Representing Penalties on Basis Expansions as a Polynomial In the Parameters (Working Title)

February 22, 2013

1 Introduction To Basis Function Expansion Methods

One of the major fields in Statistics is that of fitting functions to data. A very large class of methods are known as *Basis Function Expansion Methods*, where we estimate the output as a weighted sum of basis functions. These models generally take a form as follows:

$$\mathbb{E}[y|t] = c_0\phi_0(t) + c_1\phi_1(t) + \cdots + c_N\phi_N(t)$$

Here we have our finite set of basis functions $\{\phi_k\}$ along with our parameters $\{c_k\}$, $\mathbb{E}[y|x]$ is the conditional expectation of our independent variable y, given our dependent data x. Roughly $\mathbb{E}[y|x]$ can be thought of as our best estimate of the value of y given x.

Basis Function Methods might seem a little abstract, but they are ubiquitous. For example, Simple Linear Regression is an example of a basis function method. Take two basis functions $\{1, t\}$, so that $\hat{y}(t)$, our estimate of y given the value t, can be written as a linear combination of the two:

$$\mathbb{E}[y|t] = \hat{y}(t) = c_0 + c_1 t$$

This is the form of a simple linear regression model. If our basis consists of just the single constant function $\phi(t) = 1$ then we get a model of the form:

$$\hat{y} = c_0$$

In this case we would generally use the mean or the median of the y values as our estimate of c_0. If we go in the other direction and add a quadratic function t^2 we get a quadratic regression model:

$$\hat{y}(t) = c_0 + c_1 t + c_2 t^2$$

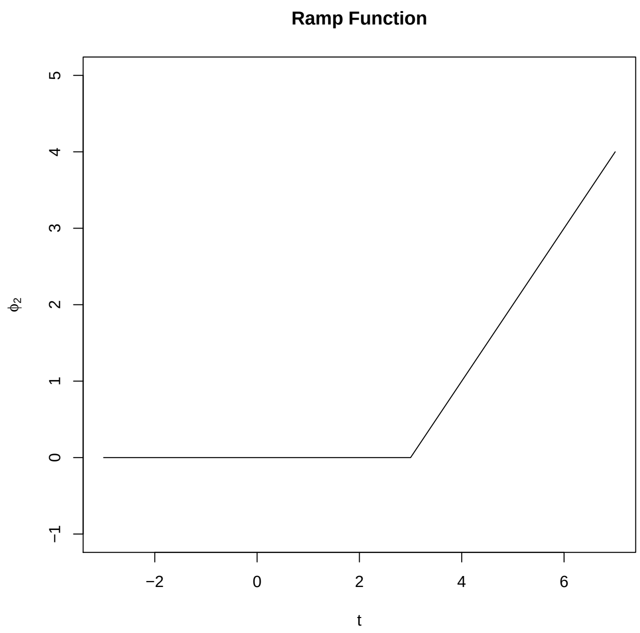

Figure 1: The Ramp Function ϕ_2

There are other choices of basis besides monomials we could use. We could have a basis of two functions $\{\phi_1, \phi_2\}$ Where the two functions are defined as follows:

$$\phi_1(t) = 1$$
$$\phi_2(t) = \begin{cases} 0 & \text{if } t < \tau \\ t - \tau & \text{if } t \geq \tau \end{cases}$$

The first function is a constant function, the second is a ramp function based at τ.