

Numerical Methods in Civil Engineering
Prof. Arghya Deb
Department of Civil Engineering
Indian Institute of Technology, Kharagpur

Lecture - 28
Differential Operators – II

In lecture 28 of our series on numerical methods in civil engineering, we will continue with our discussion of differential operators which we introduced last time.


(Refer Slide Time: 00:35)

Forward Difference & Derivative

We have seen that the k^{th} order forward difference operator is related to the derivative: $\Delta^k f(x) = h^k f^{(k)}(\xi)$, $\xi \in [x, x+kh]$

Thus $h^{-k} \Delta^k f(x)$ is an approximation to $f^{(k)}(x)$ and the error of the approximation approaches zero as $h \rightarrow 0$ and the error is approximately proportional to h (error is linear in h)

We also saw that $h^{-k} \Delta^k f(x)$ is a much better approximation to $f^{(k)}(x + \frac{kh}{2})$ than to $f^{(k)}(x)$; the error in this case being quadratic in h i.e. $O(h^2)$



So, we last time we looked at the forward difference operator and found a relationship between the forward difference operator and the derivative and we found that, the k^{th} order forward difference operator is related to the k^{th} derivative of a function through this sort of a relationship and thus $h^{-k} \Delta^k f(x)$ is an approximation to $f^{(k)}(x)$. It is an approximation, why because this is evaluated at $x+kh$ and that is evaluated at x and the error of the approximation, which is 0 as $h \rightarrow 0$, because $x+kh$ must belong to the interval x to $x+kh$ as $h \rightarrow 0$ and the error is approximately proportional to h it is linear in h .

We also saw at the end of last lecture, that $h^{-k} \Delta^k f(x)$ is a much better approximation to the exact derivative evaluated at $x + \frac{kh}{2}$ than, it is to the derivative to the actual derivative at x , because the error in that case is being quadratic in h as you reduce h error goes as h^2 .

(Refer Slide Time: 01:51)

Other difference operators

Other difference operators include the central difference operator, the average difference operator and the backward difference operator.

The central difference operator, denoted by δ operates on $f(x)$ to yield $f(x + \frac{1}{2}h) - f(x - \frac{1}{2}h)$.

The average difference operator μ operating on $f(x)$ yields,

$$\mu f(x) = \frac{1}{2} [f(x + \frac{1}{2}h) + f(x - \frac{1}{2}h)]$$

The backward difference operator is defined by:

$$\nabla f(x) = f(x) - f(x-h)$$

It is clear that $\Delta f(x-h) = \nabla f(x)$, $\Delta f(x - \frac{1}{2}h) = \delta f(x)$. Similar relationships exist between higher order difference operators as

WATEL well.

And we also talked about other difference operators being the central difference operator, the backward difference operator and the average averaging operator and we define the central difference operator by delta operating on f of x is equal to this. The average difference operator mu operating on f of x is defined like this. The backward difference operator is defined like that and we can also establish relationships between each of these operators for instance. It is clear, that delta f x minus h is equal to the backward, forward difference operator operating on f x minus h is equal to the backward difference operator operating on f x and similarly, we can get similar relationships can be obtained not only for the first order difference operators but, also for higher order difference operators.

(Refer Slide Time: 02:52)

Relationships between operators

Commonly used relationships include the following:

$$\delta^2 f(a) = \Delta^2 f(a-h) = f(a+h) - 2f(a) + f(a-h)$$

$$\mu\delta f(a) = \frac{1}{2}[f(a+h) - f(a)] + \frac{1}{2}[f(a) - f(a-h)]$$

$$= \frac{1}{2}[f(a+h) - f(a-h)]$$

Similar to the relationship between the k^{th} order forward difference operator and the k^{th} order derivative i.e. between $\Delta^k f(x)$ and $f^{(k)}(x)$, there exist relationships between $\delta^k f(x)$ and $f^{(k)}(x)$:

$$f^{(k)}(a) = h^{-k} \delta^k f(a) + c_1 h^2 f^{(k+2)}(a) + c_2 h^4 f^{(k+4)}(a) + \dots \quad (++)$$

The right hand side includes only even derivatives of f . Also, the constants c_i are independent of f but depend on k

Some common relationships between higher order difference operators well, del square f as if you apply the central difference operator two times on the function evaluated at a, that is actually equal to del square f a minus h. It is equivalent to applying the forward difference operator two times on f at a minus h and it can be shown that, this equal to f a plus h minus twice f a f a minus h similarly, if I have a function evaluated at a and I apply successively on that function. The central difference operator, first I apply the central difference operator then, I apply the average difference operator then, I get something like this.

Now, similar to the relationships between the k'th order forward difference operator and the k'th order derivative which, we just saw in the previous slide, which was this 1 similar to that relationship. You can get relationships between the central difference operator and the derivative. The k'th order central difference operator and the k'th order derivative and the relationship actually is something like this, and It says that the k'th order derivative of a function is related to the k'th order central difference operator like this, plus these terms and you can see these are terms involving h square h four and so on, and they also involve the derivative of the function higher order derivative of derivatives of the function evaluated at a.

So, if I am interested in f prime of a and I want to find how my central difference approximation to f prime of a is related to f prime of a, I use this relationship, this tells

me that $f''(a)$ is approximately equal to the central difference operator acting on f scaled by the appropriate step size plus, the error terms and how are the errors, what are the error terms well. The error terms go as h^2 , h^4 and so on. The leading order of the error term is given by h^2 and it also involves the partial higher order derivatives of f . So, $f''(a) + \frac{h^2}{12} f^{(4)}(a) + \dots$. So, the hand side includes only even derivatives of f also the constants c_k are independent of f these constants, they do not depend on the function but, they do depend on the order of the derivative, that we are interested in these constants depend on the k .

(Refer Slide Time: 05:51)

Derivative in terms of the central difference operator

From (++) it is clear that for 2nd-degree or 3rd degree polynomials,

$$f''(a) = \frac{\delta^2 f(a)}{h^2} = \frac{f(a+h) - 2f(a) + f(a-h)}{h^2}$$

Thus in this case theoretically the second order central difference operator is exactly equal to the second derivative of the function irrespective of the step size.

However in practical applications of results such as (++) it is clear computations such as above will involve round-off errors and they will have significant effect on the accuracy of the higher differences and thus their ability to represent the exact derivatives

NPTTEL

They depend on the order of the derivative. So, from this previous expression, it is also clear that for second or third order polynomials. The derivative approximation, that I get using the central difference operator is going to be exactly equal to the actual derivative, why because these terms $f''(a) + \frac{h^2}{12} f^{(4)}(a) + \dots$ are going to be 0, these are going to be 0 these are going to be 0. So, it is clear that for second degree or third degree polynomials, $f''(a)$ is equal to $\frac{\delta^2 f(a)}{h^2}$ which we have seen to be this we have seen that.

So, it is clear that for second degree or third degree polynomials, the second derivative is exactly predicted by the central difference operator acting two times on the function. Thus, in this case theoretically the second central difference operator is exactly equal to

the second derivative of the function irrespective of step size, which is very important. So, whatever is my step size, if I find the derivative using this little formula here, that is going to be in exact, that is going to give me the exact second derivative of the function, and why do I say theoretically well, it is always theoretical whenever you do numerical approximations like this is always, you will say theoretical because of the possibility of round-off because. This $f(a+h)$ is going to be calculated using finite position arithmetic. We are going to miss some terms. So that is not going to be exactly equal to the analytical derivative $f''(a)$ but, within those round-off, within the limits of those round-off errors, this is the exact derivative for second degree or third degree polynomials.

In practical applications, that is what I said again in practical applications of results such that, it is clear that computations such as above will involve round-off errors and they will have significant effect on the accuracy of the higher differences and thus, their ability to represent the exact derivatives, is that clear? why do I say, why do the round-off errors become more significant for a higher order higher differences than for the lower differences well, you have to you can see that, when once you have h^2 , once we are interested in the second derivative, we have h^2 in the bottom typically h^2 is going to be smaller than h where, third derivative that will involve h^3 in the denominator, so the round-off. So, h^3 is going to be smaller than h^2 . So, the round-off whatever the round-off error is going to be more important for h^3 than it is for h^2 because, h^3 is much smaller than h^2 . Is that clear?

(Refer Slide Time: 09:20)

Estimating function values

It is apparent that using the difference operators to estimate the derivatives of a function requires knowledge of the function values at various points on a grid (x_0, x_1, \dots, x_n)

We denote by $\text{int}(x_0, x_1, \dots, x_n)$ the smallest interval containing the points (x_0, x_1, \dots, x_n) and by f_k, f'_k, f''_k the function value and the first and second derivative at grid point x_k

Suppose our knowledge of the function and its derivatives is incomplete i.e. we know the values of the function and possibly ~~some~~ of its derivatives at some points on the grid but we do ~~not know~~ the value of the function at say a grid point x_i

It is apparent, that using the difference operators to estimate the derivatives of a function requires knowledge of the function values at various grid points. So, I have to have my grid and now, we are looking at 1 D later on. We are going to generalize this into high dimensions to high dimensions but, even in 1 d you can see that in order to evaluate these derivatives, I must know the function values are those grid points.

The question is what if I do not know some of those functions values that are, what we are going to look at next. So, we denote by interval x_0, x_1 to x_n the smallest interval containing the points x_0, x_1 through x_n and by f_k, f'_k, f''_k , the function value and the first and second derivative at grid point x_k . So, this is the some nomenclature suppose our knowledge of the function and its derivatives are incomplete, that is why we know the values of the function and possibly some of its derivatives at some points on the grid but, we do not know the value of the function, say at some other point in the grid, say at point x_i and it turns out that for a certain difference formula, that I have got to use to approximate my derivative. I need the function value at x_i . So, what am I going to do? So, next we are going to discuss at a sort of systematic procedure to get those function values to at the unknown points at the unknown grid points.

(Refer Slide Time: 10:53)

Estimating function values

We would like to use the function and derivative values at the grid points, where they are known, to set up an approximation formula which would yield the function value at x_i .

This value should exactly match the function value at x_i obtained by fitting a polynomial of the highest possible order through the known grid point values (x_0, x_1, \dots, x_n) .

Since the polynomial of the highest order that can be completely defined using n grid point values (x_0, x_1, \dots, x_n) is of order $n-1$ this means that the approximation formula must yield the same value of the function at x_i as predicted by this $(n-1)^{\text{th}}$ order polynomial.

So, what we would like to do is we know certain values, we would like to use the function and derivative values at the grid points, where they are known to set up an approximation formula, which would yield a function value at x_i , where the function value is not known. Now, the criterion is that, this value should exactly match the function value at x_i obtained by fitting a polynomial of the highest possible order through the known grid points x_0, x_1 through x_n . So, I know the function values at x_0, x_1 through x_n . So, the best approximation I can get, if I can fit a polynomial through those points.

So, what I am saying is that I want to come up with an approximation formula for finding the function value at x_i but, criteria is that my approximation formula must give the same value at least, the same value as I would get if I fit a polynomial through those $n+1$ points. So, using $n+1$ points, I can fit the best highest order polynomial, I can fit through those points must be order of n . So that is must of order n . So, since the polynomial of highest order, that can be completely defined using n grid point values n , there is little error there. So that is actually but, let us suppose that, this is up to x_{n-1} . So, x_0, x_1 through x_{n-1} is of order $n-1$ this means that, the approximation formula must yield the same value of the function at x_i as predicted by this $(n-1)^{\text{th}}$ order polynomial. So, if I have the highest order polynomial, the value that it would predict at x_i must be given by my approximation formula.


(Refer Slide Time: 13:05)

Errors from the estimation

However any such approximating formula will have errors associated with it

These errors will include truncation errors. For example in our example just considered, if the function which actually gave rise to the known values at the grid points were of a higher order than $n-1$ e.g. of order n or higher, then the value we predict at x_i will be inexact

The magnitude of the error in this case will be given by the truncation error, with the leading order term in the error given by a term of order n

 NPTEL

Now, any such approximation formula will have errors associated with it. Why will it have errors? Well suppose to I have these function values at those points x_0, x_1 through x_{n-1} suppose, the function values are actually generated by a function which is of a higher order than my polynomial that a highest order of polynomial, which was n . So, suppose or rather $n-1$. So, if it is higher than that, if I actually generated the function values using a higher order polynomial then, the value that I am getting with my approximation formula is going to be in exact, because my approximation formula can only give the highest. It can only as accurate as the highest order polynomial, that can be fit through those points suppose, I generated values at those points using a higher order polynomial then, my approximation formula is not going to give me the value at x_i , because that x_i . The actual value is going to be predicted by that much higher order polynomial, is that clear?

So, these errors will include truncation errors for example, in our example just consider, if the function which actually give rise to the known values at the grid points were of a higher order than $n-1$. For example, of order n or higher then, the value we are going to predict at x_i will be inexact, because our value will only be as good as $n-1$. It will not predict, it cannot be exact if my value at x_i was actually predicted using a polynomial of a higher orders. The magnitude of the error in this case will be given by the truncation error with the leading order term in the error given by a term of order n , because that is the term which I am throwing away, that is the term which my

approximation formula cannot capture all lower order terms, my approximation formula can capture terms higher than that n minus 1, which is the higher first term, which I cannot capture is of order n . So, my error is of order n .

(Refer Slide Time: 15:28)

Errors from the estimation

Other sources of error include round off errors due to uncertainty in input data (known function and derivative values at the points (x_1, \dots, x_n)) as well as round off errors that occur during the computation.

One way to reduce the truncation error is to reduce spacing between grid points: the grid points become closer while their number is held constant. The truncation error will reduce with grid size h : in our example if the grid size reduces from h to $h/2$, the error will reduce by a factor of $1/2^n$.

However the work to compute the function values often increases sharply as the grid size becomes smaller. In addition, round off errors become more and more significant as grid size increases.

So, other sources of error include round off errors due to uncertainty in the input data for instance. I know only my function values up to the round off errors at the known points at the known grid points. I do not know the function values exactly. I only know them up to the round off error. So, those are the errors that will inevitable. So, one way to reduce the truncation error is to reduce the spacing between the grid points. So, if we make the grid points closer and closer while their number is held constant the truncation error is going to reduce with grid size as you make your grid finer and finer the truncation error is going to reduce with grid size in our example, if the grid size reduces from h to h by 2, the error will reduce by a factor of one-second n , because the error goes as n .

So, if I reduce the grid size from h by $h/2$, my error will be reduced by a factor of one-second. So, what am I saying is that, I keep my number of known points the same but, suppose that, at that locations at which those whose function values are known become closer and closer. So, my grid size becomes smaller and smaller, in that case my error is going to go down and my error. In this case is going to go down by one-second to the power n .

However, the work of work of computing function values often increases sharply as the grid size becomes smaller and smaller the work of the computation becomes more expensive, because I have to deal with smaller numbers in addition round off errors become more and more significant as grid size increases the round off errors become larger and larger.

(Refer Slide Time: 17:30)

Increase in round off with grid size

For example, if machine precision is 10^{-3} (say), then it makes no sense to reduce grid size h below 10^{-3} because any reduction in truncation error due to smaller grid size will be offset by the round off error being nearly as large as h

The solution is the Richardson extrapolation technique which enables calculation of the limiting value of a quantity as the grid size approaches zero, without actually having to reduce the grid size to zero

Suppose the current grid size is h and suppose we can evaluate function values at this grid size which denote as $F(h)$. We would like to know the limiting value $F(0)$ as the grid size approaches zero, without having to make the grid size extremely small

For example, if the machine precision is say 10 to the power minus 3 then, it makes no sense to make the grid size h below 10 to the power minus 3, because any reduction in truncation error due to the smaller grid size is about is because the round off error is going to be nearly as large as my grid size. So, it makes no sense to make the grid size smaller than the round off error then, whatever increase I am getting in reduction in truncation error because of the smaller grid size is going to be offset by the round off error, because it is going to be offset by the. So, if the machine precision is of the order of 10 to the power minus 3, it makes no sense to reduce the grid size below 10 to the power minus 3.

So, what do we do suppose I know I can only. So, I am saying that even if you make your grid size smaller and smaller, it is not going to be of much use because after a certain point your round off error is going to annihilate, whatever increase you increase in whatever reduction in truncation error, you achieve by reducing the grid size. So, what am I going to do, well it turns out that there is a method called the Richardson

extrapolation technique. The Richardson extrapolation technique which comes to array why does it come to array well basically we want to find the function we want to find get an estimate when the grid size actually become 0 we want to get an estimate of the function value at x_i at unknown point I want to find the function value at x_i when the grid size become 0 I know that when the grid size actually become 0 then my truncation error is going to go to 0 as well.

So, I want to find the value of the function at x_i , when the grid size goes to 0 but, I know that in actual practice if I actually go on trying to reduce the grid size to 0. I am not going to get anywhere, because at after a certain point my round of error is become going to become. So, large that it is not going to help. So, then I use this extrapolation technique. So, what do we do, well suppose the current grid size is h and suppose, we can evaluate the function values at this grid size, which we denote as F of h suppose for a grid size h . We can find out the function value at x_i , which I which I am going to call F of h . So, we would like to know the limiting value F of 0. So, I would like to know the function value at x_i for a grid size 0. So, for F of 0 as the grid size approaches 0 without having to make the grid size extremely small, without having to go through that pain and ultimately futile exercise, the painful an ultimately futile exercise of making the grid size extremely small In order to know.

(Refer Slide Time: 20:38)

Richardson extrapolation

In order to do this, we have to know how the function F behaves as $h \rightarrow 0$

Suppose we know that as $h \rightarrow 0$, $F(h) = a_0 + a_1 h^p + o(h^p)$ [$r > p$]


In the above $a_0 = F(0)$ is the known quantity of interest. We know p and r but have no idea of what a_0 or a_1 might be

According to Richardson extrapolation, we can calculate a_0 and a_1 if we know F for two step lengths h and qh ($q > 1$)

$$F(h) = a_0 + a_1 h^p + o(h^p) \quad (*)$$

$$F(qh) = a_0 + a_1 q^p h^p + o(h^p) \quad (**)$$

Hence from $(*) - (**)$: $a_1 = \frac{F(h) - F(qh)}{h^p(1 - q^p)}$

 NPTEL

So, we can use Richardson extrapolation provided I know, how the function behaves as h becomes smaller and smaller, suppose I know that as h goes to 0 F of h is given by this relationship F of h is equal to a_0 plus $a_1 h$ to the power p plus order h to the power r r greater than p , suppose I know that the behavior of the function as h goes becomes smaller and smaller is reasonably predicted by this equation by this equation with a h to the power p and the order of the error is of the order of h to the power r . So, basically I want to find out a_0 , because when h goes to 0 F of 0 is equal to a_0 . So, I want to find out a_0 . I do not know, what is a_0 is I have some idea of p n r but, I do not have any idea of what a_0 and a_1 . I know that as my h goes become smaller and smaller, the function value goes as h to the power p , the function value go to the changes as h to the power p . I have some idea of that factor p but, I do not know a_0 and a_1 .

So, according to Richardson extrapolation, we can calculate a_0 and a_1 , if we know F for two step lengths h and qh where q is greater than 1, how do we do that well we evaluate F of h for at h at step size h and we evaluate F at step size qh using these F of h and F of qh . I can solve this set of equations to find out what is my a_1 ? I can solve these two set of equations to find out my a_1 .

(Refer Slide Time: 22:43)


Richardson extrapolation

Substituting the value of a_1 in (*) we get:

$$a_0 = F(0) = F(h) + \frac{F(h) - F(qh)}{(q^p - 1)} + o(h^r)$$

If we know a more complete expansion of $F(h)$ in terms of powers of h , and unknown coefficients, then one can repeatedly use the Richardson extrapolation technique to get better and better estimates of $F(0)$

The repeated application of Richardson's extrapolation becomes simple when the step lengths form a geometric series $h_0, q^{-1}h_0, \dots$

 NPTEL

Once I found out my a_1 , I can find out my a_0 . Substituting the value of a_1 in this equation I can find out what is a_0 ? And once I know a_0 , I know what $F(0)$ is. So, it turns out that a_0 comes out as F of h plus F of h minus F of qh by q to the power p minus

1 with this error term. So, if I can find out the function value at h for step size h , if I can find out the function value for step size q of h and then I know that as the step size becomes smaller and smaller, I know how the function behaves. So, I know p then, I can find out what is the function value at for step size 0 .

So, this allows us to actually compute the function value at 0 , step size up to this error term, it allows me to compute that function value at 0 step size even if I without reducing the step size, making the step size infinitesimally small. So, if we know a more, we know only after the first term. The first term in the, in how if I know how F behaves near the 0 step size, only after the first term in that series but, if I know more terms in that series then, I can do this repeatedly. Let us see how we can do that.

If we know a more complete expansion of $f(h)$ in terms of powers of h and unknown coefficients, then one can repeatedly use the Richardson extrapolation technique to get better and better estimates of $f(0)$. So, repeated application of Richardson extrapolation becomes simple, when the step lengths form a geometric series h_0 by some factor h_0 by square of that factor and so on. Here, we have in this example. We have just use Richardson extrapolation once, because we only know the leading order term in the series, if I know more terms in the series. I can use Richardson extrapolation repeatedly and once I use it repeatedly. This error which was there of order of h^r , it is going to become smaller and smaller. So, I can know the value of $F(0)$ more exactly, if I use this repeated Richardson extrapolation but, in order to use that repeated Richardson extrapolation, I must know how $F(h)$ behaves up to I know the series expansion for higher orders of in terms of higher order of h also.

(Refer Slide Time: 25:47)

Repeated Richardson extrapolation

Suppose as $h \rightarrow 0$, $F(h)$ goes as the following:

$$F(h) = a_0 + a_1 h^{p_1} + a_2 h^{p_2} + \dots + a_k h^{p_k} + a_{k+1} h^{p_{k+1}} + \dots$$

where $p_1 < p_2 < p_3$

Then if we set $F_1(h) = F(h)$ and iterate to obtain:

$$F_{k+1}(h) = F_k(h) + \frac{F_k(h) - F_k(qh)}{q^{p_k} - 1}$$

then it can be shown that

$$F_n(h) = a_0 + a_n^{(n)} h^{p_n} + a_{n+1}^{(n)} h^{p_{n+1}} + \dots \quad (*)$$

This shows that as n the number of iterations increases, $F_n(h)$ comprises other than a_0 terms that involve higher and higher powers of h

Thus as n increases, $F_n(h)$ becomes a better and better estimate to $F(0)$

Suppose, we know that as h goes to 0 $F(h)$ goes like this $F(h)$ is equal to a_0 plus $a_1 h$ to the power p_1 plus $a_2 h$ to the power p_2 plus $a_k h$ to the power p_k and 1 with all these coefficients being increasing. So, $p_1 < p_2 < p_3$ is greater than p_1 , $p_2 < p_3$ is greater than p_2 and so on.

Then, if we set $F_1(h) = F(h)$ and iterate to obtain $F_2(h)$ is equal to $F_1(h)$ plus $F_1(h) - F_1(qh)$ divided by $q^{p_1} - 1$. So, here we use this little iteration formula and we iterate repeatedly and we start with saying that $F_1(h)$ is equal to $F(h)$ and then, we use that to calculate $F_2(h)$ is equal to $F_1(h)$ plus $F_1(h) - F_1(qh)$ divided by $q^{p_1} - 1$. So, that will give me $F_2(h)$ then, again I use that to calculate $F_3(h)$, $F_4(h)$ and so on. Then, it can be shown that $F_n(h)$ if I do this n times $F_n(h)$ the series for $F_n(h)$ is actually, given by a_0 plus $a_n^{(n)} h^{p_n}$ plus $a_{n+1}^{(n)} h^{p_{n+1}}$ plus \dots . This coefficient is most important h to the power p_n then, plus $a_{n+1}^{(n)} h^{p_{n+1}}$ and so on.

Now, let us go back and look at this $F(h)$ was given by this $F(h)$ was given by this and you can see the leading order term as is of power p_1 but, after I have iterated n times $F_n(h)$ after n times $F_n(h)$ the leading order term is of power p_n p_n is much higher than p_1 as we saw before, this shows that as n as the number of iterations increases $F_n(h)$ comprises other than a_0 terms, that involve higher and higher powers of h thus as I do as

this iteration repeatedly F_n of h becomes a better and better estimate to a 0 is equal to F of 0 , because these error terms become smaller and smaller, how can we prove this?

(Refer Slide Time: 28:22)

Proof for Richardson extrapolation

To prove this we recall that (*) holds for $n=1$. Suppose it holds for $n=k$. Then $F_{k+1}(h)$ can be calculated from $F_k(h)$ as above and it will involve the same powers of h as the expansion for $F_k(h)$

Next we note that in the expansion for $F_{k+1}(h)$ the coefficient of h^{p_k} is:

$$a_k^{(k)} + \frac{a_k^{(k)} - a_k^{(k)} q^{p_k}}{q^{p_k} - 1} = 0$$

This means that $F_{k+1}(h)$ will not contain any terms of order h^{p_k} and will contain terms of order $h^{p_{k+1}}$ or higher. This implies that the result holds for $n=k+1$, and this completes the proof by induction

NPTL

Well to prove this we recall that star meaning this. This we said holds for F_1 , if it holds for F_1 . Suppose it holds for n is equal to k then, we can calculate $F_{k+1} h$ from $F_k h$. We can use this formula. Let us suppose that it holds for F_k then, we can calculate $F_{k+1} h$ using that formula. And it will involve the same powers of h as the expansion of $F_k h$ why? Because $f_{k+1} h$ is directly computed from $F_k h$, so we can expect that $F_{k+1} h$ will involve the same powers of h as F_k . Now, let us see what is.

Let us know that in the expansion for $F_{k+1} h$, the coefficient of h to the power p_k . Let us find out what is the coefficient of h to the power p_k I know that F_k is the leading order term in F_k is given by h to the power p_k , Let us find out what is the coefficient in F_{k+1} of h to the power p_k . Well it turns out that, it is like this $F_{k+1} h$. So, that is a_k to the power p_k . So, a_k to the power a_k to the power k . h to the power p_k will have a_k in front. So, that will be a_k the coefficient minus $F_k q h a_k$ if I evaluate F_k with $q h$. It will be h to the power p_k times, q to the power p_k . Now, the coefficient will be $a_k q$ to the power p_k .

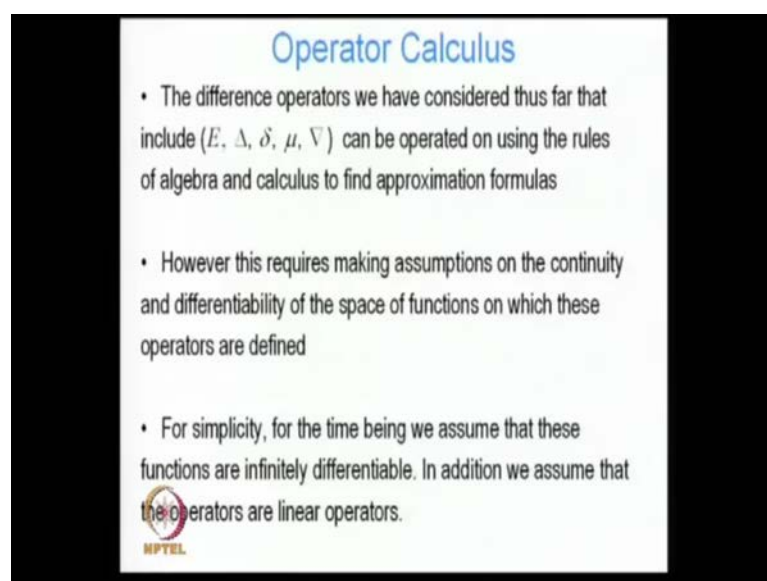
The coefficient of h to the power p_k , if I use the step size $q h$ is going to be h is going to be $a_k q$ to the power p_k . So, it will be a_k minus $a_k q$ to the power p_k divided by q to the power p_k minus 1 .

And of course, this term which is h^k , which is coming from that term, the coefficient of h to the power $p - k$ will be $\frac{k!}{(p-k)!}$. So, if we add these two, add all these terms together you will find that, this is becoming 0. This is becoming 0. So, what does this tell me this tells me that although I generated $F_{k+1}(h)$ using F_k I did not use anything else.

I just got $F_{k+1}(h)$ using F_k of h although, I generated F_{k+1} using F_k it turns out that when I go through this operation. I have got rid of the first term in the x series expansion for F_k . The next terms are there h^{p-k+1} is there but, the coefficient of h^{p-k} has become 0. So, in this expansion for F_{k+1} the leading order term is h^{p-k+1} it is no longer h to the power $p - k$. So, that means that every time I do this, I annihilate more and more terms.

This means that, $F_{k+1}(h)$ will not contain any terms of order h to the power $p - k$ and will contain terms of h to the power $p - k + 1$ or higher, this implies that the result holds for $n = k + 1$ and this completes the proof by induction. So, if we know that, it holds for F_1 . We assume that, it holds for F_k if it holds for F_k , it has to hold for F_{k+1} . So, since it holds for 1, it holds for 2, it holds for 3, and it holds for any n . So, we see that as we do this repeatedly, what is proof tells us is that if we apply the repeated, if I apply the Richardson extrapolation repeatedly then, I am going to get higher and higher I am going to get F_0 with greater and greater accuracy the error is going to reduce.

(Refer Slide Time: 32:54)



Operator Calculus

- The difference operators we have considered thus far that include $(E, \Delta, \delta, \mu, \nabla)$ can be operated on using the rules of algebra and calculus to find approximation formulas
- However this requires making assumptions on the continuity and differentiability of the space of functions on which these operators are defined
- For simplicity, for the time being we assume that these functions are infinitely differentiable. In addition we assume that the operators are linear operators.

NPTEL

So, why did we go into that, because we said that, we have to evaluate the function values at the grid points and suppose I do not know, if I have to evaluate my certain, if we have to evaluate a certain difference operator and I do not know the function value at a required grid point then, I can use this approximation method to find out the function value at the grid point and but, the approximation method is only going to be as good as the highest order polynomial that I can fit. So, there are inevitably going to be errors and those truncation errors are going to be of order higher than the highest order polynomial that I can fit but, I have this.

So, there are two ways of reducing those truncation errors 1 is to reduce the step size but, then I found that if I reduce the step size below certain value that is useless, because my round off errors are going to dominate and they are not going to give me any further advantage. So, then I talked about the Richardson extrapolation technique, which allows me to find out a more exact value of the function at x , even close to what I would get if I used a step size of 0, if I use the step size of 0 without encountering those round off errors, that is without actually evaluating the function value at for 0 step size. So, that was the idea that is brief recapitulation for we discussed.

So, now let us look at another way, another there is some very important well and very important way of obtaining difference formulas so up to till now. The difference formulas we have got, we just defined them but, there it can turn out that by using something called operator calculus. It is possible to come up with these approximation formulas for those derivatives. So, different operators we have, consider thus for can be operated on using the rules of algebra and calculus to find more approximation formulas however, this requires making assumption on the continuity on derivability of the space of functions on which these operators are defined for the sake of simplicity, we assume that. Our functions are infinitely differentiable for the sake of simplicity.

(Refer Slide Time: 35:35)

Operator Calculus

If the operators are linear and P is a generic difference operator and f and g are functions that belong to the domain of the operator i.e. possess appropriate differentiability properties, then:

$$P(\alpha f + \beta g) = \alpha P f + \beta P g \text{ where } \alpha \text{ and } \beta \text{ are scalars}$$

The operators are also considered to be commutative i.e.

$$(PQ)f = (QP)f = Q(Pf)$$

Two operators P and Q are equal if $Pf = Qf$ for all f

Next we consider an infinite series of operators. Expanding function $f(x+h)$ in Taylor series about x we get:

$$f(x+h) = f(x) + hf'(x) + \frac{h^2}{2!} f''(x) + \frac{h^3}{3!} f'''(x) + \dots$$

NPTEL

And we also assume that our operators are linear, if the operators are linear what does it mean? it means that, if P is a generic difference operator anything like Δ or whatever P is any generic difference operator and f and g are functions that belong to the domain of the operator, that is they possess appropriate differentiability properties then, P operating on αf plus βg is equal to αP operating on f plus βP operating on g where, α and β are scalars. We also assume that, the operators are commutative that is, if I operate on f with first P first Q and then P .

That is the same thing as if I operate on f with first P and then, I operated on the result with Q and then, I make a further statement which is that, when can I say that two operators are equal well. I can say that two operators are equal if I operate, if they give me the same result for all functions f . So, if I operate with P on a space of functions the space of functions include all possible functions with the appropriate differentiability properties and I find that the result is equal to what I would get. If I operate with the operator Q on that same space of functions then, I can say that my operators P and Q are the same.

So, I know the equalities terms from the result. So, P is equal to Q , if the result that I get if I operate with P on all the functions belonging to the function space of interest is the same as I would get if I operate with Q on that same function space. So, let us look at some examples of the use of operator calculus in order to do that, let us look at our

Taylor series of $f(x+h)$. So, $f(x+h)$ is equal to that you know from Taylor series. And this, I can rewrite this in terms of the shift and the differential operator. How well what do you have on the left insight? you have $f(x+h)$. So, I can say that as the shift operator operating on f and the left hand side, I have the differential operator.

So, I can write this as E of f of x plus 1 plus h times, the differential operator plus hD operating two times on f of x , which I denote as $h^2 D^2$ then, I have $h^3 D^3$ operating three times on f of x which I denote as $h^3 D^3$. So, you see $h^3 D^3$ cube f of x . So, that I am denoting as $h^3 D^3$ operating three times on f of x . So, $h^3 D^3$ to the power cube divided by factorial three operating on f of x .

Thus it is clear. Now, look at this series, what is that series? that is an exponential series exponent or to the raise to the power hD . So that, thus it is clear, that E is equal to hD . So, operate a calculus allows us to get this relationship between the shift operator and the differential operator. The shift operator can be expressed as the exponential of the differential operator scaled by the step size recall that we showed earlier that, the shift operator is actually 1 plus, the forward difference operator we showed that. So, this means that, E raise to the power hD must also be equal to 1 plus Δ we take log of both sides \log of E to the power hD is equal to $hD \log$ of 1 plus Δ , I can expand it in series \log of 1 plus Δ if I expanding series I get Δ minus Δ^2 by two plus Δ^3 by 3 and therefore, hD operating on f of x .

Now, I am look at looking at this operating on f of x is equal to $h f'$ of x by definition, because D is nothing but, the differential operator is equal to $h f'$ of x is equal to this operating on f of x . So, what does this tell me? This tells me the operator calculus, tells me that the exact derivative is actually equal to this series operating on f of x , this series which involves the forward difference operator operating on f of x . So, not only is does it involve Δf of x it involves $\Delta^2 f$ of x , it involves $\Delta^3 f$ of x and so on. So that is sort of, because you in order to a estimate. The first derivative is tells me that, you must take like infinite differences, which tells me got then, how is that going to work. If I have to take infinite differences in order to estimate my first derivative sole thing is unworkable.

(Refer Slide Time: 41:25)

Use on polynomials

The above example shows how operator calculus can allow us to establish relationships between various operators, in this case the D and Δ operators.

Let us consider a function $f(x) = e^{\alpha x}$ which satisfies our requirement of being infinitely differentiable. From (*) we get:

$$hDe^{\alpha x} = \left(\Delta - \frac{\Delta^2}{2!} + \frac{\Delta^3}{3!} - \dots\right)e^{\alpha x} \quad (**)$$

It can be shown that (**) holds for all polynomials. But recall our earlier result: $\Delta^k f(x) = h^k f^{(k)}(\xi)$, $\xi \in [x, x+kh]$.

If f is a polynomial of degree ' n ' we will have only n terms on the right hand side of (**) because terms involving Δ^{n+1} or higher will be equal to zero, since $f^{(n+1)}$ and higher order derivatives are zero.

But, we must remember that, these differences go like this h to the power k f of k . This differences $\Delta^k f(x)$ goes as h to the power k times f derivative of f . So, if as n goes so, it becomes smaller and smaller.

(Refer Slide Time: 41:53)

An important relation

Writing this in terms of the E (shift) and D (differential) operators we get:

$$Ef(x) = \left(1 + hD + \frac{(hD)^2}{2!} + \frac{(hD)^3}{3!} + \dots\right)f(x)$$

Thus it is clear that $E = e^{hD}$ i.e. the shift operator can be expressed as the exponential of the differential operator scaled by the step size

Recall that we showed earlier that $E = 1 + \Delta$. Hence $e^{hD} = 1 + \Delta$

Taking log of both sides, we get:

$$hD = \ln(1 + \Delta) = \Delta - \frac{\Delta^2}{2} + \frac{\Delta^3}{3} - \dots$$

Therefore, $hD[f(x)] = hf'(x) = \left(\Delta - \frac{\Delta^2}{2} + \frac{\Delta^3}{3} - \dots\right)f(x) \quad (**)$

The higher contribution of these higher order differences is smaller, because they are proportional to h to the power k . So, for my step size, if suppose I have my step size 0.1. So, and then the contribution of Δ^n is going to be 0.1, raise to the power n 0.1 raise to the power 10. So that is going to be something like 10 to the

power minus 11. So, eventually those higher order terms are not going to contribute much. So, I can throw out those, I can ignore those contributions.

(Refer Slide Time: 42:32)

Use on polynomials

The above example shows how operator calculus can allow us to establish relationships between various operators, in this case the D and Δ operators.

Let us consider a function $f(x) = e^{\alpha x}$ which satisfies our requirement of being infinitely differentiable. From (*) we get:

$$hDe^{\alpha x} = \left(\Delta \cdot \frac{\Delta^2}{2!} + \frac{\Delta^3}{3!} \dots \right) e^{\alpha x} \quad (**)$$

It can be shown that (**) holds for all polynomials. But recall our earlier result: $\Delta^k f(x) = h^k f^{(k)}(\xi)$, $\xi \in [x, x+kh]$.

If f is a polynomial of degree ' n ' we will have only n terms on the right hand side of (**) because terms involving Δ^{n+1} or higher will be equal to zero, since $f^{(n+1)}$ and higher order derivatives are zero.

So, the above example shows, how operate a calculus can allow us to establish relationships between various operators in this case, the D and the delta operator. So, for an example, let us consider the function $e^{\alpha x}$ is equal to $e^{\alpha x}$, why do we consider it? Because we assume that, the function has to be infinitely differentiable. So that $e^{\alpha x}$ is infinitely differentiable. So, from this, we get $hDe^{\alpha x}$ is equal to Δ^2 by factorial 2 plus Δ^3 by factorial 3 and so on times $e^{\alpha x}$. It can be shown that, this holds for all polynomials.

Recall our earlier result $\Delta^k f(x)$ is equal to that, if f is a polynomial of degree n . We are going to have only n terms on the hand side of star. So that is one saving. So, if the function f is a polynomial of order n . So, the higher order differences are going to disappear, because higher order differences are proportional to the derivative higher order polynomials after a certain order of $n+1$ derivatives, $n+1$ in a higher polynomial of order $n+1$ in a higher derivatives are going to be 0. So, those things are going to disappear. So, this differences also going to get truncated but, so that is 1 saving.

So, for a higher order polynomial, we know that after a certain order, the differences are going to be by definition 0 but, again even the higher differences that does not means, that we have to calculate differences up to order n and absolutely not, because those differences are going to be proportional difference of order n is going to proportional to h to the power n. So, it is contribution for a reasonable choice of step size is going to be negligibly small. It is going to be probably less than round off. So, we can make sensible choices of how many terms we need to take. So, if f is a polynomial of degree n, we will have only n terms in the hand side of star, because terms involving delta n plus 1 or higher will be equal to 0, since f n plus 1 in higher order derivatives are 0.

(Refer Slide Time: 45:05)

Relation between D and δ

Thus (*) gives a formula for the exact numerical differentiation of a n^{th} order polynomial involving the operator Δ

We can similarly establish relationships between the derivative operator and the central difference operator. Recall:

$$\delta f(x) = f\left(x + \frac{h}{2}\right) - f\left(x - \frac{h}{2}\right) = (e^{hD/2} - e^{-hD/2})f(x) = \frac{\sinh(hD/2)}{2} f(x)$$

Hence $\delta = 2 \sinh \frac{hD}{2}$ since two operators are equal if their action on any function is equal.

Expanding x in terms of a series in $\sinh x$:

$$x = \sinh^{-1} x = \frac{1}{2} \sinh^3 x + \frac{1.3}{2.4} \sinh^5 x + \frac{1.3.5}{2.4.6} \sinh^7 x + \dots$$

NPTL

Thus star gives the formula for the exact numerical difference differentiation of n'th order polynomial involving the operator delta, it is exact but, it is we are never going to use it exactly, because those higher order differences are going to give me smaller and smaller contributions. We can similarly, establish relationship between the derivative operator and the central difference operators. We got relationship between the derivative operator and the forward difference operator we can similarly, establish relationships between the derivative operator and the central difference operator how well let us see. We know that del f of x is equal to f of x plus h by two minus f of x minus h by 2, this I can think of as a forward difference operation.

Remember, I know sorry not the forward difference operation and my apology. That is the shift operation I know there is a relationship between the shift operator and the exponential. I know, there is a relationship between the shift and the differential operator using the exponential function. So, in this case if I look at this again $f(x+h)$, I can think of that as a shift operator and now the shift operators has step size h , my shift operator with step size h and I know that my shift operator with step size h is related to the difference operator by e to the power hD . So, my shift operator with step size h is related to my differential operator by e to the power hD .

So, this I can write as $f(x+h)$, I can write as $e^{hD} f(x)$. I can write as $e^{-hD} f(x)$ and this operating on $f(x)$ this operating on $f(x)$, what is this? That is nothing but, my sine hyperbolic D , sine hyperbolic D divided by 2, sine hyperbolic D divided by 2. So, that times $f(x)$, thus what do we get. So, operate a calculus tells me that, the operation of the differential of the central difference operator on $f(x)$ is equal to sine hyperbolic h sine hyperbolic of the differential operator by 2 divided by 2 operating on $f(x)$. So, hence δ is equal to $2 \sinh(hD/2)$ since, the two operators are equal if their action on a function gives me the same result. So, this central difference operator operating on $f(x)$ is equal to this operator, operating on $f(x)$. So, I can say that my central difference operator is equal to this operator, because we said before that, the equality of the operators means, if the equality of the results on the same function.

If I operate with the with 2 different operators on the same function and the results are the same then, I will going to say that my operators are going to be the same and since, I said that, I must be true for any arbitrary function. My $f(x)$ is any arbitrary function, which satisfies the appropriate differentiability requirements. So, we get this expression δ is equal to that and then, why if I go back. So, then I take a step back and I say that if I can always expand x in terms of a series in sine hyperbolic x . I am not going to show that, you can see some books on good book, good mathematics books on infinite series, you will see that, you can always expand x in terms of a series in sine hyperbolic x .

(Refer Slide Time: 49:41)

Relation between D and δ


But $\sin \frac{hD}{2} = \frac{\delta}{2}$. Hence

$$\frac{hD}{2} = \frac{\delta}{2} - \frac{1}{2} \left(\frac{\delta}{2}\right)^3 + \frac{1.3}{2.4} \left(\frac{\delta}{2}\right)^5 - \frac{1.3.5}{2.4.6} \left(\frac{\delta}{2}\right)^7 + \dots$$

Hence finally :

$$\frac{hD}{2} = \delta - \frac{\delta^3}{24} + \frac{3\delta^5}{640} - \frac{5\delta^7}{7168} + \dots (***)$$

This relation is similar to (***) for the forward difference operator :
 it gives an expression for the exact differential operator in terms
 of the central difference operator and its higher orders



So, since we can do that but, we know that sine hyperbolic D by 2 is equal to half of my difference operator. So, if I expand h D by 2 h D by two, I can expand h D by 2.

(Refer Slide Time: 45:05)

If it treat x as h D by two, I can get h D by 2 in terms of a series in sine hyperbolic h D by 2 sine hyperbolic h D by 2 but, sine hyperbolic h D by 2 is equal to delta by 2, sine hyperbolic h D by 2 is equal to delta by 2. So, I can get in the series for h D by 2 in terms of delta rather delta by 2.

(Refer Slide Time: 49:41)

So that is exactly this series. So, h D by 2 is equal to delta by 2 minus half delta by 2 cube by t2 and so on. Hence finally, we get a series like this h D is equal to delta minus delta cube and so on. So, this gives me a very similar relation similar series to the series I got between the forward difference operator between the forward difference operator and my differential operator, the series that I got between my forward difference operator and difference operator. So, there I got h D in terms of delta minus delta square by2 and so on and so forth.

Here, I am getting h D in terms of the central difference operator that was in terms of the forward difference operator. This is in terms of the central difference operator gives an expression for the exact differential operator in terms of the central difference operator

and it is higher order. So, here by using this operator calculus, we are getting these series expressions series for my exact differential in terms of my difference operators but, this should not scare you, because you should realize that, these higher order difference operators are going to have negligible contribution.

Their contribution is going to scaled with the step size. So, it is not that if you have to find out the exact differential. I have to find out all these terms in the series absolutely, not it only, we only need to compute some terms in that, see only some terms in that series are going to contribute significantly meaning. It will not be lost in round off only some terms in the series are going to give significant digits, which are not going to get lost in round off but, this gives some insight, it is not maybe. It is not that practical, nobody is actually going to use that series but, give some insight into how they behave. Similarly, by squaring this series twice, we can get a expression for the differential operator 2 times where, if I square it we get an expression for D square which is my second derivative, my differential operator acting 2 times.

(Refer Slide Time: 53:07)

Relation between D and δ

By squaring (***) twice, we can get a similar expression equivalent to application of the differential operator twice, and by extension n times:

$$(hD)^2 = \delta^2 - \frac{\delta^4}{12} + \frac{\delta^6}{90} - \frac{\delta^8}{560} + \dots$$

Sometimes direct application of (***) may present certain difficulties. That is because the leading order term in the expansion of hD in terms of the δ operator is δ

Recall that $\delta^2 = \frac{1}{2}[f(x+\frac{h}{2}) - f(x-\frac{h}{2})]$. Thus evaluating the derivative in terms of δ requires us to know the function values at mid-points of the intervals. If the function values at the mid points of the intervals are not known, that adds to the complexity

NPTTEL

So, I can get h D square similarly, analogously I can get a series in terms of the central difference operator for my n'th derivative as well n'th derivative as well there are some problems with direct applications for the particularly for the central difference operator there might be some problems why is that, because the central difference operator allows us requires us to know the function values at mid step suppose I only know the function

values at $h/2$, $h/3$, $h/4$ if I have to use the central difference operator. I have to know the function values at $h/2$, $h/3$, $h/4$ and so on and so forth, so using the central difference operator sometimes present difficulties. So, next class, before moving on to something and it is larger things. We will talk about how we can use a combination of the average difference and central difference operator to get around this problem.

Thank you.