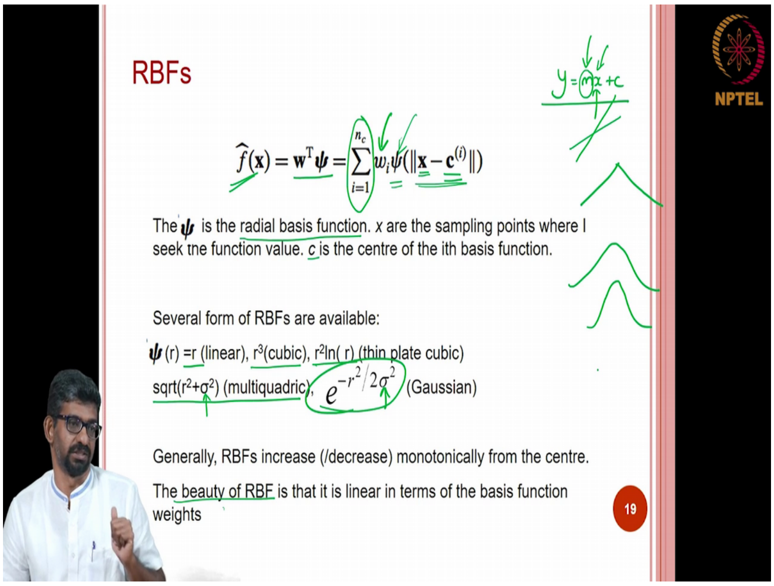
# Surrogates and Approximations in Engineering Design

**Prof. Palaniappan Ramu**
**Department of Engineering Design**
**Indian Institute of Technology, Madras**

## Lecture – 13
## Radial Basis Function – 2

(Refer Slide Time: 00:17)




Now, we will go back and we will see what the math; is simple it is only representation we are not even looking at the math in detail. It is only a representation, ok. Similar to the problem that what we have seen I want to get w T psi. I am making this a generic approximation I want to identify f hat. What is this psi? The psi is a basis function in our case it is the radial basis function. What is this x? This x is the sampling points meaning the new points, where I seek the function value I am just making a generic formulation here, ok. So, this x could be anything that you want, wherever whichever point that you want. What is the c? C is the centre of the i-th basis function.

So, in the example that we discussed there were three basis functions that we use three triangles and this c i since, I am calling it i in the superscript is the centre of the i-th basis function, and you also please remember that we are doing a sum here, ok. So, if your x so, if your x was in the right of the third function, then I will look for c 3 that is what it means, ok. So, I look for c 3 yeah. So, this is a psi it is going to take some shape will see what shapes are possible, but this is nothing, but your x, ok.

So, this is like this psi is an equivalent of your slope in our y equals m x plus c. See, you cannot directly compare, but for you to relate I am giving this, ok. Do not meaning do not do not record this, ok. This is for you to immediate relate immediately, ok. You give me any new x, with the same m I will be able to find that new y for you ok, but as my m changes what is going to change? The line is going to change.

So, is the case here the psi tells me the m part of it in ones not them part of it, but m x plus c kind of stuff, m x is what this guy is about and then this x is nothing, but this one. So, I am just giving the x in a relative sense imagine that this x is not directly x, but you normalize this x and give, ok. So, if I am saying x is 120 you can feel it as 20 because you are normalized version is 100 is equal to 0, that is all. This is in a normalized sense that is it.

Now, there are two things, you need to find your w that is one thing which can be done using a simple maximum likelihood kind of an estimate, so, we will see. What are the forms that this psi can take what are the forms that we have already been exposed to through the examples and then we saw something like this, correct or I mean it was yeah nice fat these are the two forms.

But, there are also different forms that are they are all governed by some equation, psi of r r is just the difference normalized sense let us say it is linear the one that we discussed ok, it could be cubic r cube. So, for every distance for 1 it will be 1 for 2 it will be 8, ok. Then there is something called a thin plate cubic. So, it goes in this form r square times log of r. There is some multi quadric something that we are interested in general is there our Gaussian.

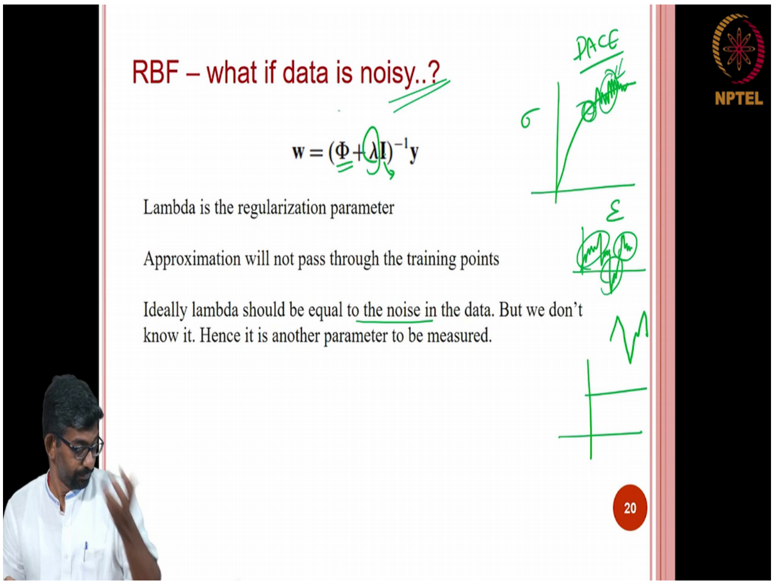
So, the Gaussian if you see is slightly different from the top. The top requires only one parameter, but you know with one parameter. I have only some control, two parameters I am likely to have better control, ok. I can control the shape as well as the size in this I can only control the shape if you look at it in the first one r, r cube, r square I can only control the shape. Whereas, in the second one I can also contain the size that is the advantage of doing, but you will have to find this guy that is another problem. There is another you are introducing another parameter trying to get more control, you are trying to introduce another parameter it is basically features, ok.

If you want to do more you need more features on your cell phone, ok. A basic phone yes you can get it for how many rupees 1000 rupees you can get a Samsung, Nokia 105, but it will let you only call and receive calls ok, but you want to do internet, you want to have Google map, you want to go elsewhere you want to do Whatsapp, you want more features then you need to pay more that is what the idea is. You need to pay more in terms of computational sense ok, ok.

So, these are there stuff that we have already discussed generally RBF'S increase or decrease monotonically from the centre the beauty of RBF or any of the approximation that we are talking about the however, complex your expressions are you are still doing only a linear superposition. You are only going to solve for these w's in a linear regression sense it is not a non-linear. But, your the psi could be the most complex, ok.

Sometimes I do not even remember the error functions that are built I just choose Gaussian ok, but it is still from a solution sense it is still only a linear regression. Because, you remember the Grammian matrix that we discussed about write x transpose x inverse that is what you are going to do, that will still remain the same that nothing will change. Because that is only a function of your x the gram matrix is only that part, that is not going to change. So, that is the beauty of all these guys,

(Refer Slide Time: 06:42)




But, anyway the question is that is great, but what is the psi that I will choose because you do not have the original function, just like I showed you in the example. The original function could have been anything, how do I know, ok. This is the challenge whenever we go you know even with industry practitioners we go people generally ask, how many design of experiment points should I use? No one knows, if I know that then I will copyright that and I will be one of the richest props or at least in that community I will become one of the richest person because invariably anyone who is trying to do design using computer models today will use it design of experiment.

And, there is a very underlying question is how many points should I use for these many dimensions. There is no thumb rule at all. That is when the designer comes into picture you need to have some understanding of what the problem is, ok. I know that this is a complex cantilever borrower problem, but do you think still need to know me give you a cantilever kind of an effect then it is going to vary only in terms of l cube. What is maximum it will vary? Maybe 3.5, non cubic it is 3.5, good enough I will use a fourth order it, at least that the engineer should be able to feed in. Because, depending on the order of the equation my number of coefficients that I need to find is going to vary for which the number of data that I want are more, accordingly I need to do then always you have a computational budget.

Computational budget will let you only find two unknowns. But, you are prescribing a fourth order polynomial, I cannot do that. So, then what I need to do is I need to find a balance. Can I fit something that I know is not accurate, but at least will give me give me an idea on how far I am from the original predictions. So, then I will make those decisions this is what people have been able to do.

So, the whole idea right now that we spoke about this their data this is parallel to the question that you asked. The data is not noisy; the data is kind of a deterministic data which is. So, the one thing that you need to appreciate is these ideas come out of days which is design and analysis of computer experiments. So, computer experiment there is no noise you give the input it will give you the same output any MATLAB code, anything ok. So, you give a wrong output it will maybe a wrong input it will give a wrong output, but that is fine, but the wrong output will be consistent every time. Unless you have a random number generator inside which does something field, if it is an analysis stuff it is only deterministic in sense.

So, the assumption that so far we discussed is there is no noise in the data per se, but that is not true in experimental sense. Does it mean that RBF can only be a used for applications like the computer simulation. No, that is not true. You can use it for any kind of a function approximation which was which was which is what makes it meaningful and sensible, right. So, the question is what happens if the data is noisy? Then, what it means is you do not want you do not expect the curve to go through all the points, ok. Unless you have a 100 percent confidence on the data you have you do not want that because you know the data also has errors correct the points that you estimated also might have errors. Then you want to do only an average sense.

See, you remember please remember the stress strain curve that I told you in the beginning. Whatever you got in the book is a lie. They did not tell you exactly how they got this, ok. It was actually like this and then why did not I go through the points? I only averaged it because if I went and did it in a different machine, I might get a slightly bigger or a smaller spectrum, ok. I changed the guy who tightened the specimen I might get a slightly different one, ok. Even my Young’s modulus might change accordingly, little bit. The material the moment I put the second specimen I might get a slightly different result.

So, the curves that you see are average curves, which is not explicitly mentioned. So, you do not want to follow any of these lines because these data themselves are noisy. See anything that you see in one sense that is also good. Any electrical signal that you see has noise one thing that is related to your life is what? Your ECG signal. Have you seen your ECG signals or anyone's ECG signal? Ok, this signal has noise, but it also captures your PQRST which is in your lead the lub dub of your heart right, but that should be very clear like this is very clear lub dub lub dub and there is a small. So, it is PQRST – PQRST that is what, but then the signal is always like this you put a strain gauge and get some strain signals it is always like this.

So, it is corrupted with noise, ok. Noise means not the noise that we know as noise, this could be electrical noise it could be the or fitting me the way in which you have fastened it and all that. What happens if there is no noise especially in your electrical signal? The story is over ok. So, there is always scope when there is noise for improvement and for betterment, ok. So, this is why I like optimization and data fitting in general, because they are very reflective of your life if you do not get some results, you should not crib about it rather you should look at this as an opportunity.

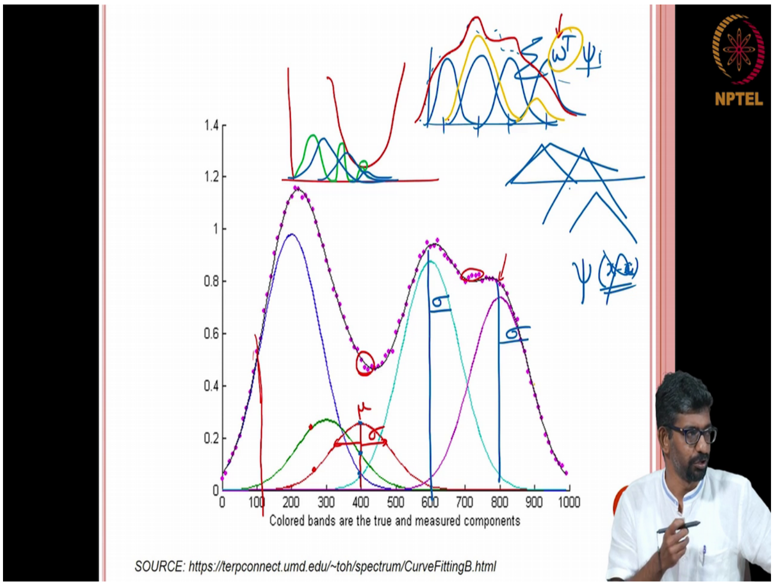
Because, if you if you say no if you get your results then the job stops there is no further scope in the research ok, but do not look for not getting yourself you should get results, but if you do not get results do not worry about it, ok. So, non-linearities offer you opportunities to improve and grow. So, then what people do in such cases is if the data is noisy then you do not want the data I mean the fitted model to pass through all the points, not necessarily ok, I want only an average fit. In that case we use something called a regularization parameter this is more like why you call like trying to preserve the pattern I am trying to preserve the topography kind of an idea. So, it is this is widely used in machine learning technique this is the regularization schemes.

So, this is just your identity matrix the phi is nothing, but the matrix of your basis functions and depending on your see ideally this lambda what you are trying to do this you are trying to take your basis functions and you are trying to add the noise to it. Ideally what should be your lambda then it should be the noise, ok. But, you do not know the noise if you know the noise you would to begin with itself we will do a better job you do not know the noise. So, ideally the lambda should be equal to the noise in the day you get the point or not ok.

So, usually it should be the noise in the data, but we do not know it. Hence it is introducing another parameter then what happens people will try to tune the lambda. So, you are introducing a third parameter, I mean from whatever we discussed so far. From whatever we have discussed so far you introduced a new parameter which is the regularization parameter. So, you will have to choose a regularization parameter in such a way your error metric will go to minimum that is all the overall ideas, right. In our least square also what did you try to do? I am fitting and I am trying to compare it with my original realizations. And, you please appreciate the point that even in the polynomial regression it does not go through all the points, it is only the best fit.

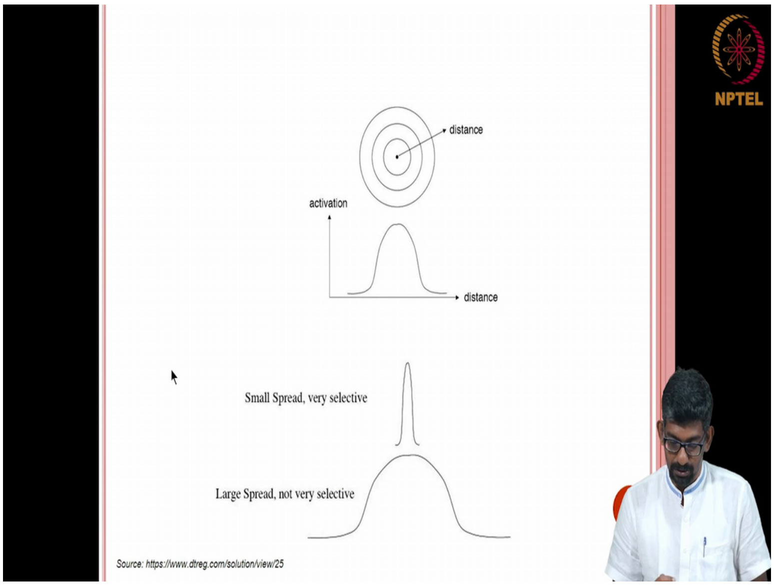
The question is how do you define what your best fit? Your best fit was not that the point should go through all the points. My overall error should be minimum. So, that criteria will still remain the same in this, what you will do this along with your the basis function parameter you will also tune your lambda such that your overall error is going to be less. So, that does not necessarily mean that your curve should pass through the points that you have because the points themselves might have errors. So, forcing them to go through that means, that it will still be around us. So, you do not want to do it, ok. So, that is one way to deal with noisy data.

(Refer Slide Time: 15:25)



Most of these curves are taken out of the internet and the Foresters book has enough examples for you to study.

(Refer Slide Time: 15:33)




So, there are different types of radial basis functions that I expose to you at least from picture perspective we looked at the equations, right. So, there is a linear variation, there is a cubic, there is a thin plate cube, there is multiquadric, there is an inverse multiquadric the meaning one over this guy and then the Gaussian function, there could be other types of functions as well.

So, we solve these things. For instance just to give you an idea in a Gaussian sense. So, let us say that the pink dots that you see are the bunch of data that I had and I wanted to approximate that function. So, what we are using here is we are using 1 2 3 4 5 radial basis functions, ok. So, we are taking them and then we are getting this kind of an approximation and when I add them up, this is simple if you see here if you see here up to here you can see how this linear superposition works right; up to here it is the same guy, there is nothing.

So, the function the approximated function just follows the blue curve, but from here I also add the green curve and the red curves value, ok. So, then you can see there is a small deviation from the blue curve. When it goes here the blue is peak, but I also get some value from here and I get some value from here. So, those two added up and it gets fitted like this is one of the best fittings, ok. But, the reason is there are also a lot of points you can see and the line does not pass through all the points. So, this is an average fitting in that sense, ok.

Student: (Refer Time: 17:07).

Those are one RBF.

Student: (Refer Time: 17:13).

They will correspond to the 400 kind of stuff; for instance this guy.

Student: (Refer Time: 17:20).

This guy is at 400 kind of stuff.

Student: (Refer Time: 17:23).

Sorry.

Student: The original (Refer Time: 17:26) some of it is not.

Student: (Refer Time: 17:28).

Student: (Refer Time: 17:36).

They will not. Why should they?

Student: (Refer Time: 17:42).

So, this is basically it is also capturing your what you call your variations and your mu value because at this point see, I see what your question is if my functions were like this I will just take the same example, I see what your question is if my function I see your question just follow. If my functions were like this then it should pass through that correct, because at this point though I am going to say RBF 1, RBF 2, RBF 3, RBF 3 is only going to play a role, but in this case what is happening is in this case what is happening is. So, at this point I am going to add all the three blues. So, they are a contribution from the other RBF’s also.

So, because if you see the fitted line I guess that you will agree, right because this is my mu value and this is the value from the green value this is from the blue this is from the other blue all put together I will be able to go here I am building upon the scaffolds that each one gives me. So, that is why the point the same reason that you pointed out in the other example we had three points, but then they ended. They ended and then only the next one started, but here it is not like that it started it going and then this guy is also starting here and coming like this. So, it will add up like that. You get the a point?

Student: (Refer Time: 19:33).

So, this is basically with the regularization parameter, yeah tell me. Sorry, the basis function?

Student: (Refer Time: 19:44).

Sigma are the same. Sigma’s are the same. Sigma means how fat they are. At least visually from this you can say that the sigma’s are the same for all of them. They are all they are all equally fat, ok. Height wise they are different, that is all and the form is the same. The form of the same the parameter that varies is only your mu.

So, where does this dwarf and the heights come into picture? Remember, this example where did they come there was one more stuff. This is just a psi that you are talking about. The fatness is a function I mean the fatness is one thing, and then I did something like this. So, this is just let us not worry about it this is just the new point prediction, right end of the psi is located around some x i’s. There was one more thing in this example, what was that? Lambda was for the regularization.

Student: (Refer Time: 20:56).

So the weights I am going to say I will give more weight for this guy because he is located around this point and there is only one out of this guy whereas, for this guy I will give a slightly lesser weight for this guy maybe is equal weight, but that weight will come out off your regression, right now we have seen only the interpolation part you understand, right. So, I was under the assumption that the regression part is understood. I did not explicitly say that. This gives you the idea about the interpolation. But, his question is how did you decide that the green is only this tall and the red is this tall? That comes from the fact that you will have to estimate your w i’s. Your w i’s are nothing, but your weights or the coefficients of each of these guys.

So, you are giving more weight for this guy you are giving less weight for that. Actually speaking all of these guys are like this was all of them are that one value to begin with. Then with respect to this point you are going to take a some of that you are going to solve that, this is what you are going to solve w transpose psi. To begin with what you will do is this the psi is the same for all of them ok, meaning sorry the weights are the same for all of them, psi is anyways the same. Psi i is the same the only thing that will differ as you are the centre, ok. It is actually x i minus c, only the centre will vary.

To begin with you will just have all of them like this let us say. I have four radial basis function, all of them will be the same to begin with; same means only these centres will be different. Of course, the how fat they are also is fixed the weights I am using the same weights for all of them. Then what I will do, here are the data points that I have something like this I will say. Then, based on the metric that I choose I will have to find out what my w's are. What this w will say is this leave it like this for this particular stuff the first one just leave your weight at that point it will say.

Then for the second one it will say probably you need to increase a little bit more. So, the mu I do not know on this yellow might not be able. So, this one it might say let us go up a little bit, so that you can capture this peak and then for this one it might say do not go that far, just go this much is enough. So, what it means is the height will be adjusted based on your w or your w is corresponds to your height that is what it is, ok. So, that is just a regression problem. So, this is what radial basis is about then that becomes a just a regression problem that is all, ok. Any question?

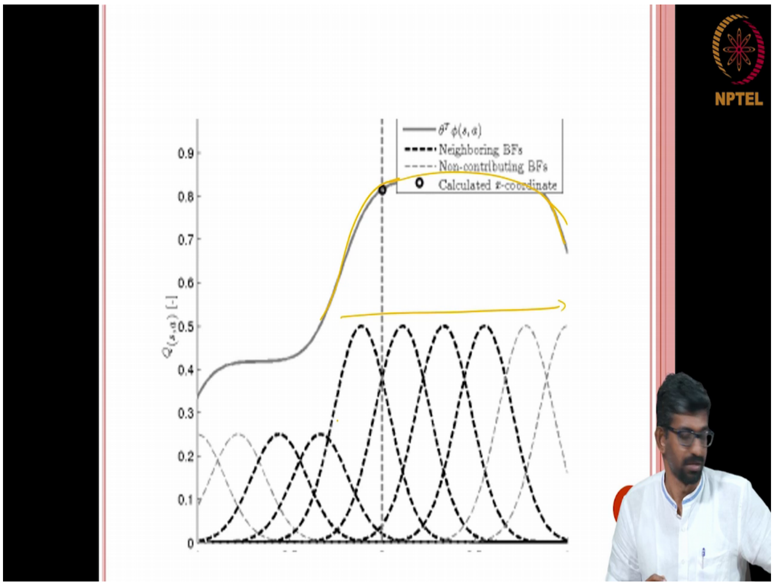
Student: (Refer Time: 23:54).

Sorry.

Student: How do you do (Refer Time: 23:58). Is it a choice or like you know there is a sigma value in (Refer Time: 24:00)?

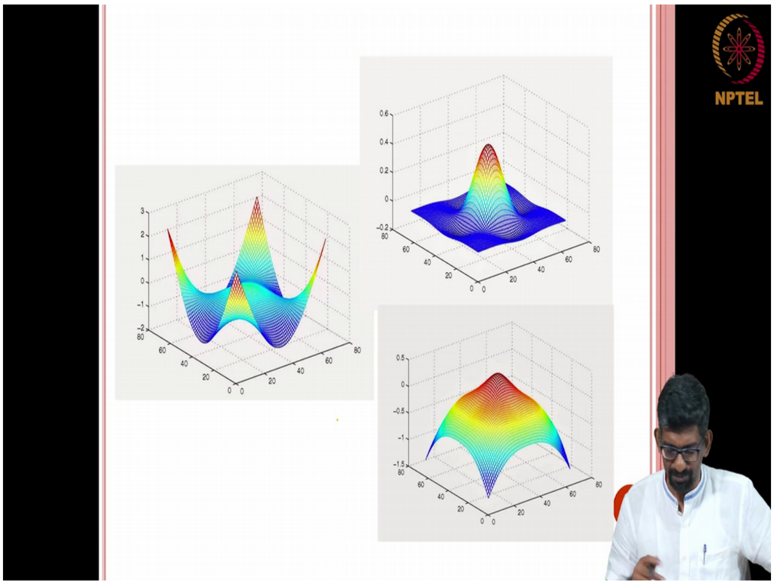
How do you choose, oh for the Gaussian distribution. I am I am sure there are there are ways to fit that and, but that is also a parameter. That is what I told you know like, yeah, that is also a parameter that you might want to do. But, the point and that is, but it is fixed for all of them ok, but in kriging what happens is it yeah, it can change for each one that is the difference between green and this ok. So, the mu also you can vary but, it is now it is one less computation if you put your mu to be at the centre.

(Refer Slide Time: 24:40)



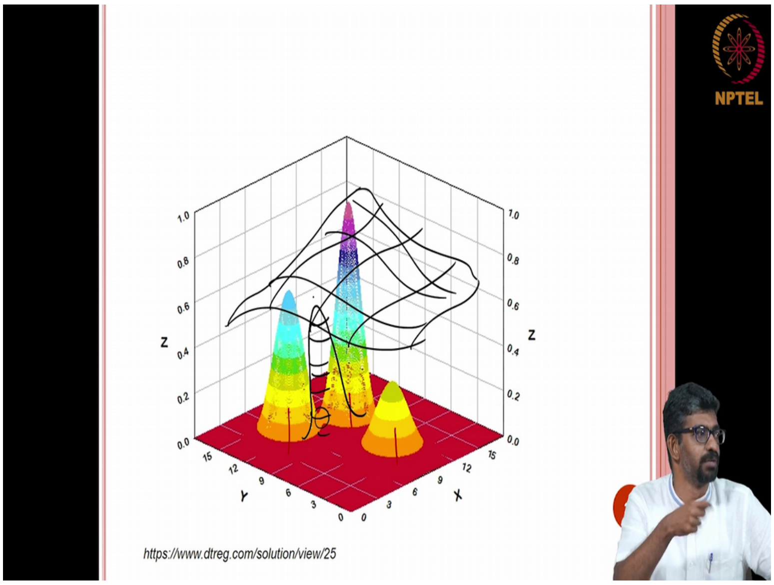

These are different such examples. For instance, in this particular case this is the actual line that they wanted to approximate. So, as you can see there are too many of these guys, it is not varying, right. So, all of them are of the same height you add them up you will get this guy and then at this point I am using multiple set stuff, ok.

(Refer Slide Time: 25:03)



And, people have done. So, this is just 1D so far we saw, 2D there are also these kind of different functions are available. This is our bi-variate Gaussian would look like, I do not know this is a quadric probably and this is quadric this is inverse quadric probably quadric means you understand it quadric is 2, quadratic 4.

(Refer Slide Time: 25:27)




And so, 3D how can you do? So, multiple Gaussians these are the different centres this is one centre, this is one centre, this is one centre, and then I can add them up. So, what I might get is I am might get a function in space, you understand, right. So, at this point it is going to add up this value in this value or I could have another function also here another Gaussian function. I could have then it will get added up.

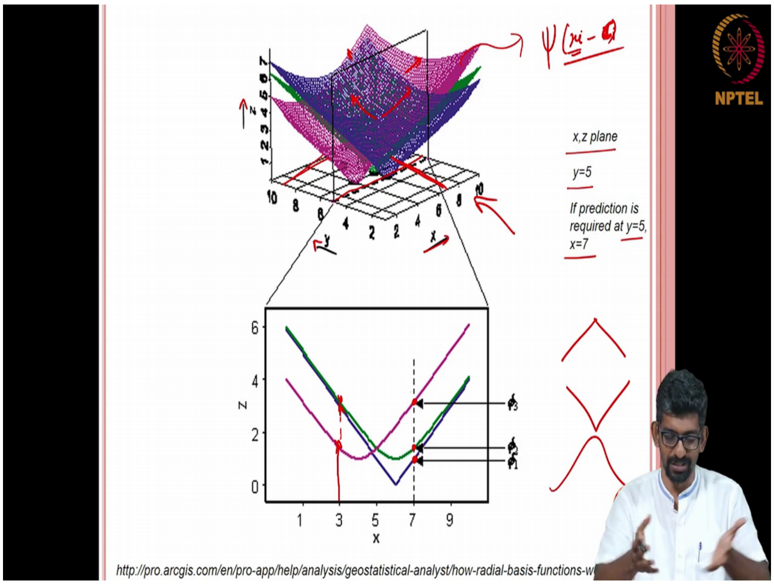
One example that I give people is imagine that there are people sitting in this room and then I throw a bedspread on not all of your of the same height, ok. So, he is taller, he is shorter, she is shorter, she is taller, ok. So, when I put bedspread on all of you, it is going to go like this and come ok. Now, randomly I took 10 people out of the thirty people who are sitting here randomly I do not have a bias, I am not taking only these people are I randomly did, what will happen? It will dip further, and it might go a little more on that side and it might dip further, depending on how there are two people who are taller and taller on that side then the function will go up like that.

So, if I were to approximate the function, imagine that I can just approximate your head this side and this side I can approximate your head. Basically the curvature is what I am trying to do, and I just take the height of you and then I take the curvature that is one thing I add that up to the second guy, I add a to the third guy I add that to the nth guy then I can approximate how the bedspread looks like linear superposition. I am just going to add each one of you and it might become an over fit because that is what was the actual one, but then I removed 10 people. So, what happens is there is an artefact, there is a sudden dip and then it goes up, but that is what that is that was not the original function. Just because I pull this person from there it went like that pulling this person is equivalent to not having the data point.

But, now when you fit what you will fit you will fit it like this, that becomes an over fit, but you do not know that it is an over fit because you do not know the original function. So, that I mean there are there are no ways to eliminate that, but you can address that ok. So, this example tells you that it is not enough if you look at it only from a output response perspective. You also have to look it up from an input response perspective, ok. You had 25 people here and you had only 5 people in this, and then you are approximating the function as a whole of 25 people that is not because in this side of the column there were only 5 people as against 10 people in each of these columns, ok.

So, here it is sparse representation of this column compared to these two people. So, your output should record that value with respect to your excess. Currently we do not, radial basis in one sense does, but still it is centre. It says in my neighbourhood how much does it weight and it does really not capture that though it is a nuclear in distance and all that it does not really capture it fully, ok. So, kriging introduces a little bit more information into that we will see how to does.

(Refer Slide Time: 29:09)




Just to give you an idea in 3D how this one was not 3D, this is still 2D for kriging for radial basis idea so, x, y and z, ok. So, we are looking at a x, y as input and let us say z as the output these are the three surfaces that I have that I have fitted, ok. So, this is a classical example like the way that I told you right for the triangular first linearly varying problem I said like imagine that your basis functions each basis function extends across your design domain that is the case here. So, it is not stopping anywhere it is going across your design domain each one of them, ok.

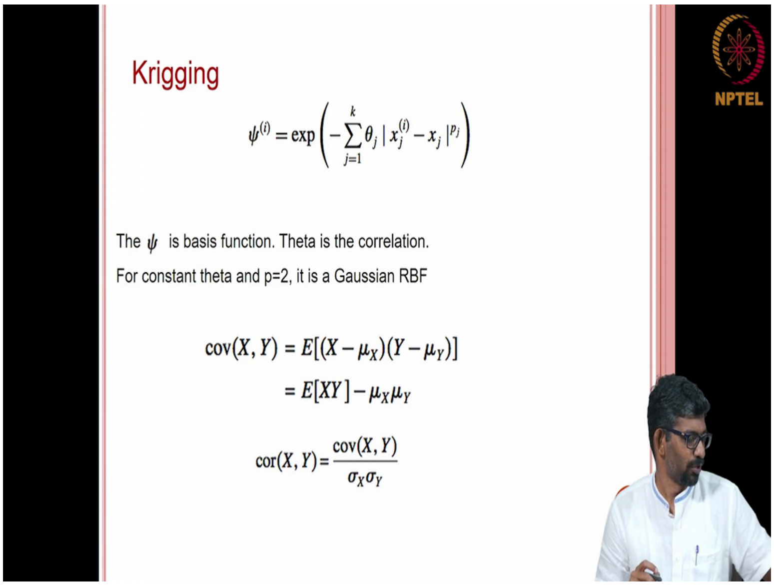
So, this is centre at some other point the red is centre at some other point, the green is centred at some other point, ok, but all of them are running across your design ok. So, now imagine that I am using a cutting plane here ok, then the visualization from x side you are you are seeing from this side ok, it is only a projection that I am putting here x versus z. So, it will look like this. You understand right, it is perpendicular to the plane of the our monitor and then you can get this guy. So, I am only looking at their projection.

So, if I am interested in y equals 5 and x equals 7 which means y equal to sorry y equal to 5, it is here and x equal to 7, it is somewhere here, ok. So, I am just taking I am putting the plane at y equal to 5 and that is what I get and then I take x equals 7. So, what will be my final value of the function? It will be a sum of these guys. If I add all these three functions I will get the representative function in space, agree right? So, if you go into the specifics of how do you do that I am just taking I am fixing why I am just cutting it so, you are able to see the cross section that is how it looks like.

For a particular x, I will take the corresponding values, I will add up the z, similarly for 3 I will be able to add up those values ok, but this is at y equal to 5. If I wanted y equals I mean I cannot do this here but I am just saying if I wanted y equal to 9, I mean get a slightly different the way in which it looks like yeah I will all get different these curves and accordingly at my x it will take different values. Imagine that you have three sheets and then you are going to cut through wherever you want and the cross sections will vary you just need to add them ok, but you do not need you will not this is only for understanding. This is only for visualization purpose, because this function is nothing, but a psi and it says x i minus x of c whatever.

So, I know the c i you just tell me where the x i is I know which basis function I need to take and then with respect to that centre I will estimate what this psi is, that is all. The psi is just some representation; it will be some x cube plus y squared plus something I will just do that straight for.

(Refer Slide Time: 32:56)




See, this could also just be a subset of the whole problem, ok. It keeps on increasing also I know, ok. Fine.