Structure-aware Learning with Random Projections

Ata Kaban
School of Computer Science, University of Birmingham, UK
Email: ata.x.kaban@gmail.com

Random projection (RP) is a simple, computationally efficient linear dimensionality reduction technique with interesting theoretical properties for learning from data. For instance, it preserves dot products between vectors and their sign to an extent that depends on their cosine similarity. This talk will show how RP can also be used as an analytic tool to better understand why learning from high dimensional data is not always as difficult as it looks. This is achieved by the ability of RPs to adapt to and exploit certain hidden low complexity geometric structure in the problem. In particular, we give new, user-friendly PAC-bounds that are able to take advantage of such benign geometry to reduce dimensional-dependence of error-guarantees in settings where such dependence is known to be essential in general. To this end, we introduce an auxiliary function class that operates on RP-ed low dimensional inputs, and a new complexity term that we define as a distortion of the loss function of the problem under RP. The latter depends on both the hypothesis and the data. We work out analytic estimates of this new complexity term in several examples of learning problems, which turn out to yield expressions that recover various regularisation schemes in parametric models, and a notion of intrinsic dimension in the non-parametric case, quantified by the Gaussian width of the input support or a bounded subset of it. The bounds tighten with the presence of good structure, otherwise recover existing dimension-dependent bounds.