Learning without a Dictionary

Sara van de Geer[1,*]

[1]*Seminar for Statistics, Department of Mathematics, ETH Zürich, Switzerland*
*Email: geer@stat.math.ethz.ch

Consider the classical problem of learning a signal when observed with noise. One way to do this is to expand the signal in terms of basis functions and then try to learn the coefficients. The collection of basis functions is called a dictionary and the approach is sometimes called "synthesis" because the signal is synthesised from the coefficients. Another learning approach, called "analysis", is based on a linear operator that describes the signal's structure. As an example one may think of a signal that lives on a graph, and the linear operator describes the change when going from one node to the next in the graph. The sum of the absolute values of the changes is called the total variation of the signal over the graph. The analysis problem can be reformulated as a synthesis problem, for which the theory is well-developed. But instead of this, we take an easier direct route to derive the statistical properties of analysis. This will be done in the context of ℓ_1-norm regularisation. For the total variation example it leads to creative cutting and pasting.