

Nonlinear Control Design

Prof. Srikant Sukumar

Systems and Control Engineering

Indian Institute of Technology Bombay

Week 1: Lecture 6 : Preliminaries: Induced Matrix Norms and Signal Norms

Alright, so Banach spaces, Hilbert spaces, we have already done this setup on what kind of space we want. So, we also need notions of matrix norms like I said. So, we work with what is called induced matrix norm. It is induced by the corresponding vector norm and hence it is called the induced matrix norm. So, what is the matrix norm for a matrix A ? By the way this matrix does not need to be square or invertible or anything like that. You can compute the norm for any matrix.

So, the P -norm for a matrix is basically defined using the supremum, yeah, supremum over x not equal to 0. The P -norm of A times x divided by the P -norm of x . I am sorry, this is a matrix norm. Yeah, this is how you, this is a notation for the matrix norm.

Yeah, and this is the definition for the matrix norm. This is exactly how you define it. It is the supremum over all x not 0, $\|Ax\|_P$ and divided by norm of x , $\|x\|_P$. So, obviously you have to imagine that the x has to be compatible with A . I cannot really take arbitrary x and arbitrary vector spaces.

I mean, yeah, if it is a 1 cross 2 matrix, I cannot take a three-dimensional vector. This is so, so x has to be something that you can actually make this multiplication happen. Okay. But remember that the vector, the size of the vector in the numerator and the size of the vector in the denominator do not have to be the same. Yeah.

You understand that you can compute the P -norm for any vector space, right? So, it does not matter what is the size of the vector itself. Okay. So, you compute the P -norm of the numerator and P -norm of the denominator since x is not 0. Therefore, this is, you know, this is a valid operation.

Yeah. And you look at the ratio. This ratio and when you look at the supremum, it is somehow the maximum. It is a generalization of the maximum. Okay. So, when you look at the ratio and you take a supremum, you are essentially measuring the maximum magnification that a matrix gives to a vector.

Okay. Somehow measuring the maximum magnification in some sense. That is how, how much the vector gets, you know, enlarged in some sense, lengthened in some sense. Yeah. By a particular matrix. That is really what you intuitively get out of a matrix norm.

Okay. Makes sense because vector norms also measure the size of the vector or length of the vector. Similarly, here the matrix induced norm measures somehow the size of the matrix but in how it influences the particular vector space. Okay. If you change the vector space, you might get a different answer.

Alright. So, therefore it is induced. Yeah. Remember. Okay. So, what is the supremum? The supremum is defined as the least upper bound.

Okay. This is the best definition of supremum you can have. It is called the least upper bound. Yeah. So, of course you can imagine there is an upper bound always.

Yeah. But then you are looking for the least upper bound. This is what is the supremum. And the more formal definition is that the supremum for a set S in \mathbb{R} closure which is basically $\mathbb{R} \cup \infty$. Yeah. Is the smallest value of y such that for all x in S , x is less than equal to y .

Okay. Exactly the least upper bound written in mathematical terms. Okay. It is not just an upper bound. Okay. If I take this set, if I take this $(0, 1)$ open interval as my set S , then you know that $1, 2, 3, 4, 5$ all of them are upper bounds.

Everything is an upper bound. But the supremum is the least upper bound. Okay. Yeah. Just to you know sort of do a thought experiment, if I say the supremum is actually $1 - \epsilon$ for some positive ϵ .

Okay. But then I know that $1 - \epsilon + \delta$ can also belong to $(0, 1)$. Correct. I just have to choose my δ appropriately. As long as δ is less than ϵ , $1 - \epsilon + \delta$ also belongs to this $(0, 1)$ open interval. But then $1 - \epsilon + \delta$ is greater than $1 - \epsilon$.

So, $1 - \epsilon$ cannot be a supremum. Okay. The simple idea. So, $1 - \epsilon$ is not a supremum for any positive ϵ .

Okay. So, the only possibility is that ϵ is exactly 0 and 1 is the supremum. Okay. Okay.

Least upper bound. Yeah. Again for most of you this may be intuitive in this example that 1 is the supremum and why did I make this proof you know whatever funny looking proof. But in some more complicated cases you have to use these kinds of proof. Okay. Anyway, so in this case it is very obvious that for an open interval $(0, 1)$, 1 has to be the supremum. Because everything below 1 is part of the set.

That is the idea of an open set. Okay. 1 is not included in the set.

Very nice. Very good point. Okay. So, if I thought of $(0, 1)$ open interval as my vector space

then the supremum is not in the set. Okay. This is the difference between a supremum and a maximum. A maximum will always be part of the set.

Okay. So, therefore there are many scenarios of sets for which you cannot define a maximum. And hence we have to talk about a supremum. That is the whole reason for talking about the supremum. Okay. That the maximum may not exist in the set.

Right. Okay. Yes. Say that again. Absolutely. That is also possible. So that is what I am saying. So, there is couple of reasons why you talk about the supremum. One is that the maximum may not be part of the set.

Yeah. So, see whenever I talk about a vector space. Alright. Once I define a vector space you have to almost think that my world ends there. Yeah. Although you know that 2 is there, 3 is there, minus 1 is there.

But for all intents and purposes between 0, 1 is my world. That is the end. Because all my analysis everything I do is inside that set. So, anything that does not lie in that set is a problem for you in some sense.

Or yeah I don't understand it. Okay. So, supremum is a way of sort of giving this extension. Okay. It talks about but again all of this works if there is a super set in some sense. Okay. There is something beyond 0, 1 so everything works.

Yeah. So, of course like you are saying there are sets where there is no maximum if I think about set 0 infinity. Yeah. If I think of a set 0 infinity then you will say that the supremum is infinity.

Okay. So, infinity is allowed. So, therefore when we talk about supremum infinity is allowed. Yeah. When we talk real numbers infinity is not allowed. Yeah. So, if you give me a set 0 infinity I will still say there is a supremum but it is infinity.

Okay. Which is useless sort of. So, I mean mostly you can't do anything with it but anyway we still say that supremum is infinity. Alright. Let us look at another example.

In this case it is a function. Okay. It is a set created from a image of a function. Okay. So, if I look at the function f of x which is $1 - e^{-x}$ and x is actually non-negative real numbers. Yeah. And I look at the set e which is actually the image of f .

I hope you understand what is the image of f . It is just all the values that f takes. Yeah. So, the image of f is actually exactly this. Right. What is the image of f ? It is this guy because it is a nice continuous function.

In fact smooth function. Yeah. So, the image of f is exactly this. So, what is the supremum?

1.

Yeah. But it is not contained in e . Okay. So, basically when the supremum is contained in the set you can just replace the sup notation with the max notation. Okay. Also when the supremum is contained in a set folks who have done real analysis and hopefully SC639 will know that the set is, if the set contains its supremum and infimum then the set is what? Set is a closed set. Okay. If the set contains its supremum and infimum or if the other way round, if the supremum and infimum of a set are in the set then it is a closed set.

Okay. So, we talk about a few matrix properties also. We use a lot of these symmetric matrices, symmetric square matrices to design these Lyapunov functions and so on and so forth. So, we like to know some properties of these. First all eigenvalues of symmetric matrices are real.

Yeah. Most of you should know all of this. A is said to be, A is positive definite if and only if any one of these is satisfied. Yeah. That $\alpha^T A \alpha$ is strictly positive for non-zero α . All eigenvalues of A are strictly positive.

There exists a non-singular decomposition Q . Looks like that and every principal minor of A is positive. Okay. So, a symmetric matrix is positive definite if this happens. Yeah.

We never talk about definiteness of non-symmetric matrices. Okay. One can extend that definition but remember the eigenvalues of non-symmetric matrices are possibly complex.

Okay. So, it does not give you nice results. Okay. So, whenever we say positive definite matrices we are invariably talking about symmetric matrices. Okay. Alright. The other thing for symmetric matrices is this inequality which we use extensively.

I mean this inequality has some name also but I forget it. Yeah. But basically it says that if you take a quadratic product or a quadratic form using this matrix. So, this is what is a quadratic form. Yeah. Then it is lower bounded by $\lambda_{\min} \alpha^T \alpha$ and upper bounded by $\lambda_{\max} \alpha^T \alpha$.

Okay. So, very very important inequality. Very simple but very important inequality we keep using this regularly in our Lyapunov analysis. Okay. So, please remember this. The other important thing to remember is that this expression is virtually impossible to compute by hand.

Yeah. If I ask you to compute a supremum of Ax over x and all that you will actually have to write some code and do some kind of a search or do some optimization to actually find this answer. Yeah. So, very painful. So, there are actually simpler formulae here.

Well known formulae for particular matrix norms. Matrix induced norms. So, the infinity

norm is max absolute row sum. The 1 norm is max absolute column sum and the 2 norm is largest singular value. So, square root of $\lambda_{\max}(A^T A)$. Okay. So, these more or less these three are the ones that get used more often than not.

So, we are not concerned. And then you have this anyway the Cauchy Schwartz inequality. I have said find the general proof but actually the general proof is here. There is a proof here. So, we will look at that later very quickly. But the Cauchy Schwartz inequality is a general inequality for all norms and obviously it is valid also for matrix induced norms.

Matrix induced norm is also a valid norm. By the way as soon as I made a norm definition for matrices, I hope you understand that matrices are also vector spaces. So, I mean they form the linear norm linear space. Yeah. So, this is evident because superposition works. I mean if you take any two same dimensional matrices and then you take a linear combination, it is the same dimension matrix.

Yeah. So, matrices also form a vector space. I mean not different dimensional matrices and all but yeah as long as your matrix dimension is fixed, you are fine. Okay. Yeah. So, of course there are simple examples and I do not know if I should go to these but yeah I mean you have the row sum, the column sum and so on.

So, one norm, infinity norm and two norm are rather easy to compute. Yeah. Basically you just have to apply these formula. Yeah. Very straightforward. All I have done is compute the absolute row sum, absolute column sum, taken the maximum absolute row sum that is the infinity norm, taken the absolute maximum column sum that is the one norm and then the two norm is you have to compute $A^T A$ and compute the largest eigenvalue.

Yeah. Little bit more work. Alright. Great. Great. Any questions? Yes. Alright. So, the nice thing is, I mean I hope you could have guessed by this formula itself. Yeah. That though we use the supremum there, it sort of turns out to have some kind of a maximum expression eventually. Yeah. So, these for when you are talking about real valued matrices, so these nice things happen.

The definitions are meant for general vector spaces and supremum is by definition. I mean I cannot say why not.

I mean and cannot be the maximum. Yeah. For general vector spaces. Okay. So, but for again real valued symmetric matrices, well actually this I apologize. No, no I think these are separate.

Yeah. Yeah. For these you do not need any symmetry or anything. Yeah. I mean I think that is pretty clear. I think only this much is for symmetric and this is for the induced norm. Alright.

So, for the real valued ones, yes you have anyway it becomes a maximum. In general, no.

Right. I cannot make that definition. Yeah. It is the supremum. It will be what it will be. Yeah. If it is not in the set, it is not in the set.

Alright. Supremum always exists. Can be infinity. Infinity is an option. So, that is not an issue as such.

Alright. Great. Now that we have done vector norms and matrix norms.

Yeah. We have to go to signal norms. Yeah. I promise you the final, final norm. Yeah. No more. Yeah. So, see we are progressing pretty linearly.

Yeah. We have states, we have systems, we have states. Yeah. Our aim is to talk about size of states, distance of states from other states and things like that. Yeah. And now states are vectors. So, therefore we looked at vector norms. But then if you think of linear system or if you think of Lyapunov function, there are matrices involved.

So therefore we also need to talk about matrix norms. Finally, when we solve these states and create trajectories, they are functions of time.

Therefore they are signals. Right. So we have to talk about signal norms. Yeah. So the vector norms give you some kind of a point wise behavior. So once I freeze time for example. For example I say that I want to look at the behavior at 5 seconds.

Then my states are a vector. And then I can look at all these vector norms. Yeah. But once, if I want to look at the behavior of a, you know of a signal over a period of time, then I need signal norms.

Okay. And that is what these signal norms do. You will see. So if I am given a vector signal. Yeah. It is almost looking like it is a function of time. You get some non-negative real inputs and you get scalar, sorry vector valued outputs. Yeah. Then the signal norm is the P signal norm or X^P is basically the integral over 0 to infinity.

Why 0 to infinity? Because it is time going from 0 to infinity. And you take any arbitrary vector norm here. X is any arbitrary vector norm. Not necessarily the P norm.

Okay. This is an arbitrary vector norm. Not necessarily the P norm though you are computing the P norm here. Okay. Don't get confused. So the P signal norm is integral 0 to infinity. Vector norm to the power P dt and take the Pth root of this integral.

Similarly the infinity norm again has the supremum. Right. Is supremum over all time of this vector norm. Okay. Supremum over all time of this vector norm. Like I said the vector

norm is arbitrary. Does not have to be the, for example if you are computing the two signal norm, the vector norm you use does not have to be the two norm.

It can be the one norm. Yeah. The only thing you have to remember is for a single problem that you are working out, stick to the same vector norm.

Yeah. Don't switch between 1, 2, 3 and just like you know. Yeah. Then you are not going to get any consistent results. All your results will be wrong.

Yeah. But you are free to choose any vector norm. Yeah. Not restricted by this P or by this infinity. Okay. Doesn't have to be the infinity norm here. Yeah. In fact that would be wrong. So if you have infinity norm here you use infinity norm here and then if you have P norm here you use the P norm here.

So basically you ended up using different vector norms in the same problem. That's not okay.

You have to use the same vector norm for the entire problem. Yeah. You can stick to any vector. Yeah. Is that clear? Yeah.

Yes. Okay. Alright. So like I said in all the above definitions x signifies any vector norm. The choice does not matter. However, never switch the norm in between.

Be consistent. Okay. Whenever a P signal norm is finite we say that the signal belongs to L^p space. Yeah. So signal norms actually define a rather big space of functions. Okay. Rather big class of functions.

Yeah. So these are called this is called the L^p space and this has significance not just in control and so on. Yeah. This has significance in a much larger variety you know array of maths. Yeah. For example, whenever you talk about any kind of function approximation using series you need L^p assumptions.

Yeah. Fourier series for example. When does the Fourier series converge? It only converges when you are in some particular L^p space. Okay. So convergence of series and when do Taylor series converge? When do Fourier series converge? So these sort of you know tools require you to have some kind of an L^p space around. So what I am saying is this L^p space has much much wider application than just in control.

We are just using it for some very small purpose is what I would say. Okay. So the other thing to remember is that L^∞ that is when the infinity norm is bounded it is exactly the same as being as saying that x is itself a bounded signal.

Okay. Proof is pretty straight forward. Yeah. I am not sure if you need to I mean anyway I

can look at one side of it. So the infinity norm is this guy. Yeah. If a function is bounded it means that there exists some positive number such that the norm is smaller than that positive number.

So if the norm is smaller than M the vector norm is smaller than M . Notice this is the vector norm notation. Yeah. Now whenever I put the time argument then I am computing a vector norm because I froze time. Whenever there is no time argument notice because of because I integrated or I took supremum the time argument is vanished.

Right. Therefore on the left hand side you never see any time argument. Okay. Whenever I am computing a signal norm there is no time argument.

Okay. So this is important notation wise remember. Yeah. So no time argument here but in all the vector norms here you see the time argument. Yeah. Because without freezing time I don't have a vector I have a signal.

Right. So once I put some particular time in here I have a vector then I can compute a vector norm. Okay. So this will be our notation. Whenever I compute a vector norm the time argument will be evident will appear and if I am computing a signal norm there will be no time argument.

Alright. Great. So when I say here that a function signal is bounded it means that in the norm it is less than equal to some upper bound. Yeah. Which means that supremum itself is also upper bound. And because for all time this is less than equal to M therefore the supremum itself is upper bound.

So very straightforward. Yeah. Which means that x belongs to L infinity. Similarly if the infinity norm is some constant it means that the supremum is some constant which means that every for all time this has to be less than equal to M .

Therefore function is signal is bounded. Okay. So both ways you can prove essentially L infinity is identical to having a bounded signal. Okay. So L infinity signals are bounded signals. Everything in L infinity space is a bounded signal. Yeah. The interesting thing to remember is that if you take any two vector norms there is a property of norm equivalence.

What is norm equivalence? It means that your p norm and q norm are always relatable by constants α and β . Okay. For any value of x t . Doesn't matter what how big x t becomes or how small x t becomes. The two norms the two norms p and q norms vector norms are relatable by a constant.

Yeah. So α norm x t p is less than equal to x t q less than equal to β norm x t p . Okay. And vice versa. Yeah. If you have p in the middle you can still find some other constants.

Right. It will be one over alpha and one over beta. Okay. So the idea is vector norms can always be related to each other. Yeah. So they have what is called norm equivalence. On the other hand signal norms do not have norm equivalence. If you have a signal which is in l_1 space it may not be in l_2 space.

If it's in l_∞ space it may not be in l_2 space. So if it just because one norm is behaving nicely doesn't mean the other will. Yeah. So a very simple example is this guy.

If I take my signal as this guy. Yeah. And I use my l_1 and I use the two vector norm. I've decided to use the two vector norm. Yeah. Because I know that it's easy. The two vector norm of this guy is just one.

Right. Do you believe that. The two vector norm of $\cos^2 t$ is just one because it's $\cos^2 t$ plus $\sin^2 t$ square root which is one. Okay. So if I take the infinity norm it is just the supremum of this guy. So it is supremum of $\cos^2 t$ plus $\sin^2 t$ and it's just one.

Right. On the other hand if I take the one norm then this is integral zero to infinity of one which is infinity. I just took the two norm again here. Two norm here too. No problem. Yeah. So the infinity norm is just one which essentially indicates that it's a bounded signal and it is right obviously bounded signal but the one norm is actually infinity.

It's a. No. No. No. Boundedness is only connected to the infinity norm. One norm two norm do not have any connotations with boundedness. They are not connected to boundedness at all. They are you can think of them as different function spaces with different properties.

Yeah. The infinity norm is bounded which means that the function is also bounded. However the function is not in L_1 space at all. Yeah.

And this is a big conundrum. Right. I mean you see that it has nice infinity norm nice L_1 . Yeah. But it is sorry nice L_∞ norm but it is not in L_1 space. Yeah.

Similarly L_2 and so on. You can compute L_2 norm any L_p norm for example for that matter. Yeah. This function is definitely in L_∞ but it is not in any L_p . Yeah. Because it will always have one integrated from zero to infinity and that will essentially give you infinity. Okay. So norm equivalence doesn't work for signal norms. Alright. Thank you.