Optmization Methods for Machine Learning - Fall 2015

Assignment # 2 - Generalized RBF Network

Laura Palagi

Dipartimento di Ingegneria informatica automatica e gestionale A. Ruberti

Sapienza Università di Roma

Posted on November 2, 2015- due date November 20, 2015

In this assignment you will implement a generalized RBF neural network for regression. We want to reconstruct the function in the region $[0,1] \times [0,1]$. (Franke's function see http://www.sfu.ca/ ssurjano/franke2d.html)

$$\begin{aligned} f(x) &= 0,75 \exp\left(-\tfrac{(9x_1-2)^2}{4} - \tfrac{(9x_2-2)^2}{4}\right) + 0,75 \exp\left(-\tfrac{(9x_1+1)^2}{49} - \tfrac{(9x_2+1)}{10}\right) \\ &+ 0,75 \exp\left(-\tfrac{(9x_1-7)^2}{4} - \tfrac{(9x_2-3)^2}{4}\right) - 0,2 \exp\left(-(9x_1-4)^2 - (9x_2-7)^2\right) \end{aligned}$$

The data set $\{(x^i, y^i) : \ x^i \in \mathbb{R}^2, \ y^i \in \mathbb{R}\}$ is obtained by sampling on 100 random points x^i the function and adding a uniform noise, i.e. $y^i = f(x^i) + \varepsilon^i$ and ε^i is a random number in $[0,1]$ (use the rand function in matlab).

The data set will be divided into a training set and a test set (choose percentage of training data between 70-80%).

As RBF function $\phi(\cdot)$ you can choose one of the following

- either the Gaussian function

$$\phi(\|x - c_j\|) = e^{-(\|x-c_j\|/\sigma)^2} \quad r > 0 \tag{1}$$

  with derivative

$$\nabla_{c_j} \phi_(\|x - c_j\|) = \frac{2}{\sigma}^2 e^{-(\|x-c_j\|/\sigma)^2} (x - c_j)$$

- or the *Inverse Multiquadric*

$$\phi(\|x - c_j\|) = (\|x - c_j\|^2 + \sigma^2)^{-1/2}, \quad r > 0 \tag{2}$$

  with derivative

$$\nabla_{c_j} \phi(\|x - c_j\|) := (\|x - c_j\|^2 + \sigma^2)^{-3/2} (x_i - c_j)$$

You need to set the number of RBF units N of the hidden layer.

Question 1. Write a program (please attach a printout) which

1. implements the error function and, if needed, its gradient

$$E(w) = \frac{1}{2}\sum_{i=1}^{P}\left(\sum_{j=1}^{N} w_j \phi(\|x^i - c_j\|) - y^i\right)^2 + \frac{\rho}{2}\|w\|^2 + \frac{\rho}{2}\|c\|^2,$$

   where $\rho_1, \rho_2 > 0$ are regularization parameters to be chosen. Use a matlab routine of the optimization toolbox for its minimization with respect to both (w, c).

2. produces a plot of the function obtained.

3. Evaluate the value of the training error and of the test error.

4. Analyse the occurrence of overfitting/underfitting varying the number of neurons N and of the parameters ρ_i.

Question 2. Write a program (please attach a printout) which

1. implements a method with unsupervised selection of the centers. Sselect the centers randomly on the P points of the training set or my a cluster algorithm. Choose the weights by minimizing the convex quadratic function

$$E(w) = \frac{1}{2}\sum_{i=1}^{P}\left(\sum_{j=1}^{N} w_j \phi(\|x^i - c_j\|) - y^i\right)^2 + \frac{\rho_1}{2}\|w\|^2,$$

   using a suitable matlab routine of the optimization toolbox. Set the regularization parameter $\rho_1 > 0$ at the value you defined in Question 1.

2. Evaluate the value of the training error and of the test error.

Question 3. Write a program (please attach a printout) which

1. Implements a supervised selection of both weights and centers using a two block decomposition method which alternates the minimization with respect to weights and centers. Set the regularization parameter $\rho_1, \rho_2 > 0$ at the value you defined in Question 1. Use matlab routine of the optimization toolbox for solving the the two minimization problems respectively with respect to centers and to weights

2. Evaluate the value of the training error and of the test error.

3. Compare the behaviour of the results with respect to the other optimization methods implemented and the performance of the algorithms.

Please note that as optimization routine you can use also a minimization algorithm developed by yourself.