Regularized Learning in Reproducing Kernel Banach Spaces*

Jun Zhang

Abstract

Regularized learning is the contemporary framework for learning to generalize from finite samples (classification, regression, clustering, etc). Here the problem is to learn an input-output mapping $f : X \to Y$ given finite samples $\{(x_i, y_i), i = 1, \dots, N\}$. With minimal structural assumptions on X, the class of functions under consideration is assumed to fall under a Banach space of functions B. The learning-from-data problem is then formulated as an optimization problem in such a function space, with the desired mapping as an optimizer to be sought, where the objective function consists of a loss term $L(f)$ capturing its goodness-of-fit (or the lack thereof) on given samples $\{(f(x_i), y_i), i = 1, \dots, N\}$, and a penalty term $R(f)$ capturing its complexity based on prior knowledge about the solution (smoothness, sparsity, etc). This second, regularizing term is often taken to be the norm of B, or a transformation ϕ thereof: $R(f) = \phi(\|f\|)$. This program has been successfully carried out for the Hilbert space of functions, resulting in the celebrated Reproducing Kernel Hilbert Space methods in machine learning. Here, we will remove the Hilbert space restriction, i.e., the existence of an inner product, and show that the key ingredients of this framework (reproducing kernel, representer theorem, feature space) remain to hold for a Banach space that is uniformly convex and uniformly Frechet differentiable. Central to our development is the use of a semi-inner product operator and duality mapping for a uniform Banach space in place of an inner-product for a Hilbert space. This opens up the possibility of unifying kernel-based methods (regularizing L_2-norm) and sparsity-based methods (regularizing l_1-norm), which have so far been investigated under different theoretical foundations.

*Machine Learning External Seminar, Gatsby Unit, May 7, 2014.