Relevance Vector Machine (RVM) Implementation

July 12, 2013

1 Problem

To implement the RVM method for regression, as discussed in the chapter 7 of the book [1], we require a dataset which can be produced using the following matlab function :

```
function [ X , T , Y ] = build_data_sin( interval , n , sigpow )
        step = (interval(2)-interval(1))/(n-1);
        X = (interval(1):step:interval(2))';
        Y = sin(X);
        T = awgn(Y,sigpow);
end
```

Here $\mathbf{x} = (x_1, \ldots, x_N)^T$ is the vector containing the features, $\mathbf{y} = (y_1, \ldots, y_N)^T$ is the vector containing the true responses and $\mathbf{t} = (t_1, \ldots, t_N)^T$ contains the noisy responses. The goal of this implementation is to produce $\mathbf{y}$ given $(\mathbf{x}, \mathbf{t})$. The dataset and an RVM solution to the problem is shown in figure 1.1. Now assuming that $\tilde{\mathbf{y}}$ is the solution produced by the RVM method, the error of this solution is the computed using the following equation,

$$E(\mathbf{y}, \tilde{\mathbf{y}}) = \sum_{i=1}^{N} (y_i - \tilde{y}_i)^2. \tag{1.1}$$

While the implantation of the RVM method is rather straight forward, considering the following conditions in the problem can make the problem more interesting and provide better analysis of the method :

- Replace $sin(x)$ with more complex functions and see how the number of support vectors change as the function becomes more complex.
- How different kernels respond to the data and which kernel produces the least error according to Eq. 1.1. To do this, make sure the dataset remains the same while different kernels are applied to it.

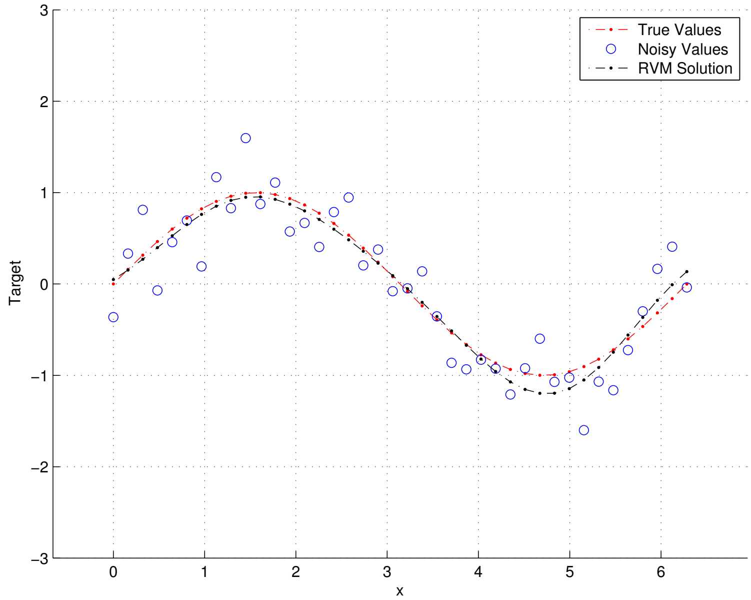

Figure 1.1: This figure shows, how the noisy data ($\mathbf{t}$) relates to the true values ($\mathbf{y}$). In this problem, while the true values are hidden from the method the task of the RVM method is to estimate them given $\mathbf{x}$ and $\mathbf{t}$.

- Here, the features x_i are sampled from an equally spaced grid. It would be interesting to see how the number of support vectors changes if the features x_i are also randomly sampled from the interval.

References

[1] Christopher M. Bishop. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006.