_k \leq ||B|| ||A|| |x|$

See, the thing is once that is done, now you can take the supremum.

Now, taking supremums, you can find out the result.

Taking supremums, we have one has

One has that the $||BA|| \leq ||B|| \, ||A||$ Clear?

So, this is true.

(c) $|(BA)(x)| = |B(Ax)|_k \leq ||B|| \, |Ax| \leq ||B|| \, ||A|| \, |x|$

taking sup one has $||BA|| \leq ||B|| \, ||A||$.

Okay.

Now, so this is more or less what we need to know on the linear transformations.

Okay.

With this, I am going to end this particular video.

Thank you.