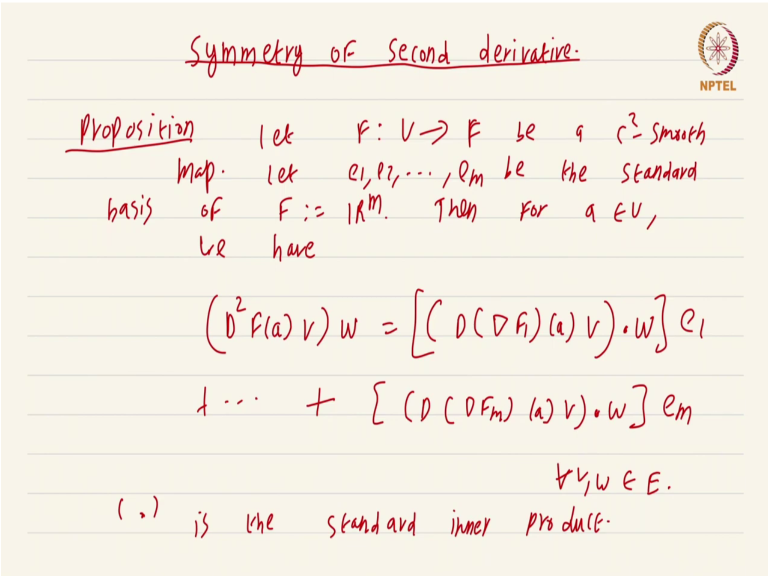
Real Analysis II
Prof. Jaikrishnan J
Department of Mathematics
Indian Institute of Technology, Palakkad

Lecture - 14.2
Symmetry of Second Derivative

(Refer Slide Time: 00:23)



In this video, we continue our study of the second derivative. We are going to show that when the map f is c 2 smooth, then the second derivative is actually going to be symmetric. Let me state the proposition that is going to form the central part of this video. Proposition, let F from U to c 2, sorry U to F be a C 2-smooth map; C 2-smooth map. Let e 1, e 2, dot dot dot e m be the standard basis; standard basis of F which is just the Euclidean space R m.

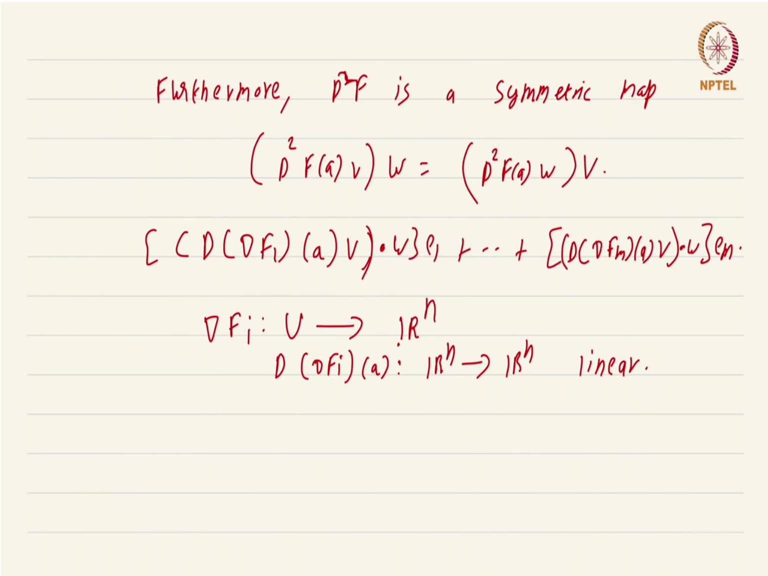
So, take a standard basis, take the standard basis for R m, then for a in U, we have an explicit expression for the second derivative.

Remember the second derivative acts on a vector v, this will produce a linear map from E to F and it acts on another vector w also coming from E and the net output of that is an vector in F and remember, e 1 to e m is the basis for F so, we are going to write down an explicit formula for this vector D 2 F a v acting on w in terms of the standard basis for R m and that expression is somewhat convoluted, but we will unpack this right after we finish stating the proposition.

This is D of the gradient of F 1 at the point a acting on v, this whole thing dot product w and this thing is the component of e 1 ok plus dot dot dot plus the final term will be D, the gradient of F m at the point a acting on v, this dot product with w e m ok and this is true for all v comma w in E.

It is a very convoluted explain expression or formula for the D 2 F a v w for the second derivative. Here, this dot is the standard inner product; is the standard inner product. So, this is this somewhat complicated looking expression for the second derivative.

(Refer Slide Time: 03:29)



And the final part of the theorem is furthermore, D 2 F is a symmetric map; is a symmetric map that is D 2 F a acting on v and that whole thing acting on w is the same as D 2 F a acting on w the whole thing acting on V. So, there are two parts to this proposition, the first part gives an explicit expression for the second derivative assuming the function is C 2-smooth of course, then the second part says that this D 2 F is a symmetric map which just means that if you act v and then at w, it is the same as acting w and then acting v.

So, let us unpack this convoluted expression ok, let me rewrite the expression here so that it is still there in front of our eyes. So, this is D of gradient of F 1 at the point a acting on v this dot product w e 1 plus dot dot dot the same thing D of gradient of F m at a acting on v dot product w e m so, this is the expression we have.

Now, first of all, note that gradient of F i is a map from U to R n right. So, it takes as input a point of the set U and its output is a vector in R n ok. So, it makes perfect sense to talk about the derivative of this map because this is starting from an open subset of an Euclidean space and landing up in an Euclidean space ok. So, it makes perfect sense to take the derivatives of these F i's, sorry of the gradient of F i's ok.

So, this expression D of gradient of F 1 of a is nothing but treating the gradient as a map from U to R n then taking the derivative. So, this map D grad F i would be a map from U to sorry it will not be a map from U, it will be a map from R n to R n that is e to e, remember e was always R n ok. So, this D gradient of F i is a linear map from R n to R n rather I should specify this what I have written is not wholly accurate, this is at the point a is a map from R n to R n linear ok.

And now, this acts on a vector v, that is this vector here to produce a vector in R n. So, this entire thing will output a vector in R n which we take the dot product. So, let me write the dot product in a more prominent way here also ok.

So, you can take the dot product with w which will give a scalar thankfully because I am putting a vector in front of it, there is a e 1 here ok. So, now, I hope it is clear, what is it that this proposition is asserting somewhat convoluted, but it needs to be done to show the symmetry of the second derivative ok.

(Refer Slide Time: 07:15)

Proof: $F = (f_1, \dots f_m)$. From an earlier prop.

$$DF(x)w = (Df_1(x)w)e_1 + \dots + (Df_m(x)w)e_m.$$

$Df_i : U \to \mathbb{R}^n$ is differentiable (why?!).

$$DF(a+v) - DF(a) \in L(E,F)$$

If $w \in E$ then

$$\{DF(a+v) - DF(a)\}w = \{(Df_1(a+v) - Df_1(a))w\}e_1 + \dots + \{\text{——}\}e_m.$$

So, the proof is going to be straightforward even though the statement looks complicated, we will just plow along straightforward lines and we will get the proof. Of course, we are going to write F as F 1 to F m; write F as F 1 to F m, we know from an earlier proposition; from an earlier proposition, we already know that D F x acting on a vector w is in terms of the basis e 1 to e m its nothing, but D F 1 x w e 1 plus dot dot dot D F m x w e m.

So, this we proved this earlier when we talked about the Jacobian matrix and all that somewhere around that time, I had stated what the derivative is going to be in coordinates. So, this is an expression for D F x w in terms of coordinates. Recall, F 1 is the component function so, it is a map from U to R so, D F 1 would be a linear functional so, D F 1 x w will indeed give you a real number so, this entire expression does make sense ok.

So, second thing I want to note down is that this was unsaid in the statement of the proposition, but I cannot leave it unsaid in the proof, gradient of F i which is a function from U to R n is differentiable and why this is differentiable? Because the function F is C 2-smooth so, I am going to leave it to you to verify the details, why is this function differentiable.

So, each gradient F i is actually differentiable, this is important because otherwise, this complicated expression on the right-hand side does not even make sense; does not even make sense so, if the functions gradient of F i's are not differentiable ok. So, the c 2-smoothness will immediately give the gradient of F i is differentiable.

Now, what is it that we have to do? We have to compute the second derivative, for that we have to take a difference of this type D F a plus v minus D F a ok. So, we have to take an expression of this type. Now, this is going to be an element of L E, F, this difference because both D F a plus v and D F a are linear mappings from E to F ok.

In fact, if w is in the vector space E, then this D F a plus v minus D F a; minus D F a acting on w; acting on w is given by; is given by what I had said earlier, this is just nothing, but D of F 1 a plus v minus D of F 1 a, this whole thing acting on w and this thing is the component of e 1 plus plus plus dot dot dot an analogous expression e m, I am not going to bore you by repeating the same thing again and again.

So, this just follows; this just follows from this, this just follows from this. So, we have an expression for the difference D F a plus v minus D F a acting on a vector w. So, this acting on a vector w is really not going to matter much as you will see in the later part of the proof ok.

(Refer Slide Time: 11:01)

Subtract the claimed expression for D^2F and consider the first term (corresponding to e_1).

$$\left[\left(DF_1(a+v) - DF_1(a) \right) w - D(DF_1)v \cdot w \right] e_1 .$$

$$= \left[\left(DF_1(a+v) - DF_1(a) \right) \cdot w - D(DF_1)v \cdot w \right] e_1$$

$$= \left[\left(DF_1(a+v) - DF_1(a) - D(DF_1)v \right) \cdot w \right] e_1 .$$

$$= \left[E(v) \cdot w \right] e_1 .$$

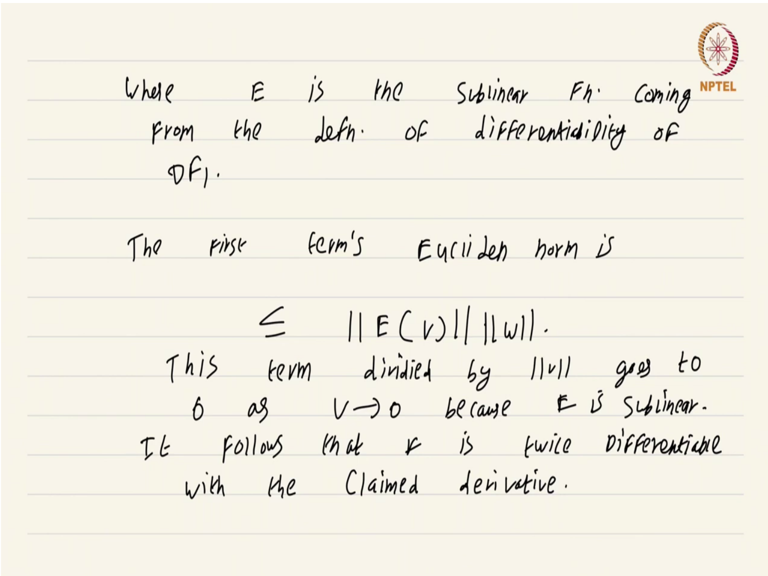
Now, what we are going to do is subtract the claimed expression; the claimed expression for D 2 F. So, we have this complicated expression in the statement for D 2 F a v w which is this complicated thing, I am going to subtract that to this ok and I am going to group the terms together in terms of the coordinate standard basis e 1 to e m and let us just consider the first term and the same analysis will hold for the rest of the terms.

So, subtract the claimed expression for D 2 F and consider just the first term; consider the first term that is corresponding to e 1; corresponding to e 1 ok. This is nothing, but this term D F 1 a plus v minus D F 1 a this whole thing acting on w minus D gradient of F 1 v dot product w and this whole thing e 1 ok.

Now, here is the part which is interesting what I am going to do is I am going to rewrite this expression because what follows involves taking the derivative of the gradient, it is natural to shift this entire thing in fact, including the W, it is natural to try to expand that in terms of the gradient. So, we can rewrite this as; we can rewrite this as gradient F 1 a plus v minus gradient F 1 a the whole thing dot product w minus D of gradient of F 1 v dot product w ok, this is starting to look really nice of course, I must write down the e 1 term ok.

This is starting to look exceptionally good because we can write this as gradient of F 1 of a plus v minus gradient of F 1 a minus D gradient of F 1 acting on v the whole thing dot product w and the whole thing is the component or coefficient or coordinate corresponding to e 1 excellent. Now, this is nothing, but E of v E of v dot product w; dot product w e 1 ok, this is just E of v dot product w e 1.

(Refer Slide Time: 14:13)



What is E of v where E is the sublinear function coming from the definition of differentiability of gradient of F 1. Remember, right at the beginning of the proof, I made the remark that gradient of F 1 is in fact, differentiable so, there is an x; there is a linear map and blah blah blah blah, there is an error function blah blah blah blah blah all that, I am just calling that error function that comes error term as E of v ok.

So, what is the net upshot of all this? The net upshot of all this is that the first term terms Euclidean norm; is less than or equal to norm of E of v; norm of E of v norm w product, this is just the Cauchy-Schwarz inequality, I have just used the Cauchy-Schwarz inequality here ok. So, and this term; divided by non v goes to 0; goes to 0 as v goes to 0 because E is sub linear; E is sub linear ok.

So, from this, I am going to leave the trivial details to you, it follows that; it follows that F is twice differentiable; with the claimed derivative; with the claimed derivative. So, this concludes the proof of the first part modulo some detail which I have cleverly left for you ok.

(Refer Slide Time: 16:33)

Symmetry: $D(\nabla F_1)$ is a real-symmetric matrix (why?).

$$\left(D(\nabla F_1)(a)\, v\right)\cdot w$$

$$= \left[D(\nabla F_1)(a) v\right]^t w = v^t \left(D(\nabla F_1(a))\right) w$$

$$= w^t \left(D(\nabla F_1(a)\, v\right) = \left(D(\nabla F_1)(a) w\right)\cdot v$$

as claimed.

Now, we have to show symmetry, that is the final part of this proposition. Well, observe that D of gradient of F 1 is a real symmetric matrix; real symmetric matrix; and I leave it you to ponder why this is so, this is again because F is c 2, because F is c 2 this D of gradient of F 1 is going to be a real symmetric matrix ok.

So, let us just look at again the first term, if I show that if you interchange v and w in the very first term, it is not going to make a difference, then it is not going to make a difference in all the other terms also with the same analogous argument. So, the first term is going to be D of gradient of F 1 of a acting on v dot product w; dot product w ok.

So, this first term and the second term are both vectors in R n, I am representing them as column vectors. So, to write down the dot product, I can also treat it in matrix notation, and I can just write this as D of gradient of F 1 of a acting on v, this will produce a column vector, I take the transpose and make it into a row vector and then, multiply by w ok.

So, of course, here I have made this identification, this left-hand side, this not left-hand side, top hand side is actually a number whereas, this expression since I am treating it as matrices, I am treating both w and this vector as matrices, this is actually a 1 cross 1 matrix, but it is harmless to identify a 1 cross 1 matrix with the corresponding entry ok.

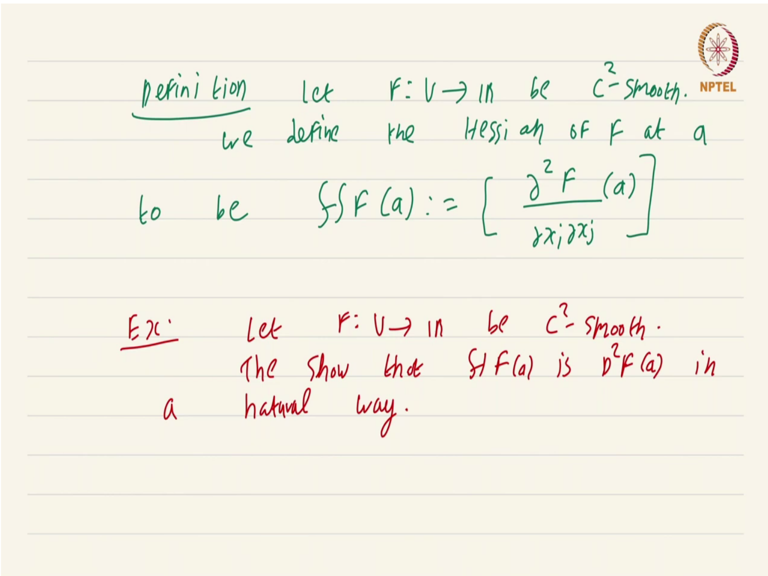
Now, what I am going to do is I am going to take the transpose of this whole thing which is going to give the same value because it is a 1 cross 1 matrix and the transpose; and the transpose as you can see is nothing, but, this is nothing, but v transpose, I am taking the transpose and writing equality v transpose D of gradient of F 1 ok.

This is not really taking the transfer this is just using; this is just using the formula for transpose of a product, I am treating everything, v here is going to be a column matrix, D this is a square matrix so, this is going to be just v transpose D of gradient F 1 a transpose, but that it is that is itself because I just said that D of gradient of F 1 is a real symmetric matrix and here, we have w ok, I have just taken the transpose ok.

Now, it is now at this point that I want to take another transpose, I want to take a transpose of this ok which is going to be itself because we are dealing with 1 cross 1 matrices. When you take the transpose, you get w transpose D of gradient of F 1 at the point a acting on v ok and this is same as D of gradient of F 1 at a w dot product v as claimed.

So, each of the coefficients corresponding to e 1 to e m would be symmetric if you interchange v and w nothing is going to change. Therefore, this whole D 2 F a is symmetric and the proof is complete ok. So, this proof will not be used in the rest of the course, but it is good to know, it is a bit abstract and a bit complicated because a lot of things have gotten compressed with the abstract language, but once you unwind everything and unpack everything, it is not that difficult.

(Refer Slide Time: 21:07)



So, one final definition and exercise so, definition this will make the second derivative somewhat more palatable and concrete. Let F from U to R be C 2-smooth; be C 2-smooth ok. We define; the Hessian H F of a to be by definition just the matrix of partial derivatives ok. So, I should probably write Hessian of F at a to be this.

So, this matrix is in fact, a symmetric because second partial derivatives are assumed to be continuous, we have in the C 2-smooth settings so, second partial derivative, the mixed partial derivatives are equal. So, this matrix is going to be a real symmetric matrix.

So, the exercise for you which will let you come and get your hands dirty on a concrete version of what has happened in this video, let F from U to R be C 2-smooth; C 2-smooth, then show that; H F is; H F a is D 2 F at a in a natural way. So, relate this concrete matrix of second partial derivatives with the abstract D 2 F, the bilinear map or the map that takes values in L E, F, it takes a vector E and maps it to a vector in L E, F, relate both of them and see that they are both related in a natural way.

So, in the analogous way, you can talk about higher order derivatives, there will be mappings from E to a very very complicated space alternatively, you can also view them as multilinear mappings, essentially there is no greater difficulty in the idea, it is just notation becomes somewhat unwieldy.

So, I will not pursue this any, further I have given references to several prominent textbooks that deal with this. So, I will conclude with just the second derivative part and the Hessian matrix. So, this concludes the our exploration of the second derivative. This is a course on Real Analysis, and you have just watched the video on the Symmetry of the Second Derivative.