Neural Networks for Signal Processing-I

Prof. Shayan Srinivasa Garani

Department of Electronic System Engineering

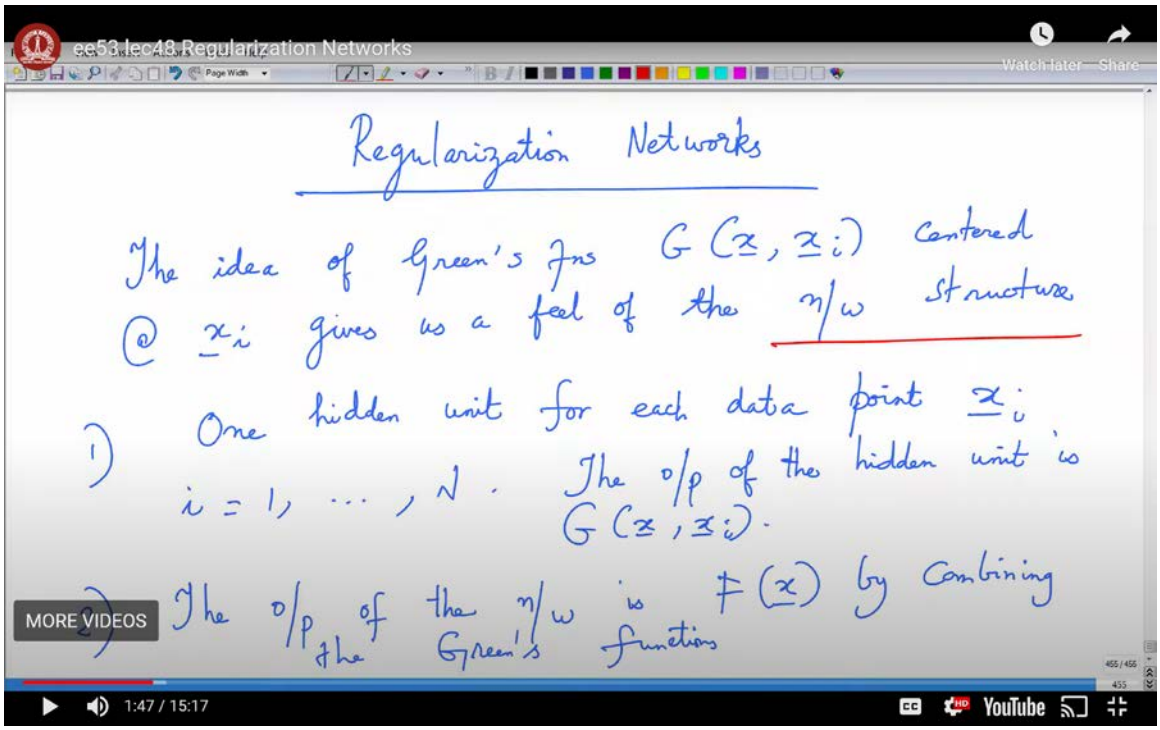
Indian Institute of Science, Bengaluru

Lecture – 48

Regularization Networks

We previously delved into Green's functions, exploring how they can be intricately linked to the solution of the approximating function we were pursuing in the learning problem. Building on this foundation, let's now transition into the study of regularization networks. This approach involves replacing the Gaussian-based units typically found in the hidden layers of a Radial Basis Function (RBF) network with general Green's function units. To understand this transformation, we'll simplify the process and examine how this network structure emerges.

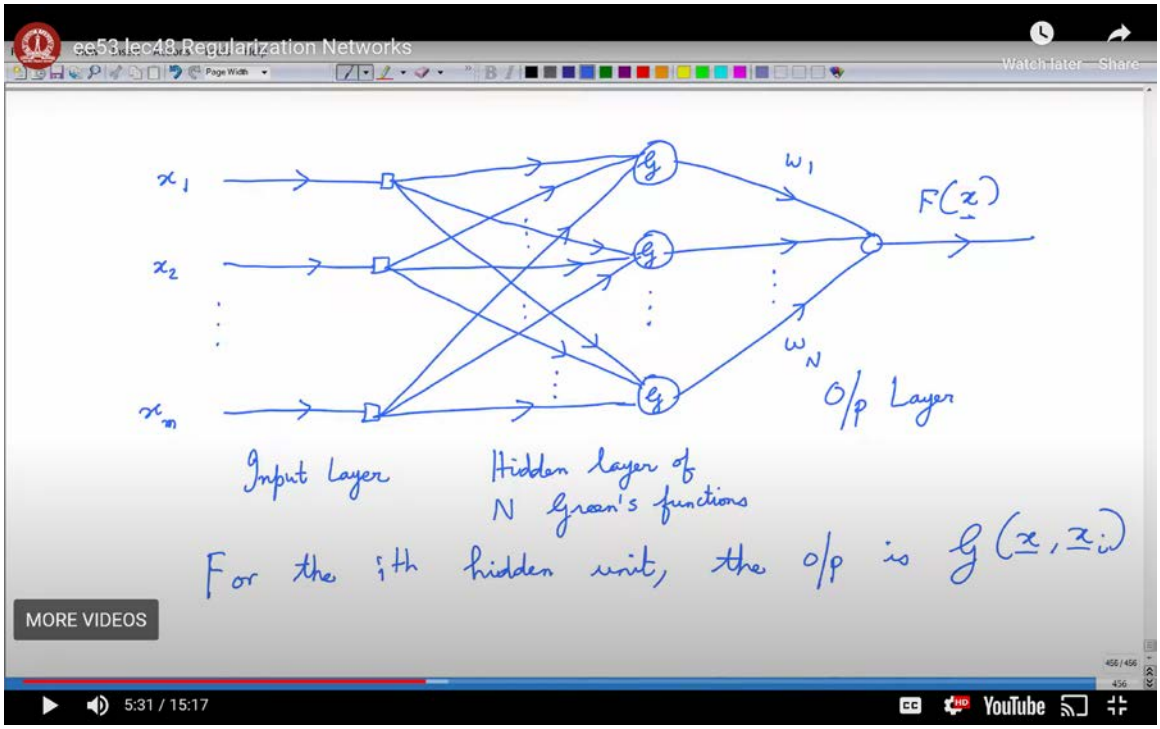
(Refer Slide Time: 01:47)



The concept of Green's functions $g(x, \xi)$ centered at ξ offers profound insight into the network's architecture, it is a pivotal element. Specifically, there is one hidden unit corresponding to each data point ξ_i, where i = 1 to n. The output of each hidden unit is essentially $g(x, \xi_i)$, and the overall output of the network is the function f(x), which is formed by combining these Green's functions, each centered at the respective data points.

Let's visualize this structure. Imagine input points $x_1, x_2, \ldots, x_m$. These points represent our input layer. For simplicity, I'll place these units here. Next, consider our hidden units, which consist of the Green's functions. Each input is connected to these hidden units, forming a fully connected network. The hidden layer comprises n Green's functions, and at the output, we linearly combine all these Green's functions with weights w_1 through w_N to yield the approximating function f(x). This forms the output layer of the network.

(Refer Slide Time: 05:31)

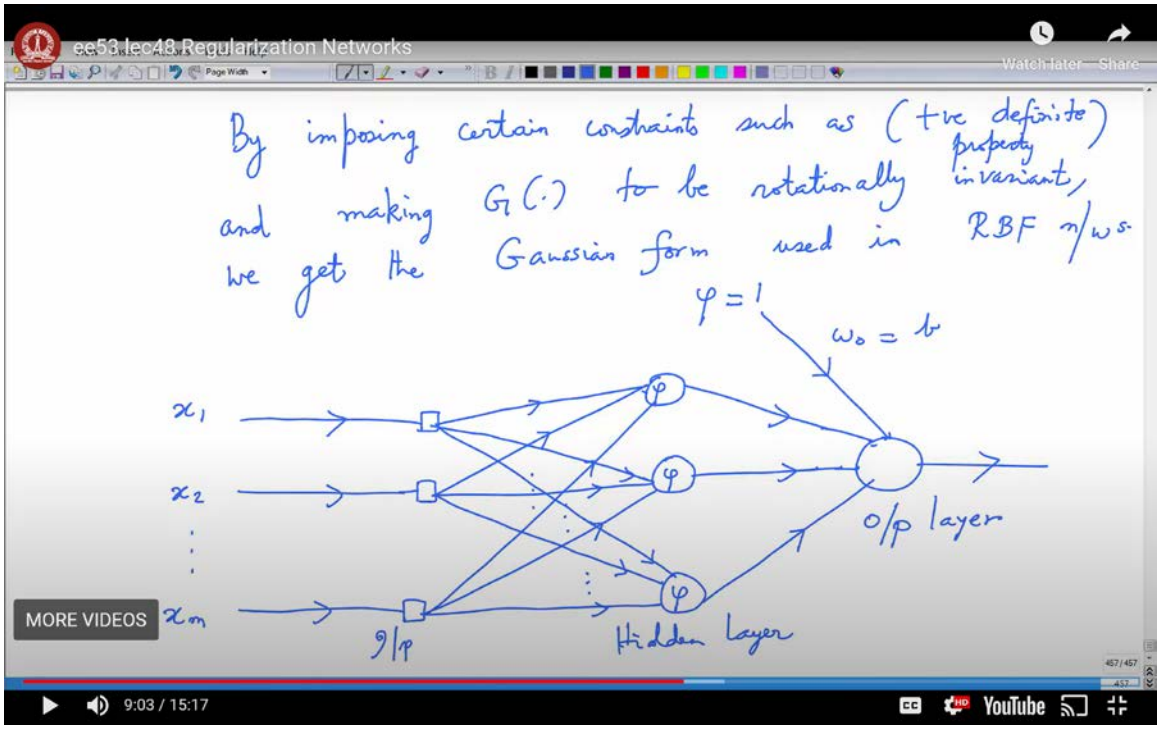


To reiterate, there is a hidden unit for each data point ξ_i, and the output of the i-th hidden unit is specifically g(x, ξ_i). In our previous discussion on RBF networks, we noted that imposing certain constraints on the Green's functions, such as ensuring positive

definiteness and rotational invariance, leads to the Gaussian form. This Gaussian form is the cornerstone of RBF networks, explaining why they are termed "radial basis functions."

Now, to recap and illustrate this network: we have input points $x_1, x_2, \ldots, x_m$, each connected to RBF units consisting of functions ϕ, which adopt this Gaussian form. The network is fully connected, with each output from the RBF units combined at the output layer, where a bias can also be introduced, forming the final output. Here, the critical distinction is that instead of using a general Green's function, we represent g as ϕ, which operates on the norm of the vector x - ξ_i. This difference is key to understanding the network's behavior.

(Refer Slide Time: 09:03)

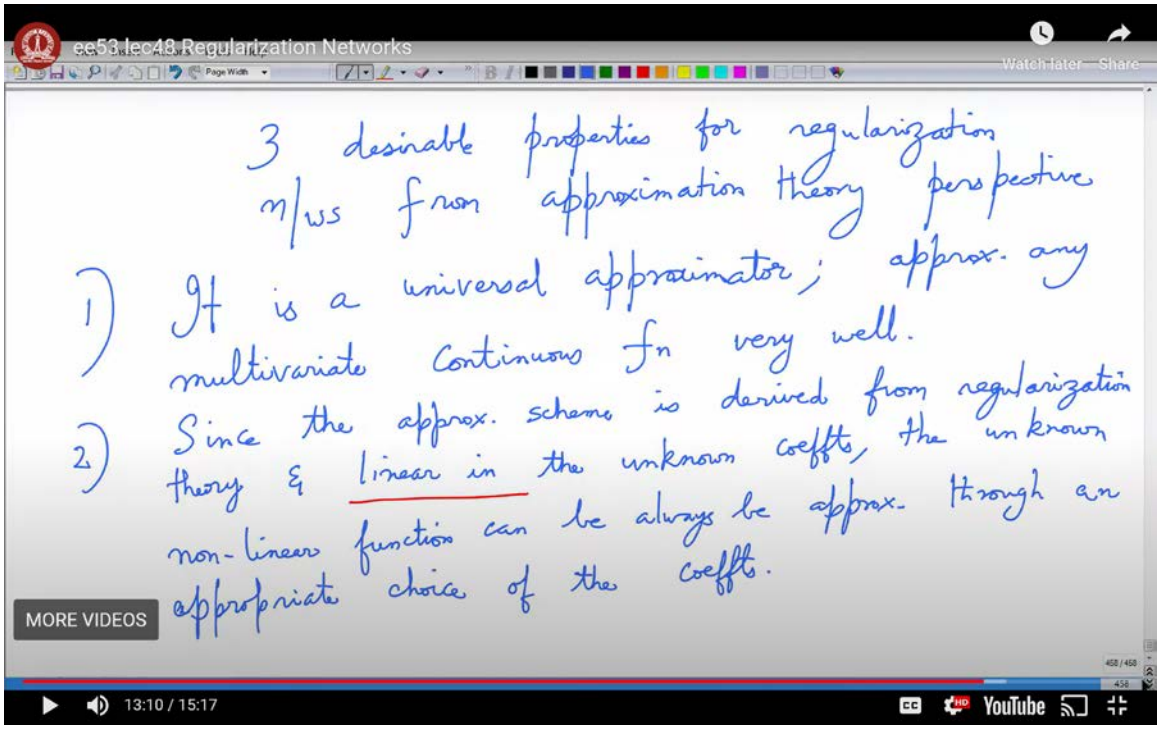


Now, let's examine this from the perspective of approximation theory. What desirable properties should this regularization network possess? There are three key properties that we should focus on, each critical to the effectiveness of the network.

First and foremost, from an approximation theory standpoint, the regularization network is a universal approximator. This means it has the remarkable capability to approximate any

multivariate continuous function with high accuracy, exactly what we seek in a robust model. Since the approximation scheme is rooted in regularization theory and is linear with respect to the unknown coefficients, it allows us to approximate any unknown nonlinear function by selecting the appropriate coefficients.

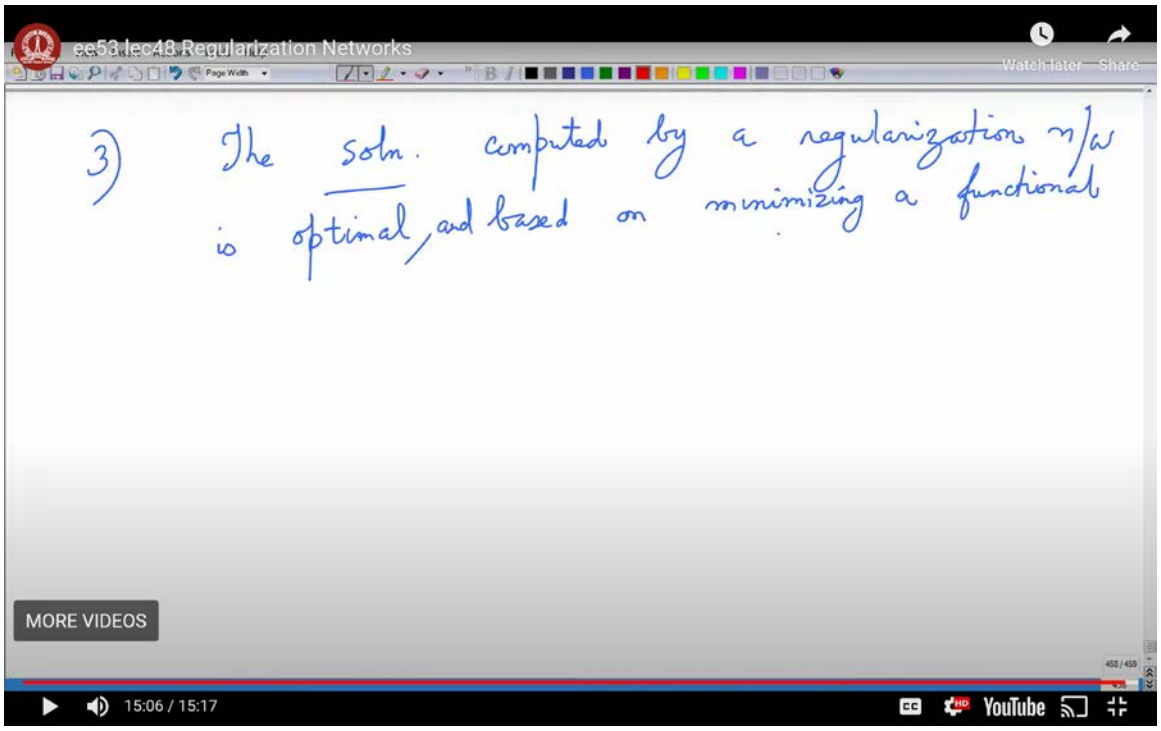
(Refer Slide Time: 13:10)



Let's underscore this crucial point: the approximation scheme is derived from regularization theory and is linear in the unknown coefficients. If the scheme weren't linear, we'd be forced to resort to nonlinear solution techniques, which can become exceedingly complex. The linearity ensures that the unknown nonlinear function can always be approximated through an appropriate choice of weights, those coefficients that combine the outputs from the hidden units.

Secondly, and of equal importance, the solution computed by a regularization network is optimal. This optimality is achieved by minimizing a functional that measures the difference between the solution and the training examples, all while adhering to certain regulatory conditions influenced by the geometry of the problem.

(Refer Slide Time: 15:06)



These three essential properties are inherent in regularization networks:

1. Universal Approximation: The network can approximate any multivariate continuous function with precision, a property shared with Multilayer Perceptrons (MLPs) and Radial Basis Functions (RBFs).

2. Linear in Unknown Coefficients: The scheme's linearity in unknown coefficients simplifies the process, allowing for the appropriate solution of these coefficients or weights.

3. Optimal Solution: The network’s solution is derived by minimizing a functional, specifically the Tikhonov functional, ensuring that it is the best possible outcome under the given conditions.

These properties make regularization networks highly desirable in various applications. With this understanding, we conclude this module and will continue to build on these concepts in our future work.