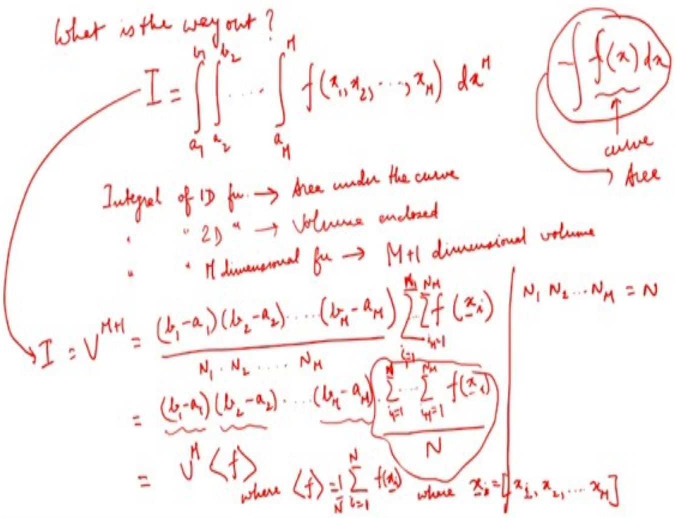
Computational Physics
Dr. Apratim Chatterji
Dr. Prasenjit Ghosh
Department of Physics
Indian Institute of Science Education and Research, Pune

Lecture - 08
Numerical Integration Part 03

What we saw so far is that, if we go from 1 dimension to 2 dimensional integral to a multiple dimensional integral so, two things are happening; one is the error in the integration becomes bad that is the quality ,where the error increases the quality of the integral becomes bad. And secondly it becomes prohibitively expensive to do these calculations.

(Refer Slide Time: 00:43)



So, what is the way out? So, this is the question we are asking. So, let us see what in general principle when we talk about a multidimensional integral what does one mean? So, what we are trying to do is, now we are going to discuss methods which will help us solve or rather get estimates of this type of integrals. So, I have m dimensional integral b 2 and the number of dimensions goes to m f x 1, x 2 all the way to x m and then I have x m dimensional.

So, this is the integrand and what I have is along each dimensions the interval is given by a m b m a for example, along the first dimension it is a 1 to b 1 second dimension a 2 to b 2 and so on and so forth. Now what geometrically what does we mean? So, we have seen that, for example if I have a 1 dimensional function say I have a f x a function which is f x basically 1 dimension. If I integrate it so, if this is a 1 dimensional function; that means, it is a curve and if I integrate it the whole integrand will give me a area.

So, what I can write is integral of 1 D function gives me area under the curve. So, integral of a 2 D function gives me the volume enclosed there. Now, suppose I have a M dimensional function; so this I can generalize to the fact that volume of M dimensional integral of a M dimensional function gives me M plus 1 dimensional volume.

So, what I can do is; I can write this integrand this I here in the following fashion. So, it is nothing, but a 'M plus 1' dimensional volume. And if we try to discretize it so, how will we calculate the volume enclosed by this particular function M dimensional function? So, basically what we need to do is if we remember the 1 dimensional trapezoidal rule we took the length along the x dimension but here we have M dimensions. So, we have to take the lengths along each of the M dimensions. So, that is basically b 1 minus a 1 into b 2 minus a 2 so on and so forth.

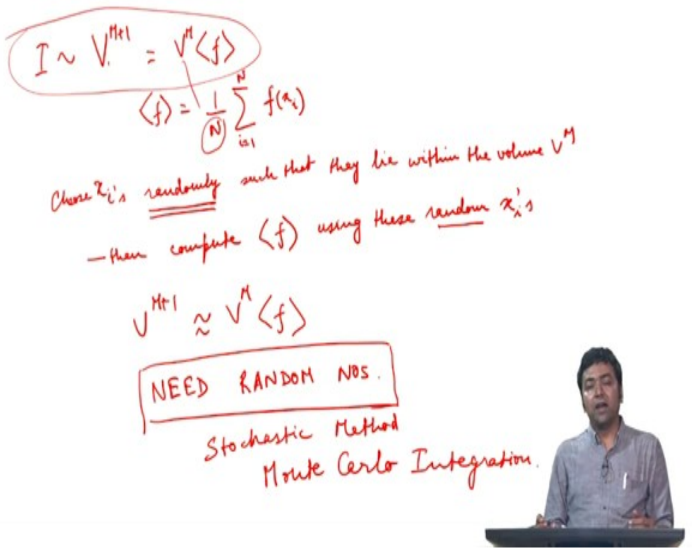
And then it goes to b M minus a M and then we need to calculate the average value of this function. So, basically f(xi) which is a vector and then we have a couple of summations. So, i1 goes 1 to M, 1 to N, N 1 i 2 and similarly i2 and then i M goes from 1 to N M and this we divide by N 1 into N 2 and so forth till N M.

So, now if we write that the product of N 1 N 2 and N M, if I write this as capital N without any suffix; so what it means is basically I can rewrite this equation as b 1 minus a 1 into b 2 minus a 2 all the way to b M minus a M. And then what I can also do is I can, I rewrite this in the following that the remaining part I can rewrite in the following way; it is a sum over i 1 equals to 1 to M all the way to i M equals to 1 to sorry this will be N 1 to N M f x i by N.

So, if I look into these terms the products along of the length along each dimension. So, this is nothing, but my volume enclosed by these sides b 1 minus a 1, b 2 minus a 2 and b 3 minus a 3. So, this I write as V M here and if I look into this term here so, this is nothing, but the average value of the function. So, that I write as f.

So, where my where f is now I am removing all the individual summations and I am putting one summation f x i into 1 by N where, x i is again where my x; where x is a 1 dimensional vector consisting of i components where i goes from 1 x 1 x 2 all the way to x M.

(Refer Slide Time: 07:08)



So, basically what this means is that my integral I am approximating it by calculating M plus 1 dimensional volume whose value is given by the volume M into the average value of my function in that volume. If we would have done it in the grid based method; so what I would have done is? That this average I would have computed on a predetermined set of grid points, which are given by N and then I sum over x i the value of the; I evaluate the value of the function at each of the grid points and do it. But this, but going about this way for a multidimensional case as we made the case before that it will be very expensive, I mean it will be prohibitively expensive I would say.

So, what people suggested is that instead of taking a nice ordered fixed set of grid points, why do not we choose the grid points randomly such, but such a way that these points lie within this these N points lie within this volume. So, the idea is choose i‘s randomly, such that they lie within the volume V m and then compute the average value of the function using again these random x i’s and then you plug this in this equation and you will get a estimate of the integral.

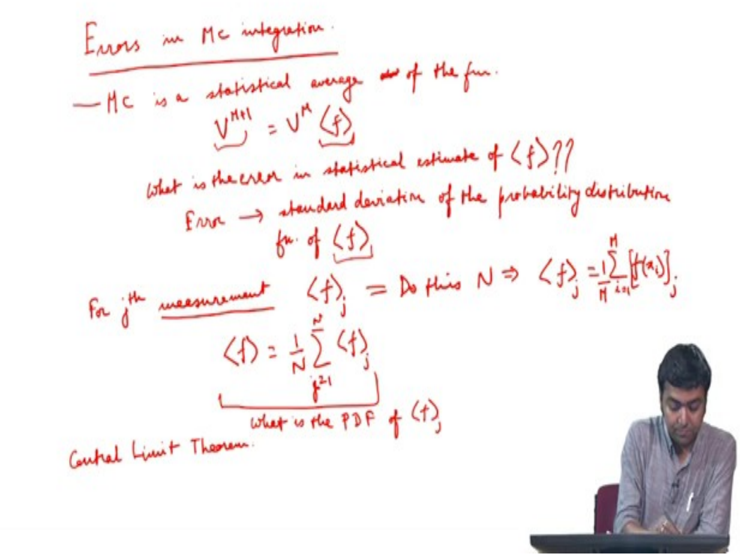
So, what we have to summarize ,this is I have my V M minus 1 is now approximately equals to V M and then the average one. So, to do this what we need is now, random numbers. Now since we are using here random numbers to determine the grid points that we will be using to compute the average of the function and so we do not know a priori which are the grid points that we are going to use. So, we only know that these grid points will lie within the volume enclosed by the sides, but we do not know where specifically these grid points will be.

So, what it means is that, now things have become sort of randomized and not deterministic as before as we had seen for example, in the trapezoidal rule. So, hence these type of integration method is called stochastic method and this integration method is known as Monte Carlo integration. So, this is a very simple idea about the very simple concept behind Monte Carlo integration.

So, what we do is suppose you are given a multi dimensional function; you generate set of numbers which will define your grid point in that multidimensional volume, these numbers you generate randomly. So, basically you choose your grid points randomly and once you have done that then you can compute the function at each of these grid points and then you can compute the average of the function and then you can estimate the integral.

So, basically what it means is that, what we are doing is we are finding using a statistical way to average the integrals. So, now, the moment you do anything numerically or when you do any experiment or when you do some calculations numerically. So, one question always comes up is what is the error bar of my results? How accurate are my results?

(Refer Slide Time: 12:12)



So, the next few minutes what we are going to do is, we are going to discuss how errors in my estimate of the integral using Monte Carlo? So, usually when one tries to make estimates in the error. So, what one needs to know apriori is the actual value of the integral and then compare how far my estimated value of the integral is from the actual value. But in case of Monte Carlo integrations I mean for the practical purposes where it is used; unfortunately we do not know or we do not have a estimate of the actual value of the integral.

So, in order to have an estimate of the error in the computed value, what we need to do is we need to do or use statistical ways to or estimate the error or in other words what is the statistical error in this method? So, just to summarize and to move forward, so Monte Carlo integration as we saw is a statistical involves a statistical average over average of the function ok. So, I have V M plus 1 equals to V M and then error. So, the error in the estimation of V M that lies in the error in the estimation of the statistical average of f.

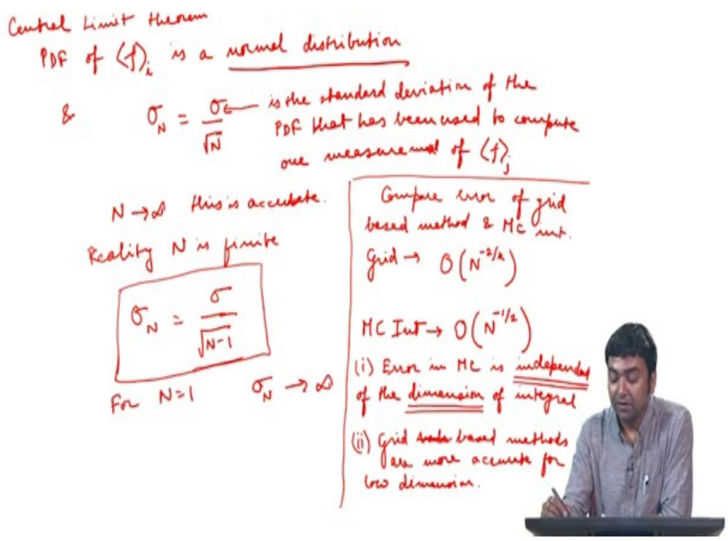
So, from the knowledge of statistics what so, what we need to know is, we need to find out what is the error in statistical estimate of f? So, basically you have to find out the answer to this question. Now, in general so when we are doing statistical estimation or computing the average value of f statistically. So, the error depends on the standard deviation of the probability distribution function of the average value.

So, suppose so, what it means is the following. Suppose you have computed say you start with a set of random numbers and then you have computed your average value of f; then you start you use another completely different set of random numbers you have computed your average value of f. Like that you do several measurements using a large number of different series of random numbers and then you plot the distribution of this quantity f. So, mathematically the way you will represent it is in the following way.

So, suppose for j th measurement, so when I say measurement, remember what I am talking is I use a set each time for each j th measurement; I use a set of random numbers which are completely different from say j minus 1th measurement and I compute the average value of the integral. So, this is what I mean for the j th measurement of the f. And then and say if I do this measurement do this for example, N number of times; so what it means is, so my f j is basically equal to say 1 by M i equals to 1 to M f x i and this is the for the j th measurement.

So, if I do that then my and if I do such measurements N times. So, the average value of the average of the function that will be given by 1 by N sum over j equals to 1 to N f j. So, this is what will be given by this. Now the question is what is the probability distribution function of this f j's? So, to find the answer to this what we will resolve to is the Central Limit theorem.

(Refer Slide Time: 18:03)



So, what Central Limit theorem states that I will not going I am not going to prove, but I will just use the statement of central limit theorem. I will just invoke basically central limit theorem. So, what it tells that is that the Probability Distribution Function of by average values of f of the values of the average of f taken from n independent measurements is a normal distribution.

So, this is what I am going to use. And as I mentioned before here in the slide that the error is basically the standard deviation of the probability distribution function of f. So, what it means is that my sigma for this case will be nothing, but the standard deviation of my normal distribution function which with it is generating. So, where this sigma without the suffix is the standard deviation of the probability distribution function that has been used to compute one measurement of f j.

So, what it other words tells us is that, if I compute the average value of the function; so if I compute it for a large number of times this goes my error will go down. So, in the limit of N tends to infinity this is accurate, my estimate is accurate, but in reality what we have is N is finite. So, this expression is modified by sigma in this fashion N minus 1.

So, one might ask the question why do we have the minus 1? So, the minus 1 comes here to stress on the fact that if you just do a single measurement then your error will be a single measurement which means that; so for N equals to 1, so my sigma N goes to infinity. So, what did other word other way tells that for a single measurement it is never possible to get a reasonably accurate estimate of my integrand ok.

So, far what we have done is we have seen in a brief way the concept of, how one does multiple integration in multiple dimension using by Monte Carlo integration method which is a stochastic method and what we also have seen is that how the error scales? So, the error scales as a function of square root of N. So, now, if we compare the scaling of the error with that of, grid based methods and MC integration.

So, to remind you the Grid based method the error scales as order N to the power minus 2 by k. While for Monte Carlo integration we see that the error scales as N to the power minus half. So, what is the major difference? So, what can we conclude or what can we learn from these two?

So, the first thing which we learn is that error in MC is independent of the dimension of the integral. So, important thing to remember, this is this does not depend on the dimension of the integral; if you have a 2 dimension or a 3 dimension or a infinite dimension the error the quality of your integral will remain same at least mathematically that is what it means.

So, but if you another thing if you look, but in contrast for the grid based method as I mentioned earlier and I am stressing it once again as we go to higher and higher dimension integrals as you try to do higher and higher dimensional integrals your the error becomes larger. And this also tells us second thing is that grid based methods are more accurate for low dimension. For example, if you consider k equals to 1 that is if you do a 1 dimensional integral; so then what you will what this tells me is that my error is much smaller for the grid based method compared to this Monte Carlo integration ok.

So far so good so, we have looked into, we have learned how to estimate the error also in the Monte Carlo integration. The next thing which we need to know to do this type of integrations is, so how to generate random numbers? Or rather know how to generate random numbers, but how to use how to get a set or set of random numbers and how to use them, because the whole method is based on this random number generation.