Learning through Deterministic Assignment of Hidden Parameters

Zongben Xu

Abstract

Supervised learning boils down to determining hidden and bright parameters in a parameterized hypothesis space based on a finite number of input-output pairs. The hidden parameters are those parameters that determine the hidden predictors or nonlinear mechanism of an estimator, while the bright parameters are those characterizing how the hidden predictors are linearly combined or the linear mechanism. In traditional learning paradigms, the hidden and bright parameters are not distinguished and trained simultaneously in one learning process (or saying, in one-stage). Such an one-stage learning (OSL) brings a benefit of theoretical analysis but suffers severely from the very high computation burden. To overcome this difficulty, the two-stage learning scheme (TSL), that could be named as the learning through random assignment of hidden parameters (LRHP), was developed in past years, which assigns randomly the hidden parameters in the first stage and determines the bright parameters by solving a linear least squares problem in the second stage. LRHP works well in many applications but suffers from the uncertainty problem: Its performance can be guaranteed only in a certain statistical expectation sense.

In this talk, we report a new scheme of TSL: The learning through deterministic assignment of hidden parameters (LDHP). Motivated from the study on the classical hard sphere problem in mathematics, we propose to deterministically take the hidden parameters as the tensor product of the minimal Riesz energy points on sphere and the best packing points in an interval. We theoretically show that with such deterministic assignment of hidden parameters, LDHP shares the same generalization performance with that of OSL, i.e., does not degrade the generalization capability of OSL. Thus LDHP provides an effective way of overcoming both the very high computation burden in OSL and the uncertainty difficulty in LRHP. We present a series of simulation and application examples to support this advantage and the outperformance of LDHP, as compared with a typical one-stage algorithm — Support Vector Regression (SVR) and a typical LRHP algorithm — Extreme Learning Machine (ELM). The study reported here paves a new road to simply and more efficiently solve any supervised learning problem.

Zongben Xu
Institute for Information and System Sciences
School of Mathematics & Statistics
Xian Jiaotong University
Xian, 710049, China
e-mail: `zbxu@mail.xjtu.edu.cn`