Introduction to Neural Networks

Johns Hopkins University
Engineering for Professionals Program

605-447/625-438

Dr. Mark Fleischer

Module 13.2: Radial Basis Functions

What We've Covered and Will Cover:

- Clustering:
  - The nature of the problem
  - Defining it
  - Modeling it as an optimization problem
- Radial Basis Functions
  - Their connection to clustering
  - An Application of RBFs

Heuristics to the Rescue!

- We can make intelligent guesses as to initializing cluster center locations.
- Then add additional cluster centers up to some number.
- How can neural networks be used here?

Basis Functions

The first level of processing regarding weights and inputs.

Up to now, we've only really considered a Linear Basis Function for a perceptron:

$$A_i = \sum_j w_{ij} x_j + \theta_i$$

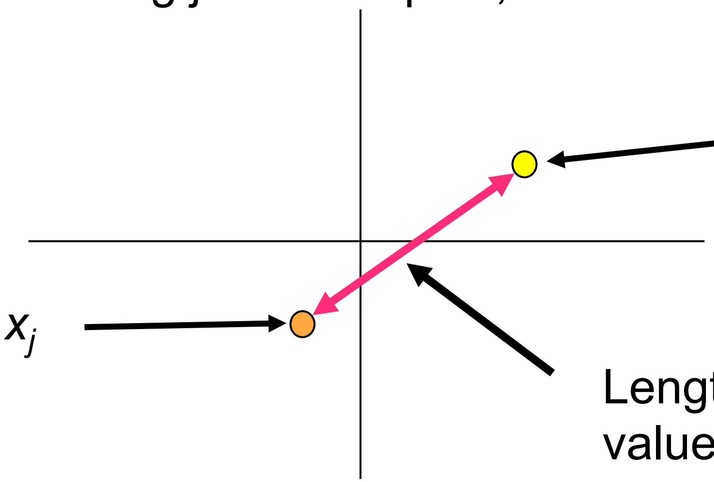
but there are others!

Radial Basis Functions

An Example:

$$A_i = \sqrt{\sum_{j=1}^{n} \left(w_{ij} - x_j\right)^2}$$

With a node having just two inputs, it looks like this:



Radial Basis Functions

- Another similar basis function, but a bit more general:

$$A_i = \sqrt{\sum_{j=1}^{n} \alpha_{ij} \left(w_{ij} - x_j \right)^2}$$

This is an elliptic basis function.

Observations

- The closer the inputs are to the weights, the smaller the value of A_i
- If we want a perceptron to 'fire' when inputs are close to the weights, what kind of activation function should we have?
- We could use a Gaussian function as the activation function.
- We could also use some threshold post-processing to make the output 1 when the inputs are within a certain radius of the weights.

Some Variations of its use

$$f_i(\mathbf{x}) = \frac{e^{-\sum_j (x_j - w_j)^2 / 2\sigma^2}}{\sum_i e^{-\sum_j (x_j - w_j)^2 / 2\sigma^2}}$$

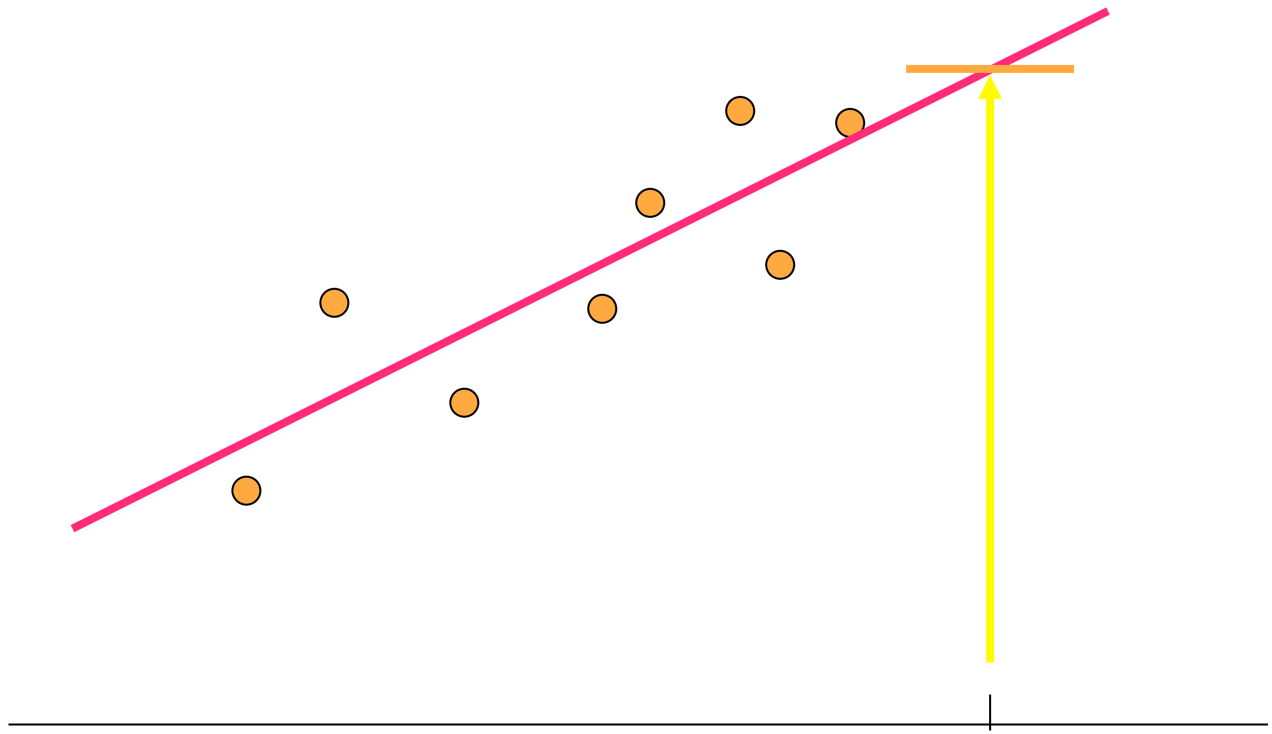
Utility of RBFs

- Can be used to approximate many functions.
- From Rojas: "The main difference between networks made of radial basis functions and networks of sigmoidal units is that the former use locally concentrated functions as building blocks whereas the latter use smooth steps."

Regression

- Regression is useful for prediction
- Also useful for function approximation
- We assume data is produced by some underlying model and perturbed by random variation (noise)
- Can NNs be useful here? Can different basis functions be useful here?

Regression—a Review



Generalized Regression Neural Network

- Attempts to perform generalized regression calculations
- Useful for many types of regression problems
- Uses four layers:

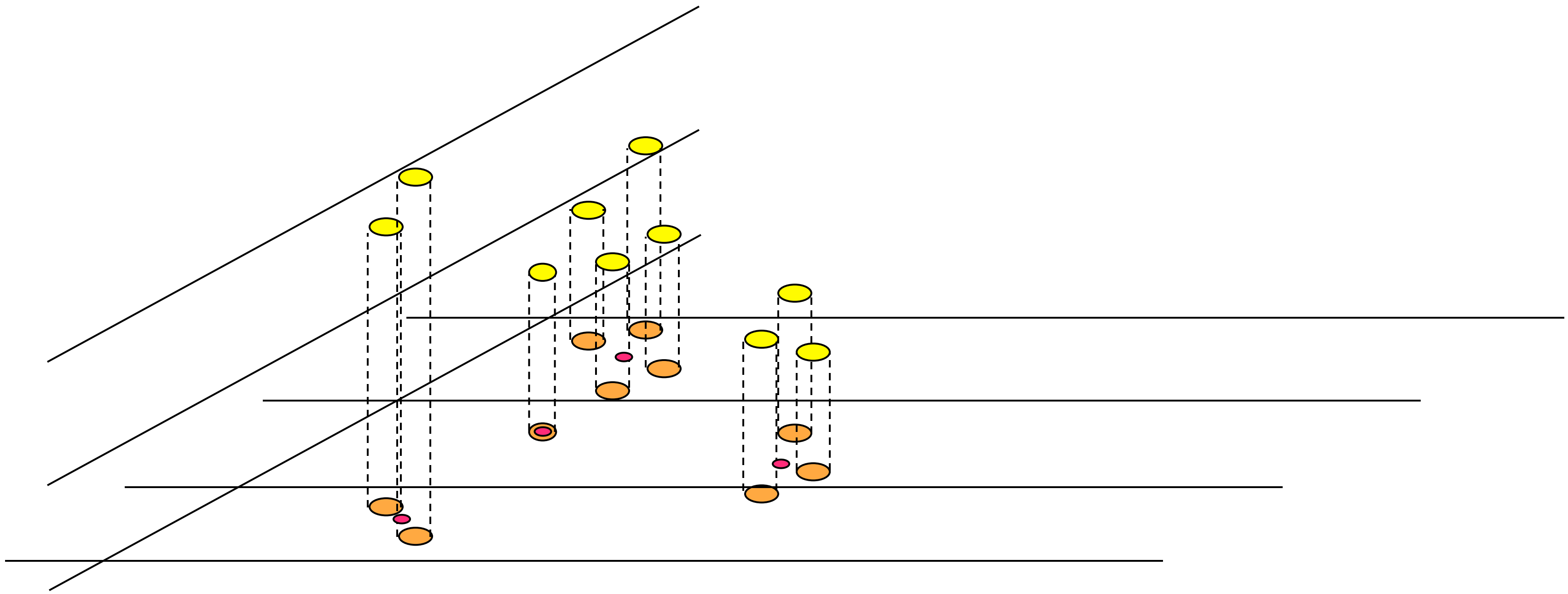
Generalized Regression Neural Network

$$E\left[y|X\right]=\frac{\int_{-\infty}^{\infty} yf(X,y)dy}{\int_{-\infty}^{\infty} f(X,y)dy}$$

$$D_i^2=\left(\mathbf{X}-\mathbf{X}_i\right)^{\mathbf{T}}\left(\mathbf{X}-\mathbf{X}_i\right)$$

$$\hat{Y}(\mathbf{X})=\frac{\sum_{i=1}^{n}\mathbf{Y_i}\exp\left(-\frac{D_i^2}{2\sigma^2}\right)}{\sum_{i=1}^{n}\exp\left(-\frac{D_i^2}{2\sigma^2}\right)}$$

Generalized Regression Neural Network



Generalized Regression Neural Network

- From Specht:

Represent exemplars or cluster centers. They could be based on RBFs with Gaussian activation.

Performs a weighted sum of the Pattern Unit outputs. Weights correspond to number of observations associated with that cluster.

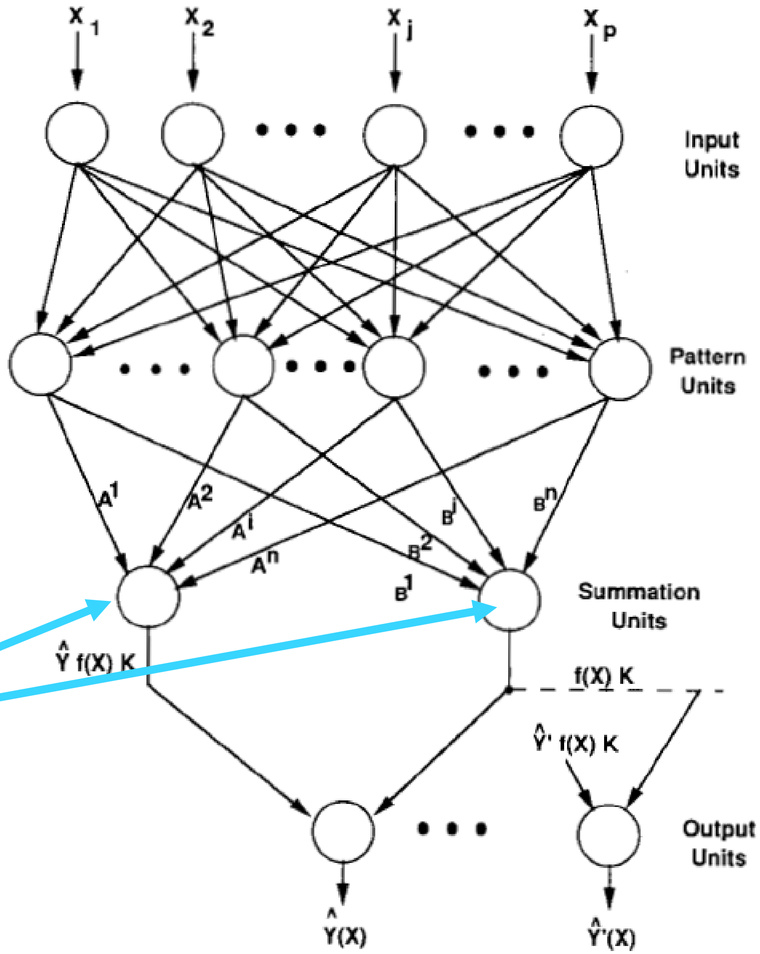


Fig. 1. GRNN block diagram.

Generalized Regression Neural Network

- Two Summation Units:

  $f(\mathbf{X})K$: "sums the outputs of the pattern units weighted by the number of observations each cluster center represents."

  $\hat{\mathbf{Y}} f(\mathbf{X})K$:This unit "multiplies each value for a Pattern Unit by the sum of samples [Y_i] associated with Cluster center [X_i]."

Generalized Regression Neural Network

- Applications of GRNNs:
  - General regression problems,
  - Adaptive control,
  - Filtering,
  - …