

The following content is provided under a Creative Commons license. Your support will help MIT OpenCourseWare continue to offer high-quality educational resources for free. To make a donation or view additional materials from hundreds of MIT courses, visit MIT OpenCourseWare at ocw.mit.edu.

PROFESSOR: Hi. In today's lesson, hopefully we will begin to reap the rewards of our digression into the subject of linear algebra. Recall that in the last few lectures, what we have been dealing with is the problem of inverting systems of linear equations. And what we would like to do today is to tackle the more general problem of inverting systems of equations, even if the equations are not linear. And with this in mind, I simply entitle today's lesson "Inverting More General Systems of Equations."

And by way of a very brief review, recall that given the linear system y_1 equals $a_{11}x_1$ plus et cetera $a_{1n}x_n$ up to y_n equals $a_{n1}x_1$ plus et cetera $a_{nn}x_n$. We saw that that system was invertible, meaning what? That we could solve this system for the x 's as linear combinations of the y 's if-- and only if-- the inverse of the matrix of coefficients of this system exists.

And in terms of determinants, recall that that means if and only if the determinant of the matrix of coefficients is not 0. If the determinant of the matrix of coefficients was 0, then we saw that y_1 up to y_n are not independent. In fact, in that context, that's where we came to grips with the concept of a constraint, that the constraint actually turned out to be what? The fact that the y 's were not independent. Meaning that we could express a linear combination of the y 's equal to 0.

So what situation were we at then? Given a linear system, if the matrix of coefficients does not have its determinant equal to 0, the system is invertible. We can solve for the x 's in terms of the y 's. If the determinant of the matrix of coefficients is 0's, the y 's are not linearly independent. In other words, the system is not invertible.

And now what we would like to do is to tackle the more general problem of inverting

any system of equations. And by any, I mean what? Now we have y_1 is f_1 of x_1 up to x_n , et cetera. y_n is f_n of x_1 up to x_n . And now what we're saying is we do not know whether the f 's are linear or not. In fact, if they are linear, we're back to-- as a special case, what we've tackled before. But now we're assuming that these need not be linear. And what we would like to do is to invert this system, assuming, of course, that such an inversion is possible.

Again, what do we mean by the inversion? We mean somehow, we would like to know that given this system of n equations and n unknowns where the y 's are expressed explicitly in terms of the x 's, can we invert this, and express the x 's in terms of the y 's-- either explicitly or implicitly? That's the problem that we'd like to tackle.

And what we're going to use to tackle this is our old friend, the differential, the linear approximation, again, that motivated our whole study of linear systems in the first place. Remember, we already know that if y_1 is a differentiable function of x_1 up to x_n , that Δy_1 is exactly equal to the partial of f_1 with respect to x_1 times Δx_1 plus et cetera the partial of f_1 with respect to x_n times Δx_n . And in terms of reviewing the notation, because we will use it later in the lecture, notice that Δy_1 was what we abbreviated to be dy_1 . And that Δx_1 up to Δx_n are abbreviated respectively as dx_1 up to dx_n . Generalization of what we did for the differential in the case of one independent variable and we went through this discussion when we talked about exact differentials in block three.

And in this similar way, we could of course express Δy_2 , et cetera, and we can talk about the linear approximations. All right? Not the true Δy 's now, but the linear part of the Δy 's, the Δy_1 sub lin, or if you prefer, Δy_1 sub lin, L-I-N. All right?

And the key point now is what? If the y 's happen to be continuously differentiable functions of the x 's in the neighborhood of the point \bar{x} equals \bar{a} -- in other words, x_1 up to x_n is equal to a_1 comma et cetera up to a_n , x_1 equals a_1 , x_2 equals a_2 , et cetera-- then near that point, what we're saying is what? That the error term

goes to 0 very rapidly. And as long as the functions are continuously differentiable, it means what? That the change in y is approximately equal to the change in y sub i times \tan .

So that what we're saying is-- and remember this is what motivated our linear systems in the first place-- that Δy_1 is approximately the partial of y_1 with respect to x_1 , evaluated when \bar{x} is \bar{a} times Δx_1 plus et cetera the partial of y_1 with respect to x_n also evaluated when \bar{x} equals \bar{a} times Δx_n et cetera, down to Δy_n is approximately equal to the partial of y_n with respect to x_1 times Δx_1 , plus et cetera the partial of y_n with respect to x_n times Δx_n , where all of these partials are evaluated specifically at the point \bar{x} equals \bar{a} . The point being what? That since we're evaluating all these partials at a particular point, every one of these coefficients is a constant.

You see in general, these partial derivatives are functions, but as soon as we evaluate them at a given value, they become specific numbers. So this is now what? On the right-hand side, we have a linear system, and the point is that we are using the fundamental result that we can use the linear approximation as being very nearly equal to the true change in y . And it's in that sense that we have derived our system of n linear equations and n unknowns.

You see, this is a linear system. The approximation-- again, let me emphasize that, because it's very important. The approximation hinges on the fact that what this is exactly equal to would be $\Delta y_1 \tan$ et cetera, Δy_n . But we're assuming that these are close enough in a small enough neighborhood of \bar{x} equals \bar{a} so that we can make this particular statement. So this is a linear system. And because it is a linear system, we're back to our special case that this system is invertible if and only if the determinant of the coefficients-- matrix coefficients-- is not 0.

And what is that matrix coefficient? It consists of n rows and n columns and the row is determined by the subscript on the y . And the column is determined by the subscript on the x . So we write that matrix as the partial of y sub i with respect to x

sub j . In other words, the i -th row involves the y 's and the j -th column, the x . All right? And that's exactly, then, how we handle this particular system. OK? Quite mechanically.

And let's just summarize that then very quickly. If f_1 et cetera and f_n are continuously differentiable functions of x_1 up to x_n near \bar{x} , then the system $y_1 = f_1(x_1, \dots, x_n)$ et cetera, $y_n = f_n(x_1, \dots, x_n)$ -- that system is invertible if and only if the determinant-- the n by n determinant whose entry in i -th row, j -th column is the partial of y_i with respect to x_j if and only if that determinant is not 0.

Now what does that mean, to say that it's invertible? It means that we can solve for the x 's in terms of the y 's. Now we may not be able to do that explicitly. The best we may be able to do is to do that implicitly, meaning this. Let me just come back to something I said before to make sure that there's no misunderstanding about this. What we're saying is, that in this particular linear system of equations, as long as this determinant of coefficients is not 0, we can explicitly solve for Δx_1 up to Δx_n in terms of Δy_1 up to Δy_n , even if we may not be able to solve explicitly for the x 's in terms of the y 's. In other words, the crucial point is, we can solve for the changes in x in terms of the changes in y . And that is implicitly enough to see what the x 's look like in terms of the y 's.

Once we know what the change of x looks like in terms of the change in y , then we really know what x itself looks like in terms of the y 's, even as I say it may be implicitly rather than explicitly. At any rate, this matrix is so important that it's given a very special name, definition. The matrix whose entry in the i -th row, j -th column is the partial of y_i with respect to x_j is called the Jacobian. I put "matrix" in quotation marks here, because some people refer to the Jacobian meaning a matrix. Other people call the Jacobian the determinant of the Jacobian matrix.

I'm not going to make any distinction this way. It'll be clear from context. Usually when I say the "Jacobian," I will mean the Jacobian matrix. I might sometimes mean the Jacobian determinant. And so to avoid ambiguity, I will hopefully say "Jacobian

matrix" when I mean the matrix, and "Jacobian determinant" when I mean the determinant. But should I forget to do this or should you read a textbook where the word Jacobian is used without the proper noun after it, it should be clear from context which is meant. But at any rate, that's what we mean by the Jacobian of y_1 up to y_n with respect to x_1 up to x_n .

And this Jacobian matrix is often abbreviated by-- you either write J for Jacobian of y_1 up to y_n over x_1 up to x_n . Or else you use a modification of the partial derivative notation. And you sort of read this as if it said the partial of y_1 up to y_n divided by the partial of x_1 up to x_n . And again, there is the same analogy between why this notation was invented and why the notation dy divided by dx was invented. But in terms of giving you a general overview of what we're interested in, I think I would like to leave the discussion of why we write this in a fractional form to the homework. In other words, as either a supplement to the learning exercises or else as part of the supplementary notes in one form or another, we will take care of all of the computational aspects of how one handles the Jacobian.

But what I wanted to do now was to emphasize how one uses the Jacobian matrix and differentials to invert systems of n -equations and n -unknowns. And I will use the technique that's used right in the textbook and which is part of the assignment for today's unit. The example that I have in mind-- I simply picked the usual case, n equals 2, so that things don't get that messy. Using again, the standard notation when one deals with two independent variables, let u equal x squared minus y squared. Let v equal to $2xy$. Let's suppose now that I would like to find the partial of x with respect for u , treating v as the other independent variable.

You see, again, I want to review this thing. When I say find the partial of x with respect to u holding v constant, it is not the same as finding the partial of u with respect to x from here and then just inverting it. Namely the partial of u with respect to x here assumes that y is being held constant. And if you then invert that, recall that what you're finding is the partial of x with respect to u treating y as the other variable. We want the partial of x with respect to u treating u and v as the pair of independent variables.

Why? Because that's exactly what you mean by inverting this system. This system as given expresses u and v in terms of the pair of independent variables x and y . And now what you'd like to do is express the pair of variables x and y in terms of the independent variables u and v , assuming of course, that u and v are indeed independent variables.

The mechanical solution is simply this. Using the language of differentials, we write down du and dv . Namely, du is what? The partial of u with respect to x times dx , plus the partial of u with respect to y times dy . And from the relationship that u equals x squared minus y squared, we see that du is $2x dx$ minus $2y dy$. Similarly, since the partial of v with respect to x is $2y$, and the partial of v with respect to y is $2x$, we see that dv is $2y dx$ plus $2x dy$.

If we now assume that this is evaluated at some point x_0, y_0 , what do we have over here? Once we've picked out a point x_0, y_0 to evaluate this at-- and I left out that because it simply would make the notation too long, but I'll talk about that more later. Assuming that we've evaluated this at a particular fixed value of x and y , we have what? du is some constant times dx plus some constant times dy . dv is some constant times dx plus some constant times dy . In other words, du and dv are expressed as linear combinations of dx and dy .

We know how to invert this type of a system, assuming that it's invertible. Sparing you the details, what I'm saying is what? I could multiply, say, the top equation by x , the bottom equation by y . And when I add them, the terms involving dy will drop out. And I will get $x du$ plus $y dv$ is $2x dx$ plus $2y dx$. In other words twice-- it's $2x$ squared. I multiply the top equation by x , the bottom equation by y . So the right-hand side here becomes $2x$ squared dx plus $2y$ squared dx , which is twice x squared plus y squared dx .

I now divide both sides of the equation through by twice x squared plus y squared. I wind up with the fact that dx is x over twice the quantity x squared plus y squared du , plus the quantity y over twice x squared plus y squared dv . Recall that I also know by definition that dx is the partial of x with respect to u times du plus the partial

of x with respect to v times dv , recalling from our lecture on exact differentials that the only way two differentials in terms of the du and dv can be equal is if they're equal coefficient by coefficient. I can therefore equate the two coefficients of du to conclude that the partial of x with respect to u is x over 2 times the quantity x squared plus y squared.

In fact, I can get the extra piece of information even though I wasn't asked for that in this problem, that the partial of x with respect to v is y over twice the quantity x squared plus y squared. By the way, observe purely algebraically that the only time I would be in any difficulty with this procedure is if x squared plus y squared happened to equal 0. In other words, if x squared plus y squared happened to equal 0, then to divide through by twice x squared plus y squared is equivalent to dividing through by 0. And division by 0 is not permissible.

In other words, somehow or other, I must take into consideration that I am in trouble if x squared plus y squared is 0. Notice, by the way, that the only time that x squared plus y squared can be 0 is if both x and y are 0. And that means, again, that somehow or other, at the point 0 comma 0 -- in a neighborhood of the point 0 comma 0 in the neighborhood of the origin, I can expect to have a little bit of trouble.

Now again, the main aim of the lecture is to give you an overview. The trouble that comes in at the origin will again be left for the exercises. In a learning exercise, we will discuss just what goes wrong if you take a neighborhood of the origin to discuss the change of variables u equals x squared minus y squared v equals $2xy$. Suffice it to say for the time being that the system of equations u equals x squared minus y squared, v equals $2xy$, is invertible in any neighborhood of a point x 0 comma y 0 except in that one possible case when you have chosen as the point x 0 y 0 , the origin.

At any rate, to take this problem away from the specific concrete example that we've been talking about and to put this in terms of a more general perspective, let's go back more abstractly to the more general system. Let's suppose now that u and v

are any two continuously differentiable functions of x and y . Let u be f of xy . Let v equal to g of xy . And what we're saying is, if you pick a particular point x_0 comma y_0 , then by mechanically using the total differential, we have that du is the partial of f with respect to x evaluated at x_0 y_0 times dx , plus the partial of f with respect to y , evaluated at x_0 y_0 times dy . We have that dv is the partial of g with respect to x evaluated at x_0 y_0 times dx plus the partial of g with respect to y , evaluated at x_0 y_0 times dy .

What is this now? This is a linear system of two equations in two unknowns. du and dv are linear combinations of the dx and dy . The key point being again-- that's why I put this in here specifically with the x sub 0 and the y sub 0 -- the key point is that as soon as you evaluate a partial derivative at a fixed point, the value is a constant, not a variable. So this is what? A linear system. We have du as a constant times the x plus a constant times dy . dv is a constant times dx plus a constant times dy .

Again, to make a long story short, I can solve for dx in terms of du and dv . I can solve for dy in terms of du and dv , provided what? That my matrix of coefficients does not have its determinant equal to 0 . And to review this more explicitly so you see the mechanics, all I'm saying is, to solve for dx , I can multiply the top equation by the partial of g with respect to y evaluated at x_0 y_0 . I can multiply the bottom equation by minus the partial of f with respect to y evaluated at x_0 y_0 . And then when I add these two equations, the dy term will drop out.

Again, leaving the details for you, it turns out that dx is what? The partial of g with respect to y -- and I've abbreviated this again-- this means what? Evaluated at x_0 y_0 times du minus the partial of f with respect to y at x_0 y_0 times dv over the partial of f with respect to x times the partial with g with respect to y minus the partial of f with respect to y times the partial of g with respect to x .

And notice, of course, that this denominator is precisely our matrix of coefficients. f sub x , f sub y , g sub x , g sub y . And the only place I've taken a liberty here is to use the abbreviation of leaving out the x_0 y_0 . And the key point is what? The only place I am going to get in trouble is if this denominator happens to be 0 .

In the two by two case-- in other words, in the case of two equations and two unknowns, notice that we can see explicitly what goes wrong when the determinant of coefficients is 0. The determinant of coefficients is just this denominator. And when that denominator is 0, we're in trouble. In other words, the only time we cannot invert this system, the only time we cannot find Δx and Δy in terms of du and dv is when this determinant is 0.

Now you see, I think this is pretty straightforward stuff. The textbook has a section on this, as you will be reading shortly. It is not hard to work this mechanically. And then the question comes up is, how come when you pick up a book on advanced calculus, there's usually a huge chapter on Jacobians and inversion? Why isn't it this simple in the advanced textbooks? Why can we have it in our book this simply, but yet, in the advanced book, why is there so much more to this beneath the surface? The answer behind all of this is quite subtle. In fact, the major subtlety is this. And that is that the notation du -- and for that matter, dv or dx or dy or dx_1 , dx_2 , whatever you're using here-- is ambiguous.

And it's ambiguous for the following reason. Recall how we defined the meaning of the symbol du . If we're assuming that u is expressed as a function of the independent variables x and y , then by du , we mean Δu . On the other hand, if we inverted this, du now-- in other words, what do I mean by inverted this? What I mean first of all is if we assume now that x and y are expressed in terms of u and v -- for example, suppose x is some function h of u and v , what does du mean now? Notice that now u is playing the role of an independent variable. For the independent variable, du just means Δu .

In other words, by way of a very quick review, notice that if we're viewing u as being a dependent variable, then du means Δu . But if we're viewing u as being an independent variable, then du means Δu . And consequently, the results that we're using hinge very strongly then-- in other words, the inversion that we're using hinges very strongly on the requirement. In other words, the inversion requires the validity of interchanging Δu and Δu , et cetera.

Now let me show you what that means more explicitly. Let's come back to something that we were talking about just a few moments ago. From a completely mechanical point of view, given that u equals f of x and y , and v equals g of x and y , we very mechanically wrote down that du was $f_x dx + f_y dy$, dv was $g_x dx + g_y dy$, and then we just mechanically solved for dx in terms of du and dv .

My claim is, is that if we translate this thing, if we translate this thing into the language of du 's, dx 's, $du \tan$'s, $dx \tan$'s, et cetera, what we really said was what? That $du \tan$ was the partial of f with respect to x times dx , plus the partial of f with respect to y times dy . And $dv \tan$ was the partial of g with respect to x times dx , plus the partial of g with respect to y times dy . And then when we eliminated dy by multiplying the top equation by g_y and the bottom equation by $-f_y$ and adding, what we found was how to express dx -- and catch this, this is the key point-- what we did was we expressed dx as a linear combination-- not of du and dv , but of $du \tan$ and $dv \tan$.

You see, notice that the result that we needed to have to be able to use differentials was not this, but this. See, we found this, not $dx \tan$ equals $g_y du - f_y dv$ over $f_x - g_y$. To be able to say-- to invert this required that this was the expression that we had, yet the expression that we were really evaluating was this one.

In fact, let me come back for one moment, and make sure that we see this. You see, notice again that the subtlety of going from here to here and inverting never shows us that we've interchanged the roles of u and v from being the dependent variables to the independent variables. So the reason that there is so much work done in advanced textbooks under the heading of inverting systems of equations is to justify that being able to switch from dx to $dx \tan$ or from $du \tan$ to du as we see fit, whenever it serves our purposes.

The validity of being able to do that hinges on this more subtle type of proof, that as far as I'm concerned, goes beyond the scope of our text, other than for the fact that

in the learning exercises, I will find excuses to bring up all of the situations that bring out where the theory is important. In other words, there will not be proofs of these more difficult things. Not because the proofs aren't important, but from the point of view of what we're trying to do in our course, these proofs tend to obscure the main stream of things.

So what I will do in the learning exercises is bring up places that will show you why the theory is important, at which point, I will emphasize what the result of the theory is without belaboring and beleaguering you with the proofs of these things. At any rate, what I'd like to do now next time is to give you an example where all of the material or the blocks of material that we've done now on partial derivatives are sort of pulled together very nicely. But at any rate, we'll talk about that more next time. And until next time, goodbye.

Funding for the publication of this video was provided by the Gabriella and Paul Rosenbaum Foundation. Help OCW continue to provide free and open access to MIT courses by making a donation at ocw.mit.edu/donate.