Gaussian Process Regression

Prof. S.R Khare

Poorvi Agrawal

28 July 2015

Regression

Regression analysis is a technique for finding out the relationship between variables. It has a lot of real life application. Be it for forecasting or to predict petrol prices for the next month, regression analysis is used widely and in almost every field. It has been classified based on the dependence of variables: Linear regression, nonlinear, logistic etc. Linear regression is commonly used approach for finding out the relationship between one dependent variable and one or more independent variable.

Gaussian Process Regression

Gaussian Process is a family of distribution where it may or may not depend on time. In Gaussian process, every finite number of random variables follows multivariate normal distribution which is normal distribution in higher dimension. Joint distribution of infinitely many random variables following above criteria is the distribution of Gaussian process. It is mainly defined by two functions: mean and covariance function. In Gaussian Process Regression, rather than specifying a function to a specific model like in linear or quadratic regression, it defines function indirectly so that it would become less parametric which implies fewer assumptions in data set, hence getting more accurate predictions. It has a wide application when it comes to noisy data i.e. unwanted, inaccurate data.

Difference in Regression and Gaussian Process Regression

Regression is considered to be an effective approach for modelling because of its simplicity in implementation and analysis. But when it comes to flexibility, it is of no more advantageous. Regression model can give poor predictions if function is not properly chosen. Gaussian process Regression is better approach than normal regression as fixed there is no need for selecting a function which makes it more accurate and flexible.