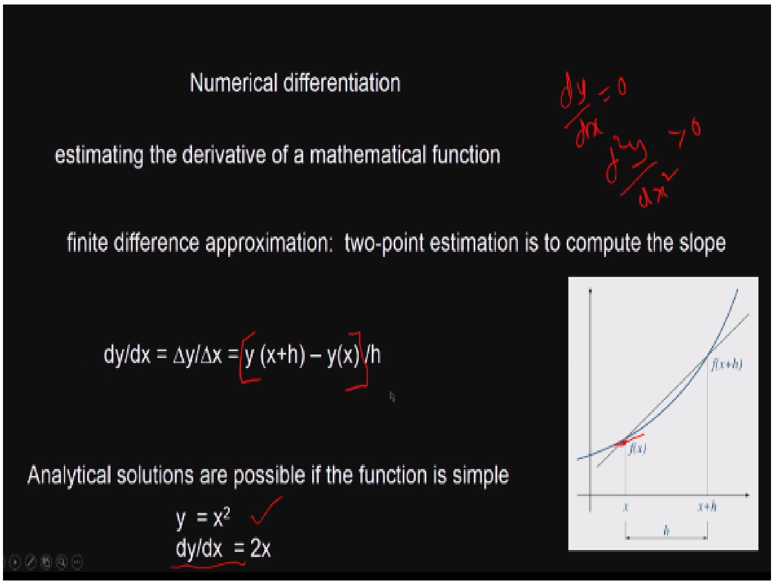
Computer Aided Drug Design
Prof. Mukesh Doble
Department of Biotechnology
Indian Institute of Technology – Madras

Lecture - 20
ODES and Numerical methods

Hello everyone, welcome to the course on computer aided drug design. We will talk little bit on numerical methods, differential equations predominantly only differential equations. I am not going to spend too much time on these numerical methods because that itself can be a separate course on its own. But we will come across numerical methods in drug discovery.

(Refer Slide Time: 00:46)



Because like I mentioned if you are going to look at minima then we need to think that dy/dx should be = 0 and d square y/d x square should be > 0. So, dy/dx is calculated, okay which we looked at in the previous class, okay. So, we need to understand how the computer is use numerical techniques to calculate that. If it is in analytical solution, then it is quite simple. For example, if you have y=x square, so dy/dx = 2x.

We must have studied in our school, okay. But then imagine a function which has got very complicated independent variables in the numerator and the denominator. So, it is not possible to do analytically and for computers they have to calculate this dy/dx, okay using numerical techniques. So, like in our Force Field method we have the distance or the bond length, bond angle, bond torsion.

So, we have 3 independent variables which needs to be, okay adjusted so that the total energy which is on the left hand side becomes minimum. So, we will have many df/dx, df/dy, df/dz we calculated. So, if it is analytical as I said it is easy like y=x square means we can calculate dy/dx = 2x. Otherwise, we need to do it numerically. So, estimating the derivative of mathematical function that is called finite different approximation.

So, the simplest one is a 2-point estimation to compute the slope at the point dy/dx as you know is approximated to delta y/delta x, okay which is equal to yx+h – yx/h, actually this is it will come out like that. So, h is the small distance, so you estimate our function y at x+h then estimate the function at y x. So, it takes the difference divided by h, here. As you can see here so, if this is your function going like that, okay.

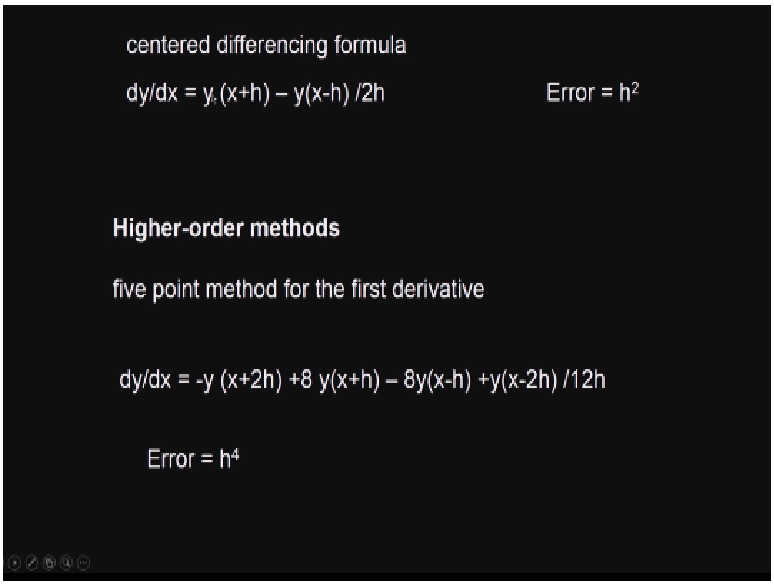
So, I want to calculate the slope at this place. So, the approximation here is you calculate the function at x point x you calculate the function at x+h and then you take the difference divided by h, actually that is a slope. This distance divided by this distance the approximated to the slope here, as you can see if h is smaller and smaller this becomes the exact slope but if x is larger and larger you see the difference here, okay. So, that is the error that comes in.

So, we need to keep a very small h, so that we get very good approximation for dy/dx at this point x, okay at this point x. But if you are going to keep a very small h you are going to too much of computation, so the computation effort also increases. So, this is the simplest method, so there are more slightly more involved method which does not require too much calculation but still can get a better approximation of the slope here.

So, as you can see here you are approximating the slope at this place, okay by taking the function at x+h and calculating the function at fx subtracting and then dividing by the distance h, okay. This is called the finite difference method. This is called a forward method. That is also a backward method but this is called the forward method. So, like I said if you have the analytical solution possible then it is most accurate.

Otherwise, in most of the cases it is not possible to have analytical solution, so we use numerical method for estimating the dy/dx, okay.

(Refer Slide Time: 05:01)



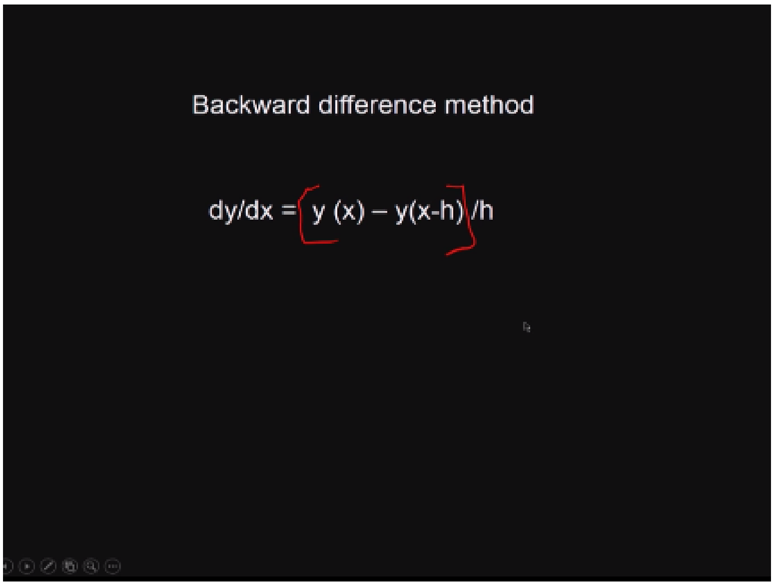
There is another method that is called centered difference formula. That means you do x+h, x-h and then divide by 2h, okay. So, this is x+h and then we have the x we have the x+h here we have the x here and suppose we have x-h here, okay. So, you calculate at x-h, so this difference will be 2h, okay, so this is at a function of xh and this is at a function at x-h and the difference is 2h.

So, you do the same calculation dy/dx = y (x+h) – y (x-h)/2h. This is called the center difference. This is slightly more accurate than the forward difference. Whereas in the forward difference we take this point and this point values subtract by take the difference / h, whereas in the center difference we take here and here and then divided by 2h, okay. So, this is like taking an average.

Then there are higher order methods like I said these methods like I showed you the forward difference, center difference have error of the order of h square. So, if h is small error will be corresponding h square. So, if I divide h one tenth then error also will become one hundredth of that. Higher order methods there is one method which is called 5-point method, as you can see here, okay this is called 5-point method, okay.

So, you calculate the function at 2h-2h, h and –h and then use this formula and then / 12h to calculate the derivative, first derivative, okay. So, you calculate the function at 2h-2h then again h-h, okay and then this has an error of the order of h power 4, okay.

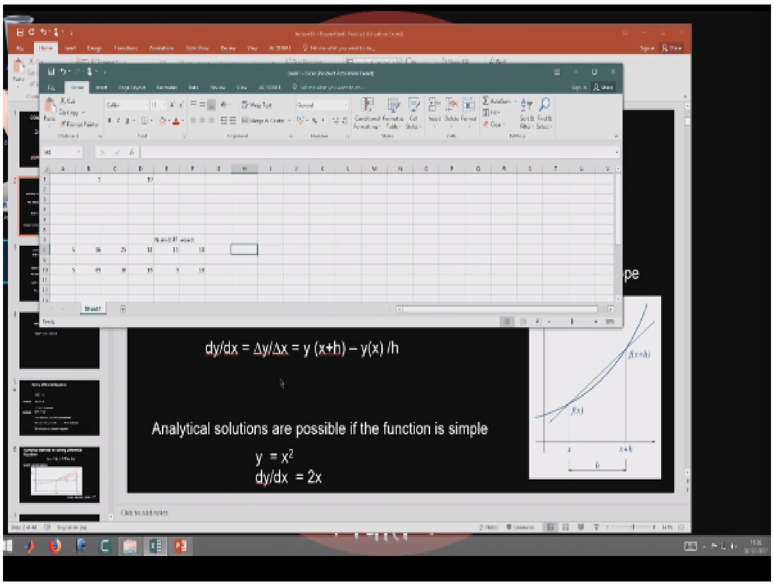
(Refer Slide Time: 07:16)



So, there is another method that is called the backward difference method. So, just like we take in the first case x+h and x-h, we take x and x-h, okay. So, this is called backward difference method as you can see here we take x+h and x here we take x and x-h. So, that is why this is called the backward difference method, okay. This also has the error of the order of x square just like the forward difference method, okay.

So, all these methods have errors in the order of x square and these higher order methods were 5 point methods as an error order of h power 4. So, software's generally use the higher order methods instead of reducing the size of h higher order methods can be more accurate and give very good estimation of the error values, okay.

(Refer Slide Time: 08:23)



So, let us look at this, for example I have y = x square dy/dx = 2x, okay. So, suppose I want to find out at x = 5, okay what do I do is in the forward method what will I do? In the forward method I will calculate y at x+h suppose I take h = 0.1. Okay, we take x = 0.01 at 5y = x square, so we can do it as a, okay, so this is at x+h because it is x square we calculate **8 09:30**+ B1 that is your h and this is at xh, so that is = 8 h to the power 2, okay.

Then what we do is we take a difference the difference is this - this is the difference and then the / h, okay. So, it comes to 10.01 but whereas if you take y = x square, okay and dy/dx = 2x, so at x = 5 the answer will come out to be 10 that is the exact value. This is the exact value, okay this is numerical differences. Do you understand this? Okay this is the exact value because at 5 dy/dx = 2x, so I put 2 *5 that comes to 10, okay.

If I do numerical differentiation what do I do? I will take the y value at 5 + 0.01 – y value at 5 and then / h. Here I have taken h as 0.01, so y at 5 + 0.01 is 25.10001 y at 5 = 25 take the difference here and then I / the h that is 0.01. So, numerical differentiation gives me 10.01 and whereas the exact distance. So, you see there is a small error of 0.01. If I increase the step length 0.02 it becomes 10.02 the error has increased.

So, if I increase it to 0.1 for example, so the error has become, see as you can see the numerical differentiation gives 10.1 whereas the exact value does not change, it is 10. So, you can see the numerical gives almost 10% error, okay. So, if I put 01 you get 1% error. So, the step length is 0.01 I get 1% error, the step length is 0.1 I get 10% error. So, if I make it still smaller, okay I get almost 0.1% error.

So, depending upon the step length the error keeps coming down, okay as you can see. So, if I keep 0.01 I get 1% error, 0.1 I get 10% error 0.001 it gets 0.1% error, okay for this type of function y = x square. So, because x square it is going to raise in a quadratic fashion, so the dy/dx can change dramatically, okay. Again it will not be constant all the time because if I take see at x= 10 I want to find out, okay.

So, it can change depending upon which point I am trying to estimate the dy/dx. So, as you can see the error depends upon the step length h, okay. You can have the 1% error; you can have 10% error larger the step length you can differentiate things faster. So, the computational time is less but the error will keep adding and adding and adding which is very dangerous.

Smaller the step length the error will be very less but you are going to have more computational. See, this is error is only 0.1% even that could be higher, okay. So, ideally I may go into 0.0005, okay so the error is 0.05%, okay. So, maybe that is okay or 0.0001, okay so the error is okay 0.01%. So, but look at the step length, so 0.301, so if I want to calculate dy/dx over a distance of 0 to 10.

That means for example if I want to calculate distance of 0 to 10 then number of calculations will be almost 1,00,000. So, lot of calculations, okay. So, the step length is very small the number of calculations will also increase, okay. But if step length is small your error also will be less. For example, if you look at this 0.0001 my error is 0.01% in my different dy/dx whereas, if I go to 0.01 y error becomes 1% which is too much.

You have to have about 0.1% error is very good to have but your step length also will keep, I mean number of calculations also keep increasing. So, that is the main point which you need to note when you are doing numerical differentiation or for example even numerical integration the step length determines the error and the step length also determines the number of calculations we have to do.

So, there are methods which use reasonably small or reasonably big step length but more accurate calculations. For example, if you look at the 5-point method for the first derivative, okay where x+2h+ this and so on. So, the y is calculated 5 different places, okay. So, the error this is in the order of h power 4, okay. We can try that also 0.01 as it says, okay. So, we say 5 it is calculated as x+2h.

So, x square is your problem, so we can calculate + 2*h, okay then you calculate at x+h that is open bracket x+. So, we have at x+2h then x+h then x-h then x-2h-2*. So, this comes to = - this + this / open bracket 12 * square coming into all these places which I missed out, yes I missed out the square term that is why, okay. So, you calculate at x+2h then you calculate x+h then you calculate x-h then you calculate x-2h, okay then use this formula x+2h that is minus of +8 * x+h – 8 * x-h -2h/12h.

So, we got 10, whereas if use a simple forward differential the simple method you get 10.01 whereas the 5-point method gives you 10x which is exactly matching with your answer, okay. So, if I make it as a point 1 for example if I make the step length very large, okay as you can see again the simple method gives you, 2 point methods give 10% error, whereas the 5-point method is still very good, okay.

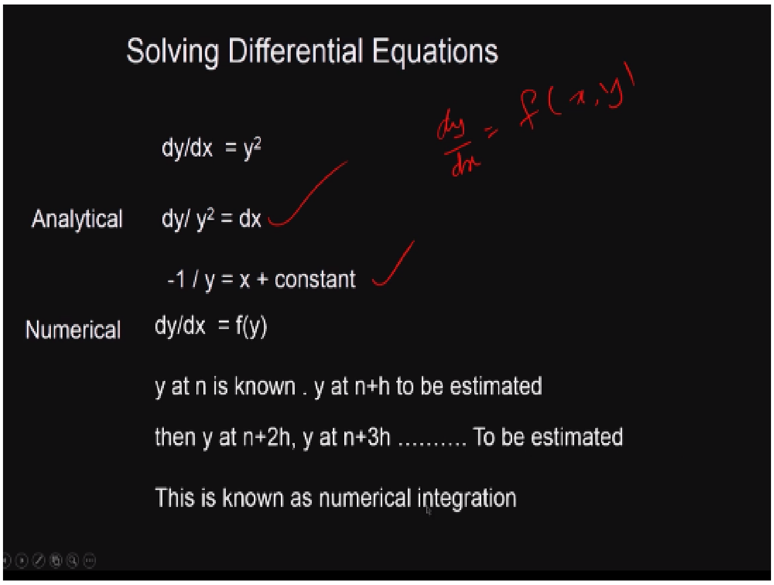
So, you still get very good answer with the 5-point method when compare to the one-point method, okay. So, step lengths can be included, okay with respect to the 5-point method that is the beauty of these types of methods. So, we can improve on the error but at the same time have reasonably large h value and the step length, so that we get reasonably good dy/dx.

So, this numerical differentiation is very important which is the backbone in many of the conformational search algorithms, when you start with the original confirmation and then you want to reach the confirmation, so that the energy is minimum. So, in order to achieve that like I said dy/dx has to be = 0 and d square y by dx square must be > 0. So, in order to calculate this dy/dx = 0 many differential equations have to be arrived at.

So, generally if the numerical methods are followed where use this type of this is a first order point estimation dy/dx = y at x+h – y at x/h, okay divided by h. And you have the centered difference that means you find out from x+h and x-h/2h or we have a 5-point method which we follow that means we take at x+2h, x+h, x-h, x-2h and then divide by 12h, okay the errors are very little here in these type of method or there is a backward difference.

But again this also has the error of the order of h square, okay. So, as I showed you that 5-point method like are always good which are better approximation of the difference getting the differentiation done using numerical, okay maintaining your h reasonably large, okay. These are needed for aiding the confirmation of minimum energy confirmation, okay where we use these type of dy/dx = 0 requirement, okay.

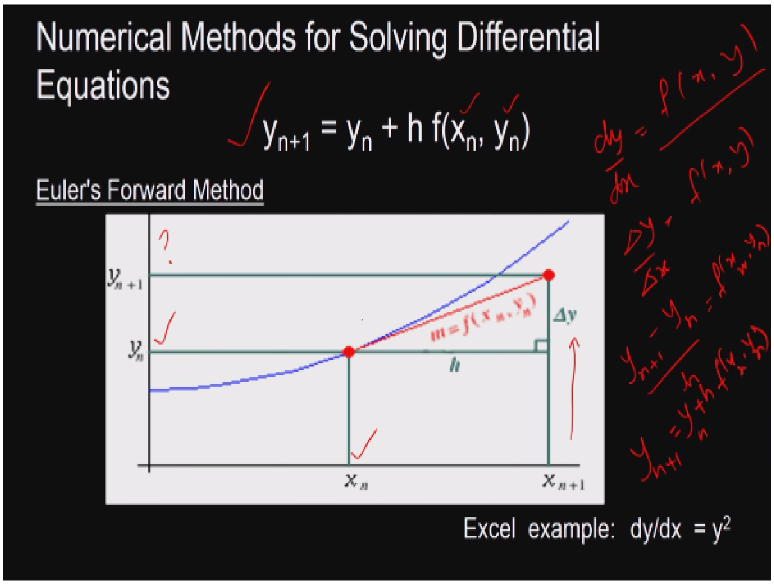
(Refer Slide Time: 22:49)



So, similarly we also have numerical integration also, okay. For example, if dy/dx = y square analytically what we do we take this here take that here take that here dy/y square = dx integrating both sides – 1/y = x+ constant, okay. So, this we must have done in our, okay finally your school or even in your engineering first year and so on actually, okay. So, if it is analytical so life is easy.

But then if you have very complicated equation here, okay then that is not possible. So, you may have dy/dx= even function of x and y very complicated equation. Then we have to solve this numerically only and there are many methods numerical methods that are available which does this type of calculation, okay. So, we want to find out y at different points at point n+h, n+2h, n+3h like that we need to this is known as a numerical integration.

(Refer Slide Time: 23:56)



So, there are many methods simples and very difficult methods. For example, the easiest one is called the Euler's Forward method. So, I know the value of y at xn, okay I know the value of y or I know yn at xn, okay like this, this one I know. Now, I want to find out at xn+1 what is the value of y that is the question, I am trying to find out. So, there are many methods are there and the simplest is the Euler's Forward method.

So, what we do? That is whatever we did numerical differentiation here we are doing the numerical integration. So, we take the h here again h is the distance or the step length, okay we calculate the function at xn, yn here. Okay, we calculate the function at this place, okay at this place and then we get the f (xn, yn), okay. Then that is the slope like here and then multiplied by h and then we say that is yn+1, okay.

This of course straight forward right. So, we have the dy/dx = f of xy, okay. So, what we do? We take delta y/delta x = f (xy), okay. So, delta y is nothing but yn+1 –yn+1 and delta x can be h f of xn * yn. So, yn+1 = yn+h * f (xn, yn), this is what is here given here, understand. So, we start from this approximation this is simplest method.

So, if I know the y value at xn value that is I know yn at xn I want to find out what is yn+1 at xn+1, okay this is my step length from xn to xn+1 this is my step length. So, what do I do? I calculate the function that is the slope here, okay because that this function is nothing but the slope here dy/dx. So, this is the reverse of the previous numerical differentiation and then I multiplied by h here. h is my distance here, okay, so I get the yn+1.

So, as again you can see this called Euler's method and again as you can see here if h is small I am going to get a nice y value. If the h is very large I may end up here where as the actual value could be here because if you have a graph like this turning upwards like this and then you would take this as a linear obviously there is a difference. So, h has to be small. So, either I can put small h or I can use different methods.

There are some very good method even with large H it can give a very good approximation for yn+1, okay from the slope here, okay. So, basically in this method you have the slope here, you have the slope here and then we multiply by the h here and then add it with the xn to get the yn+1 value, okay that is called the Euler's method. Euler's method is the simplest of all these methods.

So, these 2 numerical integration and differentiation are very important and we will continue more on this topic and there are simple methods which requires very small h to have very considerably low error value whereas there are other higher order methods which can also take in larger h but still the errors can be smaller, like I showed you in the numerical differentiation here again when we do numerical integration also.

As you can see the h is large with Euler's method you can have so much error coming into picture but there are other methods which can address this type of problem in a better way. So, we will continue more on these numerical techniques in the next class also. But as I said numerical methods is a big course of its own, I cannot cover the entire. So, I am just going to have only very faster brief introduction to this particular topic for you, okay. So, we will continue more on this. Thank you very much.