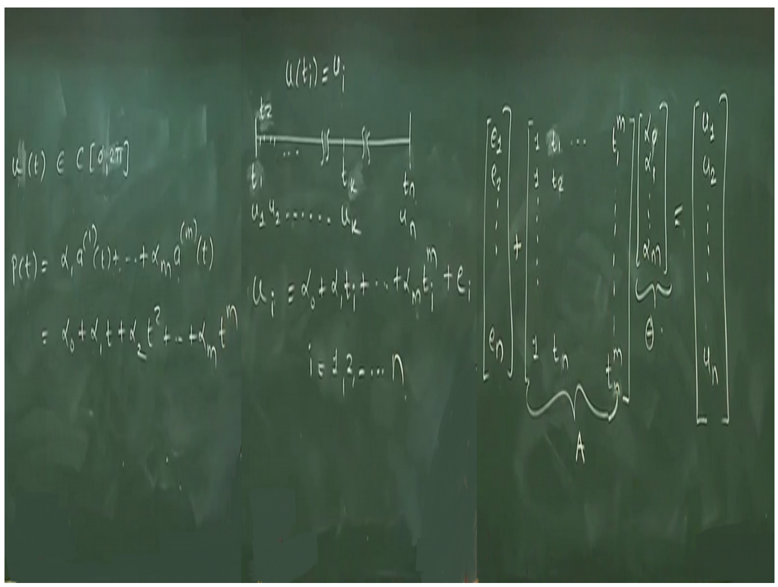
Advanced Numerical Analysis
Prof. Sachin Patwardhan
Department of Chemical Engineering
Indian Institute of Technology - Bombay

Lecture - 22
Discretization of ODE-BVP using Least Square Approximation

So in my last lecture, I discussed about Hilbert matrix and approximating a function using a polynomial function mth order polynomial function, and we saw that we get this Hilbert matrix which for large order polynomial is highly ill-conditioned difficult to invert, well the problem is not different if I just passingly mentioned.

(Refer Slide Time: 00:56)



See yesterday we saw that we had this function f t which belong to set of continuous functions, and then maybe we took some specific thing here 0 to 2 pi, and then u t which is function u t which is this, and then we had this approximation. And then we wrote the normal equation and said that particularly when you pick this to be a polynomial that is when you pick this to be alpha 0+alpha 1 t+ alpha 2 t square, for this particular case I showed that you end up into this Hilbert matrix.

And this Hilbert matrix is trouble inverting, so in principle do you get an equation you cannot solve it properly because of ill-conditioned, now ill-conditioning of a matrix will formally defined a little later, right now all that I will tell you is that the eigenvalues of this matrix or singular values of this matrix to be very precise, singular values or eigenvalues of A transpose A, eigenvalues singular values of this matrix the smallest one and the largest one are far apart and that causes problems in computations.

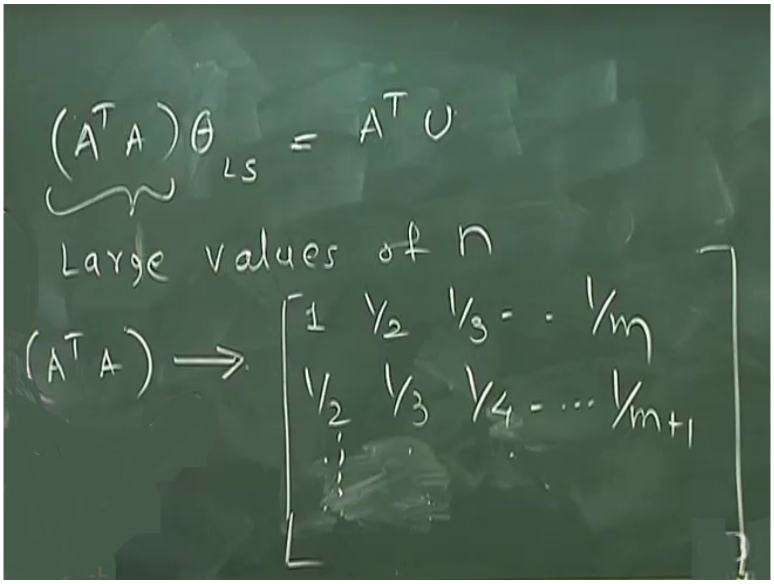
Now this is when you are trying to approximate the entire function, suppose you know this function instead of knowing the entire function what it is and are trying to approximate using a polynomial, suppose you know this function values you do not know the function, but that is more practical problem in engineering. You just know function values at different points, so t1 this is t2 and so on so tk, and in general you know value of this function at several points.

So these values will be u1 u2 uk un okay, and then you know at each point you will write this equation that is approximation or u t i=alpha 0 +alpha 1 t i +up to alpha m t i to power m, you get this equations u t i is nothing but u i, so u t i u i this +error i okay, and i going from 1, 2, up to n. You have these equations we have done this earlier okay, and we get this matrix A matrix, so we will get this huge A matrix.

So this will be 1 1, t1 t2 up to tn, up to t1 raise to m tn raise to m, and alpha 1 alpha 0 alpha m is =u1 u2 un, in this case you have this situation you are trying to fit a lower order polynomial say 10th order polynomial, and this n here number of data points is let say 1000. So you have tried to put the 10th order of polynomial in 1000 points okay, you have lot you know this function value at large number of points, simple example CP versus temperature okay.

I know CP values at many, many temperatures in some range, I am trying to fit I am trying to correlate, I will get something like this, there t will be temperature in that case, and then u1, u2 will be CP values, and of course you have this error coming up here, so this error 1 error 2 error n this vector will add here. So this is my A matrix okay so A matrix.

(Refer Slide Time: 06:30)



And we have solved this problem, we know that a solution the least square solution is given by the least square solution that is alpha least square is A transpose A or I will write it in the little different way, so this is A transpose A alpha least square, well I think we named this as theta right in the earlier this thing we are called as theta, so let us keep calling it theta so theta V square=A transpose U, U is all the values of dependent variable which we know okay we know this.

Now this as I told you is nothing but the Hilbert matrix, so this actually what you can show that this matrix in this normal equation will tend to Hilbert matrix, this not Hilbert matrix in the earlier case we directly got the Hilbert matrix. If you have large number of data points okay, what do you can show is that this matrix the proof is just given here, I am not going to write it down on the board, this matrix will tend to Hilbert matrix.

So for large values of see A transpose A this tense to 1 1/2 1/3 1/m, 1/2 1/3 1/4 okay, so large number of data points here okay. This particular matrix A transpose A okay you can look for the arguments here, this matrix will tend to this matrix which means you know basically if you want a good approximation, what will you do? If you want a good approximation you will take more data points which is logical okay.

If you take instead of taking you know 100 data points if you take 500 data points, you will get a better approximation okay, but as you take more data points and try to fit the polynomial this matrix A transpose A built into the Hilbert matrix, and Hilbert matrix is difficult to invert, because this is the ill-conditioned matrix, because this is an ill-conditioned matrix we have a problem that we cannot get reliable estimates of theta okay.

We will not get reliable estimates of theta, and because we cannot get reliable estimates of theta approximation that you develop say 10th order or 12th order is not a reliable you know is not a reliable polynomial approximation. A better way here would be to use orthogonal polynomials, instead of using orthogonal polynomials will make this matrix there are slightly different ways of constructing orthogonal polynomials for the discrete data case which are discussed in the notes.

So I am leaving it for you to read, but the idea is that if you use orthogonal polynomials then this difficulty can be avoided. There is something that we already have discussed is how do you use these least-squares approximations for fitting certain kind of correlations in chemical engineering. We discussed about Sherwood number correlation or we discussed about Prandtl number and you know Nusselt number =Reynold's number raised to something.

And if you should take a transformation then you can linearize that problem, and then solve the problem of estimating you know constants in this correlation using linear least squares, so that we have already seen. So we already seen one application of linear least square in approximation okay, so right now when I am doing CP as a function of temperature, I am trying to approximate a continuous function using a polynomial function.

This is one common application you have some algebraic static map, and then you trying to approximate using some other known-functions, and the problem is transformed to estimating this theta okay, and we also saw in any general Hilbert space how you could construct the normal equation. I am uploading extra problems today for least squares, and so you can start looking at the problems, where you will see lot of engineering applications, where you can you will be using this least squares.

The one of the things that I promise to you is that we will now move on at some point to solving boundary value problems or partial differential equations using this method, well that is what I am going to do next that is see how I can apply this method for solving a partial differential equation. So now I am going to use the ideas of least square for constructing a solution of a boundary value problem okay.

So right now I am going to restrict here to showing you 2 methods, one is method of least squares for a boundary value problem, I will not show explicitly an example of partial differential equations. I will also talk about a variant of it called Galerkin's method, now these 2 methods though we will be looking at them very briefly right now, we are not going to look at the details of this beyond a point.

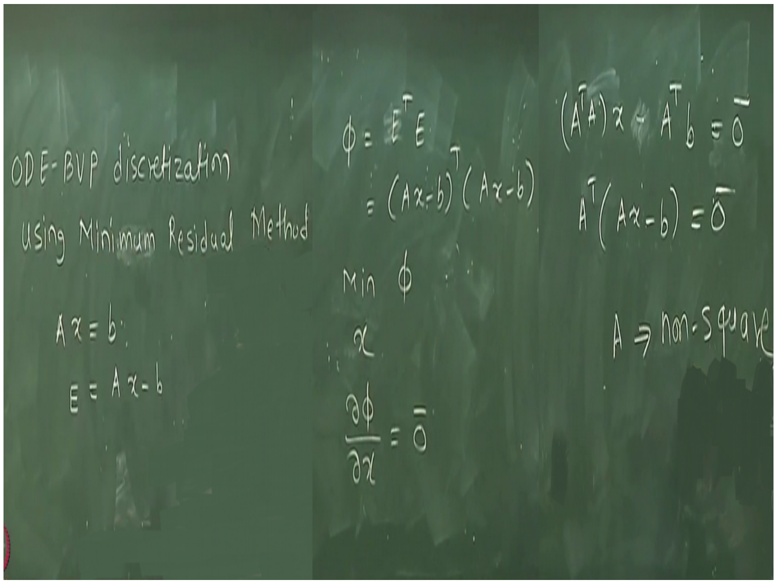
Because this method become very, very complex, and what actually you know this is actually we are getting into the method of finite element okay if you have to actually implement this you have to get into method of finite element, and finite element methods are fairly complex when it comes to computers implementation. So it will go beyond the scope of this particular course, so I have to stop at some point problem discretization, and move on into you know solution techniques.

So I am just going to give you a very, very brief introduction not really complete introduction, but at least the philosophy of this problem discretization that is say boundary value problem, how do you discretized this using least square approach, philosophy will be clear, we will not go into the details okay. So details will be I mean if I have to get into details it will take many lectures, and we have to move on and do other things, other than just discretization okay.

But this once we do this you know it is short of completes the picture, we have 3 different ways of problems discretization, one is a problem in this particularly here I am referring to boundary value problems or partial differential equations, we discretize using Taylor series approximation, we discretize using second is interpolation we got collocation methods okay, third is this least squares okay, we get this finite element method.

First is finite difference do not confuse, finite difference method, then we had orthogonal collocation, then we also had orthogonal collocation on finite elements okay, and this method but that basically is interpolation based method, and the third one is this least square method. Interpolation, we want a polynomial to pass through every point okay that is the difference. Approximations, we are trying to fit a polynomial in some least squares, we do not want to polynomial to pass through every point or the function to pass through every point okay.

(Refer Slide Time: 14:42)



So let us look at this method of least squares for solving, so this method is called as minimum residual method, and well why can you solve a problem using optimization, and just to give you a little bit of background before I move into least squares okay. We will revisit what I am doing right now on the board again, but I am just preempting something. See we know how to solve A x=b, right now from your under graduate, how do you solve A x=b? Gaussian elimination.

The best method you know is Gaussian elimination, you do you know triangularization, and then solve in the reverse direction so that is one of the efficient methods that we know, can I use the ideas of optimization to solve this problem okay. Now let us see whether we can use the ideas of optimization to solve this problem, what we wanted in optimization? We want to minimize something okay, and the minimum should be the solution okay.

Now let me take this before I get into this let me say that I want to solve A x =b okay, I want to solve this exactly okay I could say that well this problem is like saying that you let me define an error vector, which is A x-b everyone with me on this okay, and then I define a scalar objective function okay, so my scalar objective function is going to be, now well what I am going to do now is very, very similar to what is happening here.

So this is my error vector, this is A, theta and this is U, except in this particular case we had you know this matrix A was tall matrix it was not a square matrix, right now what I am writing there is a square matrix okay, so I want to solve square matrix problem by optimization is it possible. Okay, so now let me formulate an objective function phi, which is E transpose E right, I formulate an objective function E transpose E.

And then what I do is which is what is E transpose E A x-b transpose A x-b okay, and now I say that well minimize phi with respect to x, how do you get the minimum? First derivative =0 right, so dou phi/dou x=0. Can you solve this, what will you get? If I solve this if I use the rules of differentiation of a scalar function phi, is phi a scalar function? Phi is a scalar function right; I am differentiating a scalar with respect to a vector.

We have looked at the rules of differentiation of a scalar with respect to a vector, if I use those rules, what I will get is this equation A transpose A x=A transpose b, I will get A transpose A x-A transpose b=0 this is the equation that I will get, if I differentiate and set derivative of phi =0 I will get this equation okay, is this fine, just see this you will get if you expand this you will get A transpose you will get x transpose A transpose A x.

Then you will get A transpose x transpose *b you will get 4 terms, you will get one term which is b transpose b, b is not function of x, dou b transpose b/dou x=0 okay, so if you differentiate you will very easily arrive at this particular equation okay. I just want to show that this is same as A transpose A x-b=0 okay. Now if A if matrix A is non-singular okay, when will this be=0 A x-b=0 okay, so actually by solving this optimization problem you have solved you have reached the solution of A x =b.

And later at some point we are going to see a method for solving A x =b using optimization, an iterative procedure which is faster in arriving at a solution many times, then is this the optimum why is this the optimum? Why is this minimum? If you take second derivative what will you get

A transpose A, A transpose A is always a symmetric matrix, whatever is A, A transpose A is symmetry positive definite matrix very nice matrix okay, so we have reached the minimum okay.

So this is I can actually formulate a problem of solving A x =b as an optimization problem. **"Professor - student conversation starts"** yeah, is this exact solution? This will be exact solution. **"Professor - student conversation ends."** Okay, earlier we saw approximate solution, so in this situation where in the situation where you know A is a tall matrix okay, which means a is a non-square matrix it has more number of rows than the columns okay.

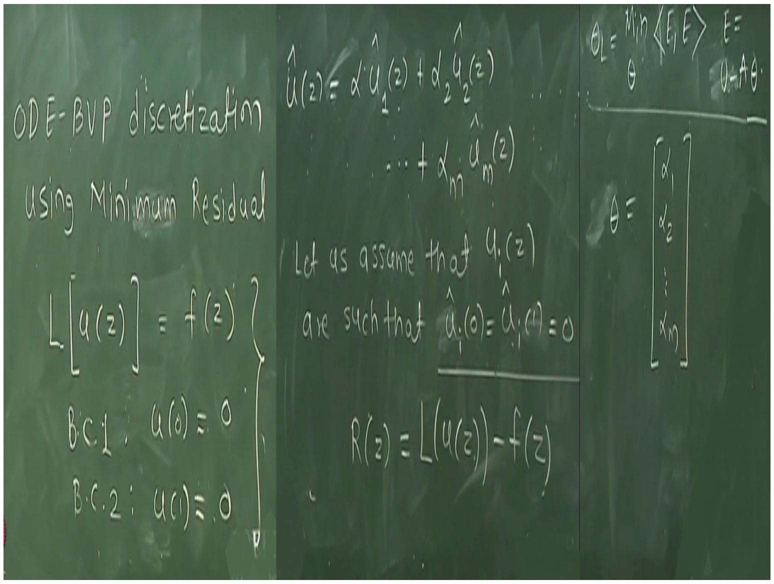
Then you will get the least square solution, when you when A is a square matrix invertible square matrix, this will give you exact solution, so look here know this will give an exact solution, if A is invertible when can you solve that problem A x=b A is invertible, if A is not invertible you cannot solve that problem right. So now we are starting with the assumption that A is invertible okay, and we want to solve the problem.

So this obviously gives me the solution if I do this okay. So optimization could be a route to arrive at a solution of a particular problem okay, what I am going to do for the case of boundary value problem is more similar to the case where A is tall okay, so a is non-square, what I am going to do for boundary value problem is called to be similar to A being non-square that is A has more number of rows than the columns okay.

It is possible to do at least formulate the problem to get an exact solution of a boundary value problem using optimization, and that discussion is included here in the nodes under the name of Rayleigh Ritz method, it is very, very nice you should read this section though I am not going to do it on the board, it takes lot of time and this actually forms the foundation of finite element method, and finite element methods are very, very useful in solving partial differential equations.

So I would say there are these 2 competitors finite difference is very easy to understand, but finite difference is not so efficient because you need large number of grid points okay, 2 methods that are competing good methods, you know which balance between computational efficiency and good solutions you have good balance of these 2 things or orthogonal collocation, collocation interpolation based methods, or these finite element methods.

(Refer Slide Time: 24:08)



So now let us move on to boundary value problem okay, now I want to consider a specific problem here boundary value problem. So there is some operator L, which is operating on u z okay, I suppose now you are familiar with these kind of things, because you are most of you are attending the other course on mathematics mathematical methods, and you have looked at these kind of problems right, L operating on u of c this=sum f z.

And boundary condition 1 this corresponds to u 0=0, and B C 2=u 1=0, so this is classical boundary value problems that you get when you are solving partial differential equation. This will typically yield if you solve it exactly, if you know the solution exactly it will yield the eigenfunctions expansions, are you doing eigenfunctions expansions, yeah so this will yield the eigenfunctions expansion.

Right now my mandate is not to solve it exactly, my mandate is to solve it approximately okay, I am going to solve it. Now here I am assuming that L is a differential operator which is operating on u, conceptually we are not doing anything different than A x =b and we are talked about it right, this is equivalent to operator A x =b we are trying to find out inverse problem classical inverse problem. So right now I am not interested in the exact solution.

I am interested in some approximation solution to this problem, and I am going to hypothesize a solution approximate solution u cap z=alpha 1 okay I am going to hypothesize a solution, now what are this u1, u2, u3 these are some known functions, these could be simple polynomial if you want it that way, but you know polynomial is not a good choice okay, this could be Legendre polynomials shifted Legendre polynomials, it could be sin and cos.

It depends on how you choose this functions, it is up to you and you have to choose this function such that the boundary condition are met of course. So I have this functions which I have chosen here, let us assume that so let us assume that we have chosen this functions for the time being okay such that at the 2 ends this satisfy the boundary conditions okay. We have chosen the functions such that at 2 boundary points you satisfy the boundary conditions okay.

So you have to choose this functions in this particular case carefully, now what I am going to do is I am going to define this residual okay, the terms used are little different but the meaning is same in the least square parameter estimation we call this error vector okay, here we call the residual vector okay, residual is between left hand side and the right hand side okay, so the residual is between L u z-f z this is my error vector.

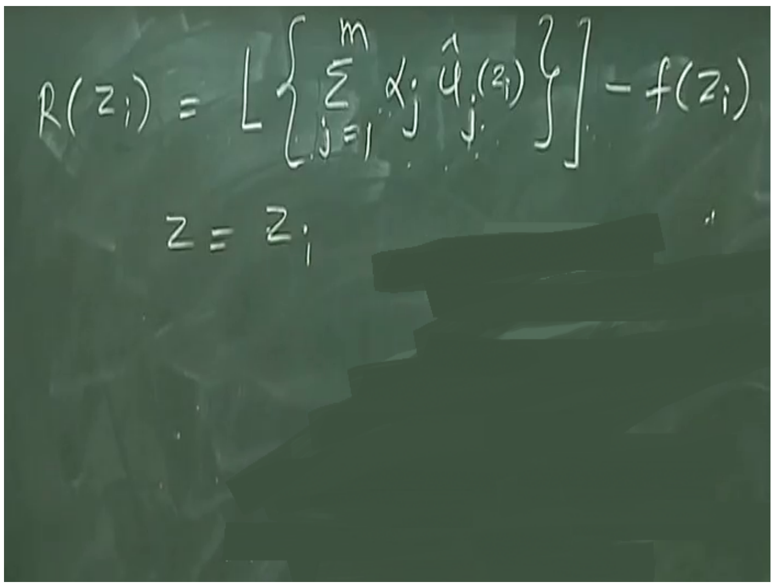
In the case of finite dimensional least squares, what did we do? How did we find the minimum? Sum of the squares of errors was minimum, what are actually sum of the square of errors? 2 norm square, what is 2 norm square? Inner product of the vector with itself right. So for the finite dimensional case we had this keep this in the background, in the finite dimensional case we wanted to obtain theta.

So theta least square was minimum with respect to the theta you know inner product of error and error right, sum of the square of errors least square estimation, error vector was where error was defined as U-A theta right let us keep this in the background okay. I just want to do the exactly same thing here for this particular problem. Now I have this functions u1 u2 u3, so what I am going to do is, I am going to pose this problem okay as minimize.

I am skipping the algebra which can be a little bit tedious I will write the final expressions, but whatever we have done till now from that you can very easily derive what I am doing, if I have to do the same thing here I will have to write this summation of these functions, and then start working with those summations, and expression will become complex but that is just an algebra, the concept is exactly identical, so you should not.

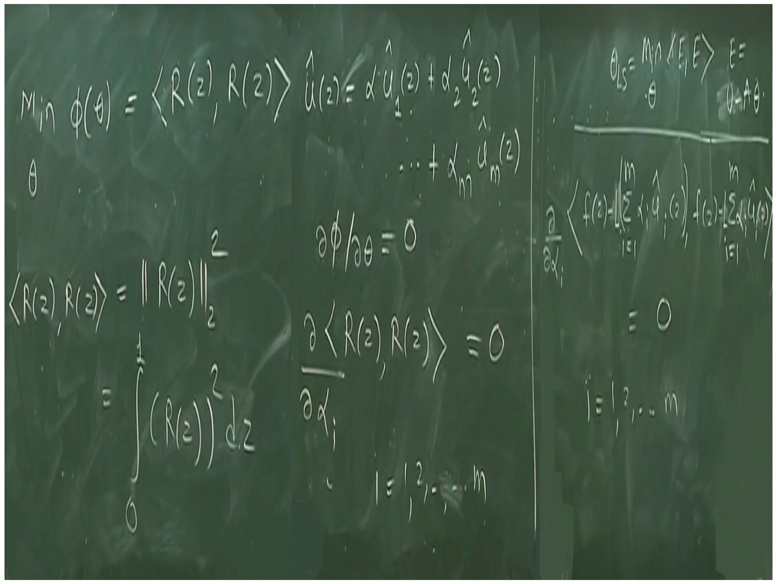
So what are the unknown here? What is the theta here? Alpha values alpha 1 alpha 2 alpha m right okay.

(Refer Slide Time: 30:19)



Now what is R u zi? Well you will have to substitute one by one each that is not the way to go about okay, so we want here R zi, what is R zi? R zi is L operating on summation I going from 1 to m alpha i ui of zi, let me do to a small modification here let us call this j, let us call this j, and let us call this j, this is residual at z=zi okay, do I need the points here, I am not going to go same way as the grid points. I am going to go in the parameters space directly.

(Refer Slide Time: 32:09)



So what I want to do here is instead of forcing the residual to be 0 at each grid point, well what I am going to do here is to pose my problem as minimize with respect to theta phi theta which is =inner product of R z, well this is something different from okay you should realize this is something different from I am not taking any grid points right now okay. So what is this, what is the inner product defined as, inner product integral say 0 to 1 right, okay.

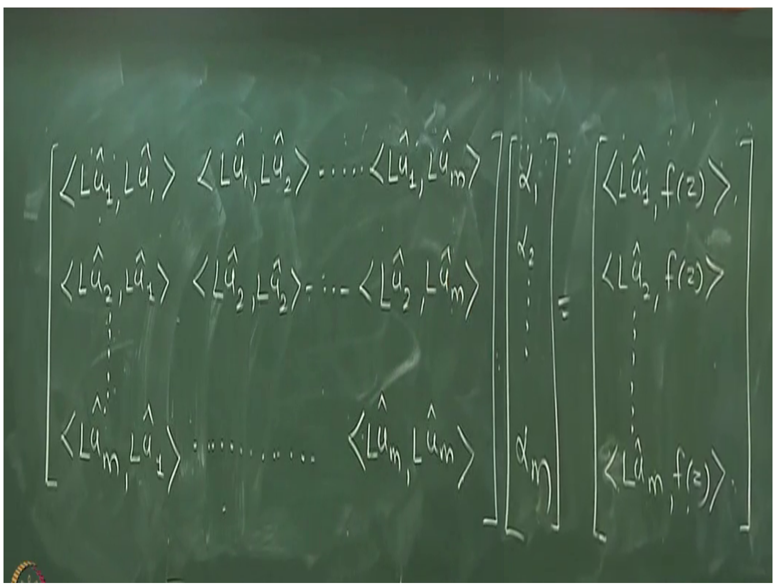
Now how do I get the equations from this? I have to differentiate this with respect to alpha with respect to theta right, so now what I have to do is to get the solution, I should see how many unknowns are there, there are m unknowns I need m equations okay, so if my phi theta is this then what is how do you get the optimum dou phi/dou theta=0 okay, so which actually means dou/dou alpha i of R=0 for i going from 1, 2 up to m, is this okay, dou phi/dou theta=0 which is same as this is my objective function error right.

This is my error square sum of the square of errors, but it is not the sum it is integral of error it is not at one point, not at few points, but over the entire domain okay. There is a difference you can see here between what I am getting the way I am connecting the solution and finite difference method are orthogonal collocation method. Here, we are setting this with respect to =0 okay, if I solve for this okay, what do I expect to get? See what is R z here?

So you have to actually solve this problem of dou/dou alpha i okay* what is this this is f z-summation i going from 1 to m alpha i u cap i right, what is R z? R z=f z-this right, this =0 okay. And this equation you have to solve for i going from 1, 2, up to m. **"Professor - student conversation starts"** (()) (36:01) yeah, L of this right, operator L operating on this, yeah, so thanks for this correction. **"Professor - student conversation ends."**

L operating on this and L operating on this, and I have to take derivative and set it =0, but if L is a linear operator if we make an assumption L is a linear operator okay, then this alpha i will come out and then you can differentiate and solve the problem very, very easily. So if L is a linear operator then what I can do is if I actually write those equations, and collect the terms together okay, I will get this equation which looks very, very similar to the normal equation.

(Refer Slide Time: 37:08)



You can work out the algebra I am just skipping it; it is just matter of doing it meticulously. You will get this equation L u1 okay, you get this equation if you rearrange, if you actually differentiate do a little bit of like a trouble of rearranging all the equations and then put them together, you will get this equation which looks very, very similar to the normal equation okay, except that operator L is operating on each one of those basis functions which we have chosen okay.

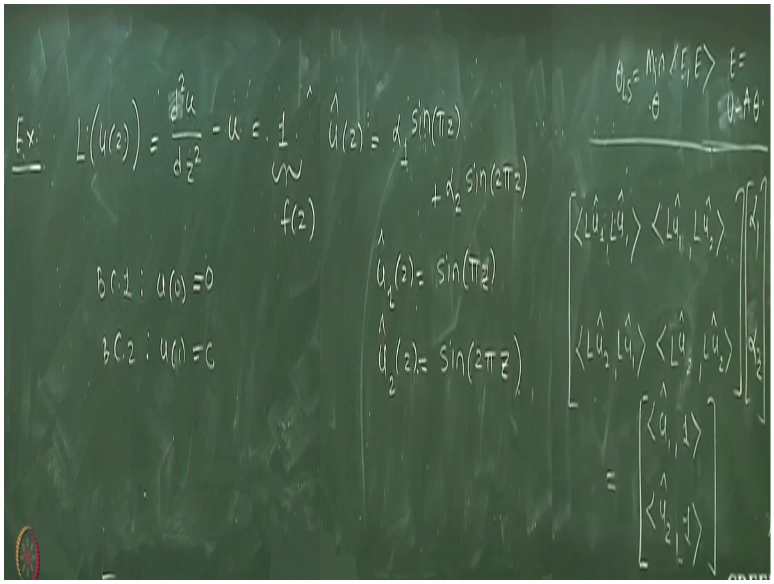
And then if I solve this then I will get alpha 1 alpha 2 alpha 3 alpha m, I am getting an approximate solution in the least square sense okay, there is one difference here between the previous methods of discretization and this method of discretization. What is the difference? Here these integrals are over the entire domain not at few points okay, I am not evaluating this integral at few points, I am evaluating this integral over the entire domain okay.

So the finite dimensional approximation comes because of the finite number of basis vector we have taken, the finite dimensional business here comes because of not because of grid points okay, here you have constructed a solution which is not, see in earlier case when you get the solution you must have done it now for the boundary value problem which we are looking at Tram problem, it is some points right finite, after all you get only few points.

Actually, you know that the solution is a continuous function, by this approach we are getting a solution which is a continuous function, I will just take a simple example, so that the ideas will become clear okay, is this algebra clear, is there any doubt about this. See here L this residual is f z-L operating on this function okay, we are taking L to be a linear operator okay linear differential equation, and then we are solving this problem of you know estimating alpha 1 alpha to alpha m.

Now let us look at a specific problem, well this particular method will work by this way, because when operator is linear, if the operator is not linear what do I do? We will come to that a little later. But right now we are just looking at a basis of the method, which is least squares method and the operator is linear, so we get this nice. So let us look at the specific problem okay that will give you better insights into.

(Refer Slide Time: 40:49)



Well I want to solve this problem of, my example is L u z this is same as d 2 u/dz square-u=1, so this is my f z okay, I want to solve this problem this is=1 okay, over the entire domain this should be =1. And my boundary condition 1 is u 0=0, boundary condition 2 is u 1=0, well u !=1 please note that you are saying that this operator operating on u should be =1 okay do not confuse okay. Now I am going to take a very, very simple solution to this problem.

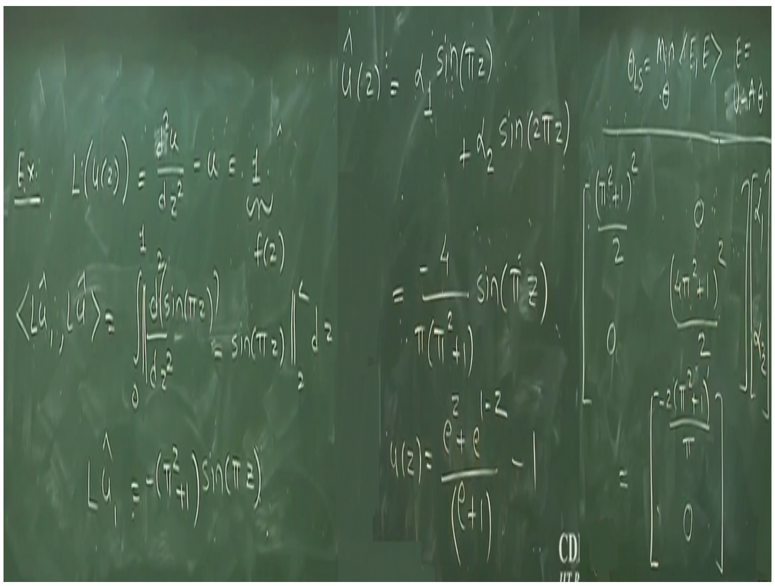
This particular problem you know you must have seen in the other course that this particular problem can be solved analytically you will get analytical solution, right now the motivation is to you know convey some concept okay, so that is more important right now. So let us take this approximate solution u cap z okay, so my approximate solution is alpha 1 sin pi z +alpha 2 sin 2 pi z okay, does this satisfy the boundary conditions?

Satisfies the boundary conditions at 0 and at 1 boundary condition is satisfied. Now I want to see if I get difference between finite difference method okay and this method is that if I get alpha 1 alpha 2, I can find out value or approximate value of the solution at every point z, not at finite number of points you get the difference, if I can get least square estimates of alpha 1 alpha 2, I get okay. So now what is the normal equation? The normal equation.

So what is my u 1 cap z= sin pi z, and what is my u 2 cap z=sin 2 pi z right, sin pi z and sin 2 pi z these are my u 1 cap and u 2 cap right okay. What is the normal equation? What is the equation that you get? Well it is not normal equation let me not call it normal equation because not exactly the normal equation. So what is the equation that you will get here? How do you get alpha 1 alpha 2?

So this is inner product of L operating on u1, L operating on u1 cap, inner product of L operating on u1 L operating on u2 cap, L operating on u2 cap L operating on u1 cap, L operating on u2 cap L operating on u2 cap, this matrix okay this matrix times alpha 1 alpha 2 will be =to get alpha 1 alpha 2 what should this be=inner product of u1 cap 1, inner product of u2 cap 1 right. So what is inner product of L operating on u1? What is L.?

(Refer Slide Time: 45:27)



So let us go back here, so what I have to do, what is the first inner product L u1, L u1 this is inner product 0 to 1, what is L u1? d 2/dz square of what is u1? d 2/dz square of sin pi z-sin pi z- this right. So this integral I can do I can just it is not so difficult to do this integrals, you have to be a little bit patient and then do these integrals okay. If you work out these integrals okay, well I will write down what you get here actually just if you take this.

What you can show is that L u1 this is equal to -pi square +1 sin pi z, you can show this very, very easily, you can also so what is L u2 okay, and then you just have to find those integrals. If you find those integrals, what you see here is you will get pi square + 1 whole square/2, 0, 0 and

4pi square + 1 whole square/2, right hand side will turn out to be -2 pi square+1/pi and 0. If I do patiently all those 4 integrals okay, then I get this equation.

I have chosen orthogonal basis functions here, sin pi z and sin 2 pi z are orthogonal, so that helps me in this particular case that help me in getting these to be 0, this comes out to be a diagonal matrix. And what happens here is that I get the final solution. So what is the final solution? If I compute alpha 1 and alpha 2 I get the final solution to be, my solution comes out to be -4/pi square+1 sin pi z, so which means alpha 2 comes out to be 0, alpha 1 comes out to be this.

And this solution is pretty close to the in this particular case the true solution can be found out, the true solution is u z equal to, this is the true solution, this is the approximate solution. True solution can be found by some different means not buy this approach. **"Professor - student conversation starts"** yeah (()) (49:13) and then force the residual to 0. So now I choose some of these functions so much standard basis know.

I can choose them from a standard basis like sin cos, I could choose them from Legendre polynomials shifted Legendre polynomials, I could have worked out this problem using shifted Legendre polynomials, if I have those shifted Legendre polynomials which satisfy the 2 boundary conditions. In this particular case sin functions are convenient because they satisfy the 2 boundary conditions.

They need not be orthogonal, they need not be orthogonal but in this particular case orthogonality always helps, we always work with orthogonal functions okay, we try to work with orthogonal functions as for as possible, orthogonality is. **"Professor - student conversation ends."** But this solution there is one difference between this solution and other solution, other solution is exact at known points, exact means it is not exact exactly satisfies the differential equation at known points.

Here, this solution is obeying the differential equation in the least square sense okay, at every point it is not =0, but if you take the least square sense it is 0 okay, integral of the square of error is 0 error is !=0 okay, integral of the residual square is 0 okay. Even in the other case we are forcing residual to 0 only at finite points, what happens to the residual at the points which are in between? They will not be 0 right, so we are only forcing at a finite number of points.

Even here you know you do not know which point exactly the residual will be 0, but solution will be crossing you know 0 at many points, so this is both are approximate solution this is a one way of getting approximate solution, and you have other way of getting approximate solution which is what we have seen earlier okay. Now the problem is that in everywhere where you are solving this your L will not be in general you know it will not be linear that could be a non-linear differential equation.

We have solving this Tram problem right it is not a linear operator okay, so what do I do, what will I get, how will I solve that problem? So in that case you know I just use the concept of orthogonality between the subspace and error the concept that we had earlier, and we will get the method of you know Galerkin method, so I will discuss it briefly in my next class. But in the next class after doing that half an hour of Galerkin method something about Galerkin method.

I want to move on to solving problems because we are only 2 classes away from the end sem and complete sem, and we will solve some problems of least square, so that you get better insight into what is happening. But the key thing here is that right now I have not use polynomial approximation, I use general functional approximation, the solution is obtained by posing the problem as the least square problem okay.

And then you are getting a solution approximate solution of the boundary value problem using the least squares approach okay, it is possible to extend this to partial differential equations, we will have to compute double integrals in x and y and so on okay, it is possible to do that in, algebra will become complex the concept is not different.