Stochastic Learning and Its Application In Hydrology

Pan Baoxiang

November 24, 2014

Outline

- Introduction to Stochastic Learning
- Two Cases
  - ICA
  - Support Vector Machine
- Discussion

Introduction

Definition:Machine learning is a scientific discipline that deals with the construction and study of algorithms that can learn from data. Such algorithms operate by building a model based on inputs and using that to make predictions or decisions, rather than following only explicitly programmed instructions.

Conditions to Use Stochastic Learning

- There is pattern
- There is not explicit mathematical expression
- Available Data

ICA

Independent Component Analysis (ICA) is a computational method for separating a multivariate signal into additive subcomponents. This is done by assuming that the subcomponents are non-Gaussian signals and that they are statistically independent from each other. ICA is a special case of blind source separation.

Problem Statement

Linear Regression of $NDVI$ from P, R, T, EP, P_{log}

Result

ICA transform among all terms.

$n_{independent\_components} = n_{variables}$

$$M_o = 0.0015P_{Mean} - 0.0074Rh + 0.0005P_{Lag} + 0.0349Ta_{Mean} - 0.0006Eps$$

$$\rho^2 = 0.6787$$

$n_{independent\_components} = 2$

$$M_o = 0.0008P_{Mean} - 0.0049Rh + 0.0008P_{Lag} + 0.0144Ta_{Mean} - 0.0001Eps$$

$$\rho^2 = 0.5856$$

$n_{independent\_components} = 3$

$$M_o = 0.0009P_{Mean} - 0.0046Rh - 0.0146Ta_{Mean} + 0.0008Eps$$

$$\rho^2 = 0.4674$$

ICA transform among same dimension terms.

$n_{independent\_components} = n_{same\_dimension\_variables}$

$$M_o = 0.0015P_{Mean} - 0.0074Rh + 0.0005P_{Lag} + 0.0349Ta_{Mean} - 0.0006Eps$$

$$\rho^2 = 0.6787$$

$n_{independent\_components} = 2$

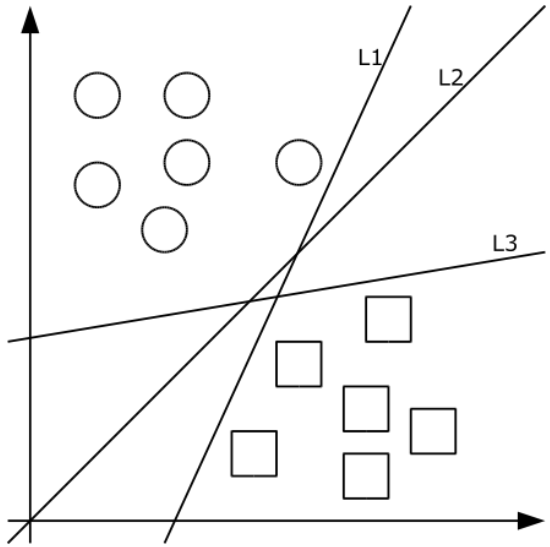
$$M_o = 0.0005P_{Mean} + 0.0011Rh + 0.0007P_{Lag} + 0.0040Ta_{Mean} - 0.0001Eps$$

$$\rho^2 = 0.5800$$

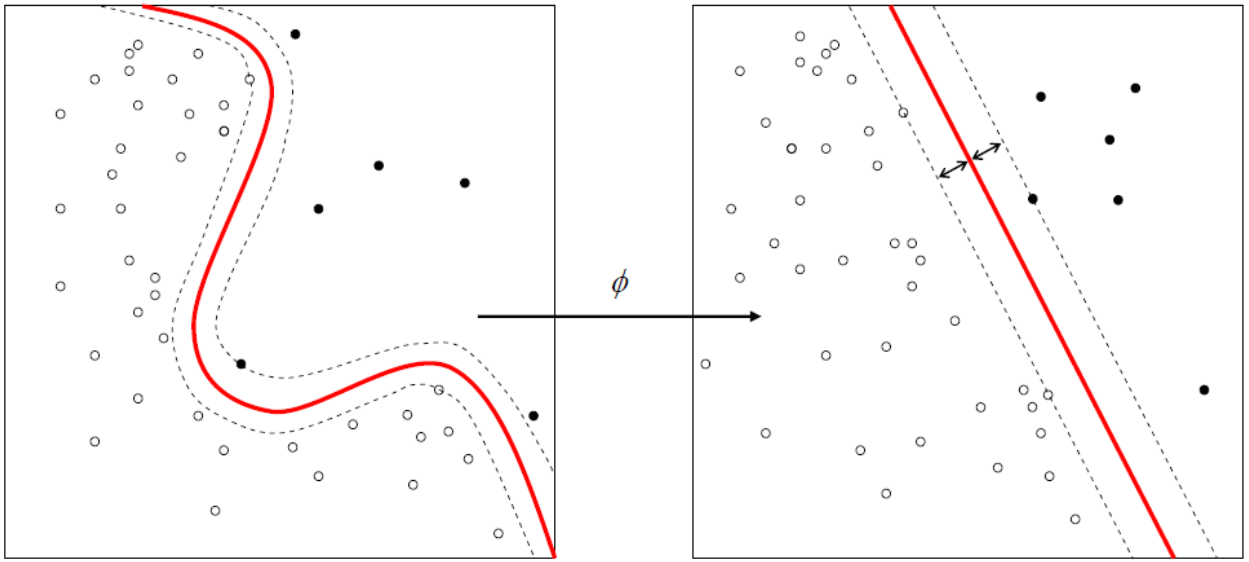
We recommend the last result for an intuitive deduction that the 2 independent components reflect the water and heat control.

Support Vector Machine

A support vector machine constructs a hyperplane or set of hyperplanes in a high or infinite dimensional space, which can be used for classification, regression, or other tasks. Intuitively, a good separation is achieved by the hyperplane that has the largest distance to the nearest training data point of any class (so-called functional margin), since in general the larger the margin the lower the generalization error of the classifier.



Non-linear Condition



Support Vector Regression

Minimize the sum of the empirical risk function and the structural risk function:
empirical risk function:

$$\frac{1}{l}\sum_{i=1}^{l}|y_i - f(x_i)|_\epsilon$$

structural risk function:

$$\frac{1}{2}w^T w$$

Example