Regression and Kernel

EE1390- Intro to AI and ML

Sadiq - EE18BTECH11051
Nawaz - EE18BTECH11052

IIT HYDERABAD

March 11, 2019

Regression and Kernels

- Kernel regressions are weighted average estimators that use kernel functions as weights
- Kernel regression is a non-parametric technique in statistics to estimate the conditional expectation of a random variable.

Regressions and Kernels

We consider the approximation of the regression function in terms of a set of basis functions $\{h_m(x)\}, m = 1, 2, 3...., M$:

$$f(x) = \sum_{m=1}^{M} \beta_m h_m(x) + \beta_0 \tag{1}$$

To estimate β *and* β_0 we minimize

$$H(\beta, \beta_0) = \sum_{i=1}^{N} V(y_i - f(x_i)) + \frac{\lambda}{2} \sum \beta_m{}^2 \tag{2}$$

For some general error measure $V(r)$. For any choice of $V(r)$, the solution $\hat{f}(x) = \sum \hat{\beta}_m h_m(x) + \hat{\beta}_0$ has the form
$\hat{f}(x) = \sum_{i=1}^{N} \hat{a}_i K(x, x_i)$

Solution

Where $K(x, y) = \sum_{m=1}^{M} h_m(x) h_m(y)$
We estimate β by minimizing the penalized least square criterion

$$H(\beta) = (\mathbf{y} - \mathbf{H}\beta)^{\intercal}(\mathbf{y} - \mathbf{H}\beta) + \lambda ||\beta||^2 \tag{3}$$

The solution is

$$\hat{y} = \mathbf{H}\hat{\beta} \tag{4}$$

With $\hat{\beta}$ is determined by

$$-\mathbf{H}^{\intercal}(y - \mathbf{H}\hat{\beta}) + \lambda\hat{\beta} = 0 \tag{5}$$

On solving this we can get

$$\mathbf{H}\hat{\beta} = (\mathbf{H}\mathbf{H}^{\intercal} + \lambda\mathbf{I})^{-1}\mathbf{H}\mathbf{H}^{\intercal} y \tag{6}$$

Solution

Here the N X N matrix $\mathbf{HH}^{\intercal}$ is the inner product of the observations i,i'. So $\mathbf{HH}^{\intercal}$ can be written as

$$\{\mathbf{HH}^{\intercal}\}_{i,i'} = K(x_i, x_{i'}) \tag{7}$$

It is easy to show that the predicted values at an arbitrary x satisfy

$$f(\hat{x}) = \mathbf{h}(\mathbf{x})^{\intercal}\beta = \sum_{i=1}^{N} \hat{\alpha}_i K(x_i, x_{i'}) \tag{8}$$

Where $\hat{\alpha}_i = (\mathbf{HH}^{\intercal} + \lambda \mathbf{I})^{-1} y$

Final Slide

THANK YOU