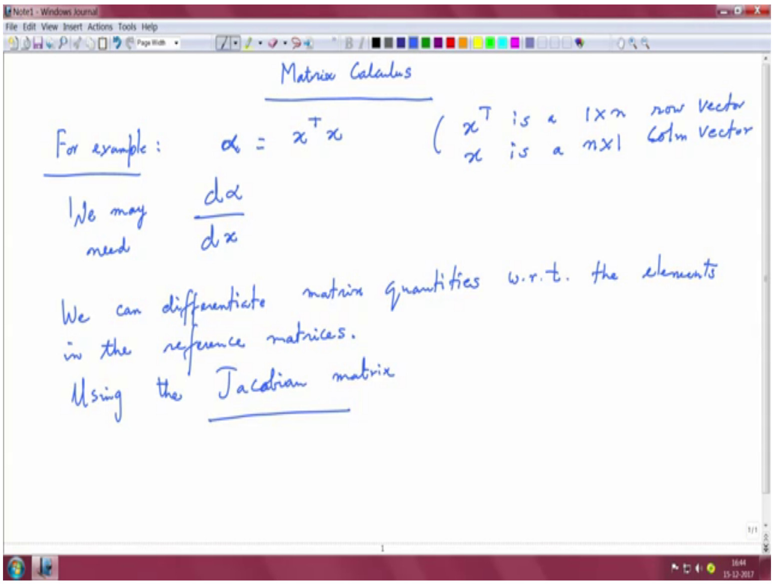
Mathematical Methods and Techniques in Signal Processing – I
Prof. Shayan Srinivasa Garani
Department of Electronic Systems Engineering
Indian Institute of Science, Bangalore

Lecture – 74
Matrix Calculus

So, we have seen cases, where it is required for us to differentiate with respect to vectors, with respect to matrices etcetera. So, this is often encountered in signal processing.

(Refer Slide Time: 00:36)



For example, a simple case would be a form of a scalar which is x transpose x, right. Imagine x is a column vector. So, therefore, x transpose is a 1 by n row matrix and x is n by 1 column I would say row vector, this is a row vector this is a column vector and we may want for some instance different take the derivative of alpha with respect to x. So, we may need the alpha by del x.

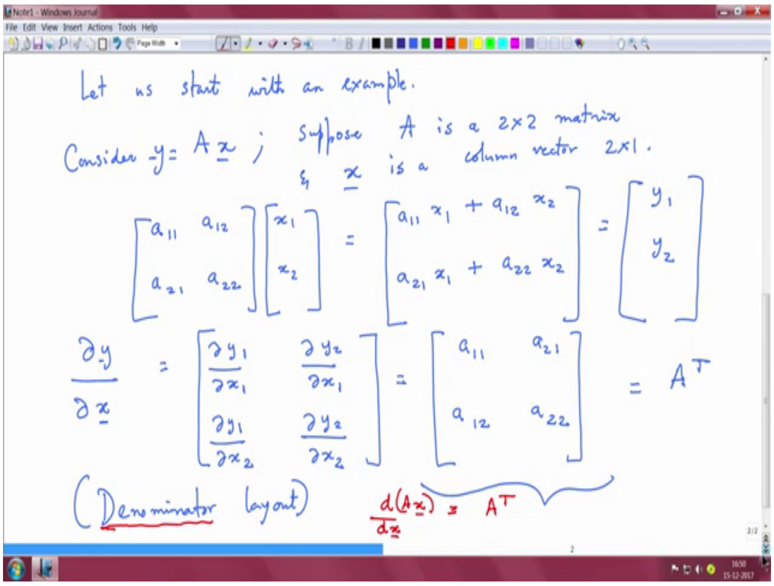
So, how to go about problems of this form and this is particularly useful when we deal with representing problems vectorially or in the form of matrices and then we have to take derivative of some function of vectors and matrices and compute these quantities and this becomes a nontrivial task.

So, unless you know how to play around with the algebra we will often get stuck and if you do in a routine way we will have to really do this as scalar derivatives over each element by element computation.

So, we will figure out some intuition behind how to go about doing this matrix calculus and then this will be useful when we discuss for example, in k l transforms we will find this often helpful this is an old concept nothing really new, but I would like to give you some idea here.

So, we can differentiate matrix quantities with respect to the elements in the reference matrices and often this is done using the Jacobian matrix. So, there are certain conventions I will present to you an intuition how this works and then we can discuss certain properties, ok.

(Refer Slide Time: 03:54)



So, let us just start with an example. So, consider A times x. So, y is some A times x suppose A is a 2 by 2 matrix and x is a column vector and this is say suppose 2 by 1 right.

So, let us just compute these quantities a is a 1 1, a 1 2 a 2 1, a 2 2 this is our standard notation right the first row first column, the first row second column, second row first column and the second row second column and then we have x 1, x 2 these are the elements of this vector x right. So, I just put this bar underneath y index indicated that they are vectors ok.

Now, if you compute this is a 1 1 x 1 plus a 1 2 x 2 this is a 2 1 x 1 plus a 2 2 time x 2 and this is the vector y composed of elements y 1 and y 2. Now, let us adopt one convention which is basically the hessian rule.

So, if I want to take the derivative of y with respect to x, so that means, I am trying to differentiate I am I am differentiating the vector y with respect to x and this is given in this format this is del by y 1 upon del by x 1 then del by y 2 upon del by x 1 that is I vary in the row I vary the coordinates of y that is I choose y 1, y 2 and then I take y 1; I differentiate with respect to x 1; I choose y 2 differentiate with respect to x 1 so on and so forth.

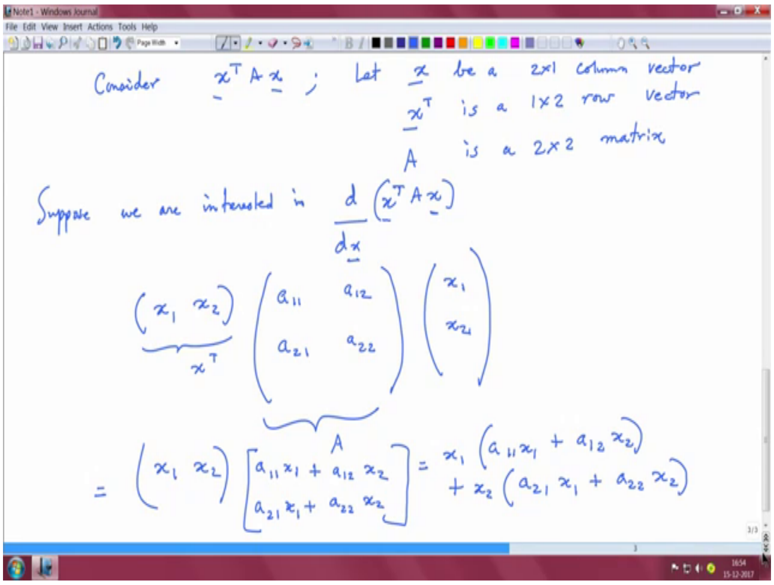
Similarly, second row I look at y 1 take the derivative with respect to x 2 then look at y 2 take the derivative with respect to x 2 and you may really question I mean should I go this way you do not have to for example, if you look at a different format then you can take y 1 del by y 1 by del by x 1 you can choose del by y 1 upon del by x 2 and so on and so forth.

So, this is a different convention. So, there are two different conventions one convention basically keeps the first coordinate of the vector constant in the row and then take the derivative we can follow this convention here which is basically the denominator layout and I mentioned this sort of carefully as long as you are consistent it is fine.

So, you basically keep the y coordinate fixed in there I mean the x coordinate fixed in the denominator and you let the y coordinate basically vary you know within the row that is across the columns. So, this is this is the convention that we will follow. So, now, if you take the derivative del by y 1 upon del by x 1 this is basically a 1 1 and del by y 2 if you take the derivative of y 2 with respect to x 1 you get a 2 1 and similarly if you populate the rest of the entries you get a 1 2 here you get a 2 2 here and observe that this matrix is basically the transpose of A right.

So, we can formulate a rule here that is if I want to differentiate the quantity A x with respect to x. So, x is a vector I get A transpose, but of course, you should remember the convention which is basically we are following denominator layout and your calculus has to be sort of consistent I mean if you are basically varying y in the row and keeping x fixed then you should be consistent with respect to that you do not want to mix once when you differentiate you want to have y varying other time, x varying I think that should not really happen, ok.

(Refer Slide Time: 09:46)



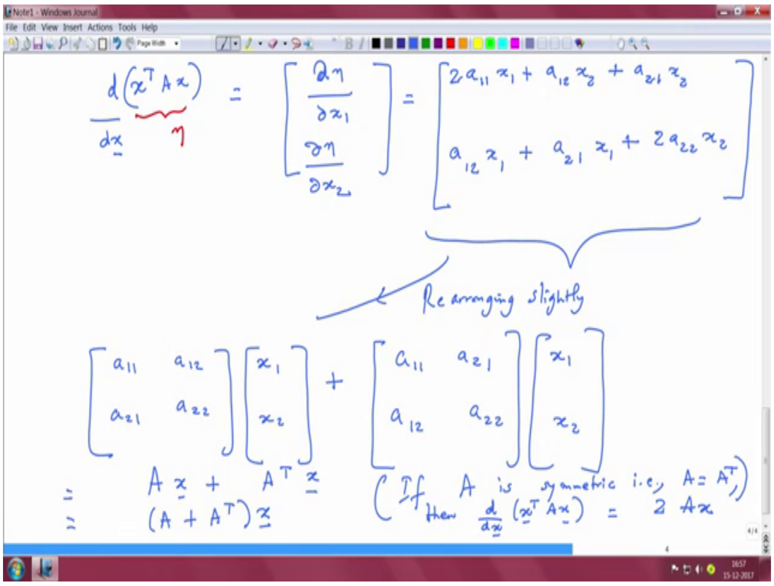
Now, let us take another example this would be a good example for us consider x transpose A x you will see this form in many many cases, ok. Now, let us look at the math for working out this example. Now, let x be a 2 by 1 column vector, ok. Now, x transpose is a 1 by 2 row vector and a is our 2 by 2 matrix like in the last example.

So, again this is a scalar x transpose a x is a scalar like possibly you think about x transpose x is a scalar I mean if you just multiply you know what makes instead of x transpose x I have x transpose a x and this is a scalar we will find this form this quadratic form in many signal processing applications. So, let us just get an intuitive feel how this works.

So, suppose we are interested in taking derivative with respect to x of the quantity x transpose A x ok. Let us work out the math this is routine computations to get a feel an intuitive feel how this works, right. I have x 1 x 2 this is my x transpose right. So, this is my x transpose now my a is a 1 1, a 1 2 a 2 1 a 2 2 right. This is my A and my x is basically x 1 x 2. Now, if I compute this quantity right I get basically x 1 x 2 let us do an interim step.

So, A x is basically what we had earlier which is a 1 1 x 1 plus a 1 2 x 2 we have a 2 1 x 1 plus a 2 2 x 2, right. When we simplify this is basically x 1 times a 1 1 x 1 plus a 1 2 x 2 plus x 2 times a 2 1 x 1 plus a 2 2 x 2 right. Now, we will adopt the same convention let us differentiate the scalar with respect to x 1 and x 2 right. So, earlier case we had a vector and we were differentiating that vector with respect to the with respect to another vector, now we are differentiating a scalar with respect to a vector.

(Refer Slide Time: 13:56)



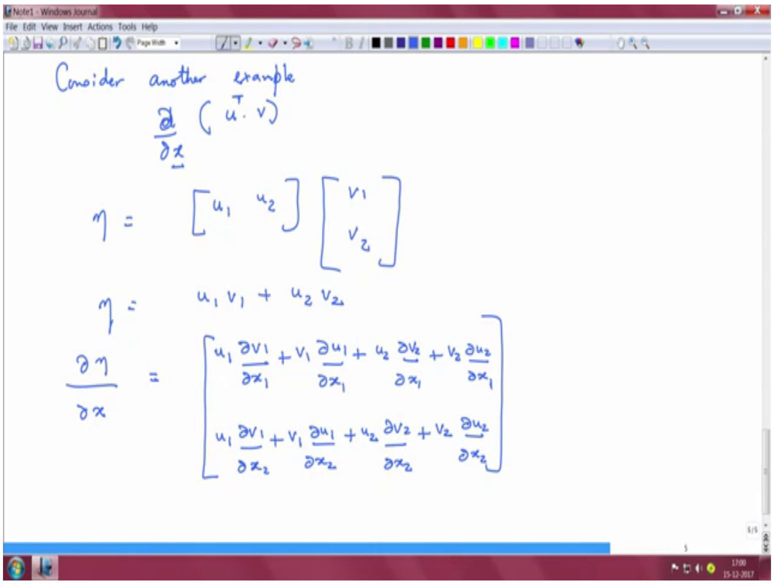
Now, derivative of x transpose A x by dx. So, this can be done as follows this is basically derivative of. So, let me call this quantity here you know the scale as eta just for convention. So, this is basically eta here ok. So, now, we have del eta by del beta by del by x 1 and then you have del by eta upon del by x 2. If you work out this carefully you will end up with the following this is going to be 2 times a 1 1 x 1 plus a 1 2 x 2 plus a 2 1 x 2 and then you have the other term which is a 1 2 x 1 plus a 2 1 x 1 plus 2 a 2 2 x 2.

Now, let us try to rearrange this right. We will end up as follows where a 1 1, a 1 2, a 2 1, a 2 2 you can just verify this that it is going to be of this form. So, you will have to have 2 times a 1 1 that has to appear and you have one term here and here another term here right I think this is straightforward to verify.

So, now this can be written as A x plus A transpose x which is basically A plus A transpose times x. So, if A is symmetric that is A equals A transpose right then derivative of x transpose A x with respect to x is going to be 2 times a times x. So, I think is a standard result.

Now, you will find this form very very helpful when you when you deal with calculus I mean I just gave you an intuitive feel how you can do this for a 2 by 2 matrix and 1 by 2 vector and you can sort of generalize this to A which is an n by n matrix x being n by 1 column vector and x transpose being a 1 by n row vector and the results are pretty straightforward ok. So, let us see one other case will give you an example again.

(Refer Slide Time: 17:52)

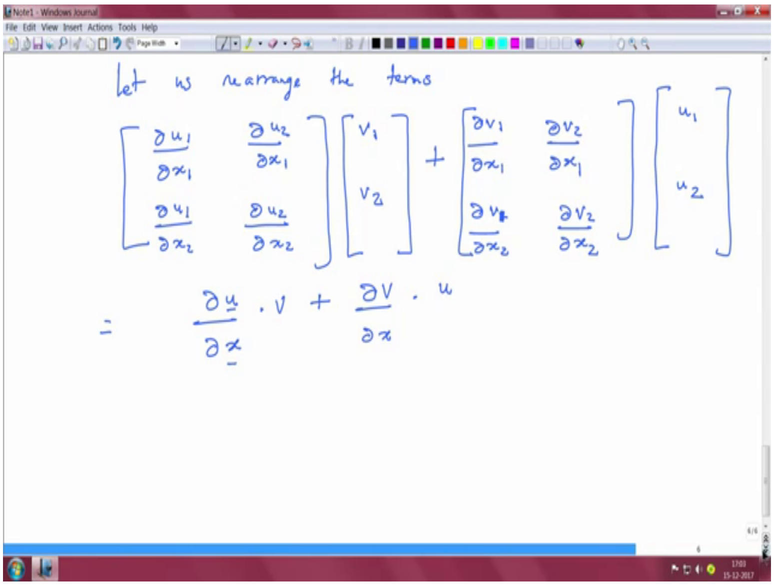


Consider another example where I want to take the derivative with respect to x of some product of 2 vectors u and v right. So, let us form this scalar which is u 1, u 2 times v 1, v 2 now that you could call this possibly u transpose following the same convention that we followed right.

Now, eta in general is say u 1 v 1 plus u 2 v 2 they are two different vectors. Now, if you are interested in taking the derivative with respect to x of the scalar eta then you will end up with the chain rule as follows, right. So, it is u one del by v 1 upon del by x 1 plus v 1 del by u 1 upon del by x 1 plus u 2 del by v 2 upon del by x 1 plus v 2 del by u 2 upon del by x 1.

So, just carefully observe that I kept this x 1 fixed throughout in the in the in the in the derivative I assumed this form. So, this is basically the denominator layout. Now, the second term of course, would be u 1 del by v 1 upon del by x 2 plus v 1 del by u 1 upon del by x 2 plus u 2 del by v 2 upon del by x 2 plus v 2 del by u 2 upon del by x 2. So, if you rearrange the terms carefully, right.

(Refer Slide Time: 20:37)

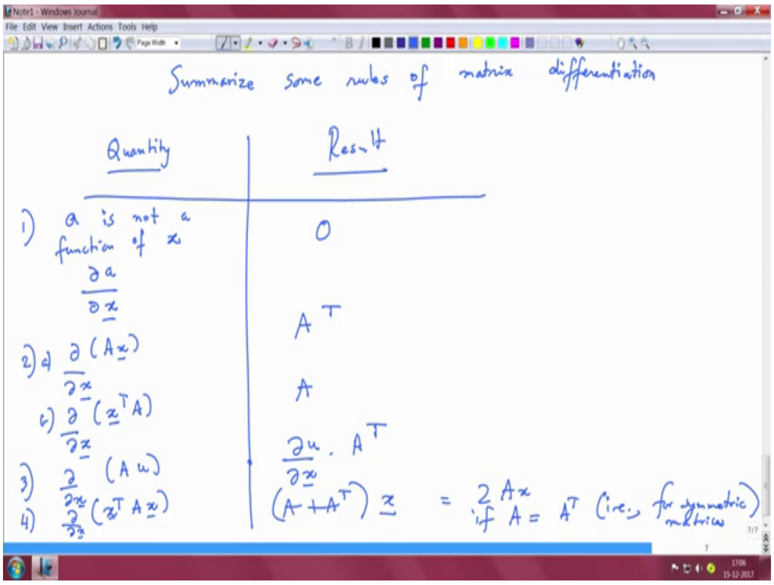


Let us try to do that you will get del by u 1 on del by x 1 del by u 2 on del by x 1 del by u 1 upon del by x 2 del by u 2 upon del by x 2 and then you will have a vector v 1 v 2 here. Similarly, you will have one more matrix as you can similarly imagine this is going to be the vector u 1 the vector u comprised of u 1 and u 2 and you can think of del by v 1 upon del by x 1 del by v 2 upon del by x 1 del by v 2 on del by x 2 and del by v 2 upon del by x 2 this is a this is v 1 here del by v 1 upon del by x 2 del by v 2 upon del by x 2, ok.

So, this is basically written in the form del by u upon del by x times v the in dot product with the times v plus del by v upon del by x times u.

Of course, we have to see that this is not the dot product it is just a normal multiplication you will find is very very useful. So, I think now since you have gotten a feel an intuitive feel how to go about doing the calculus I will give you some quantity of a table comprising of these derivatives you can just directly verify those as a part of a homework exercise and basically you can convince yourself that these results are correct, ok.

(Refer Slide Time: 23:15)



So, let me summarize some rules of matrix differentiation [noise. Now, let us see the quantity and then the result right. First is a is a constant that is not a function of x and if I am interested in del by a upon del by x the result should be 0.

Then, we saw that del by a x x is a vector upon del by x and this is A transpose and you could do the other way around you could take x transpose here right, that is another possibility if you did del by of x transpose a with respect to del by x you should get A here ok, two different variations A and B then del by del by x A times u, where u is some function of x and is a constant matrix you get del by u upon del by x times a transpose then we have this important result del by of x transpose A x upon del by x this is A plus A transpose times x and when A is symmetric which is equal to 2 A x if A is A transpose that is for symmetric matrices.

So, I think using these basic rules for doing matrix calculus. It is very easy for you to take a derivative of a scalar with respect to a vector or scalar with respect to a matrix and so on and so forth and we will explore these results when we discuss you know some optimization that requires taking derivatives of certain quantities with respect to vectors.

So, with this we conclude some basics on Matrix Calculus that is required for your mathematical methods.