Distributionally robust sample average approximation

Andy Philpott
joint work with Eddie Anderson and Dominic Keehan

Workshop in honor of the 60th Birthday of Adrian Lewis
Seattle, 9–10 August, 2022

Sample average approximation (SAA) is a widely used technique in stochastic programming settings where we seek to minimize the expected cost of a decision in an environment affected by random variables with a known probability distribution. A sample of N points is randomly drawn from this distribution and an approximate problem is formulated and solved based on these points and the empirical distribution that assigns a probability of $1/N$ to each point. SAA can be viewed as a machine-learning technique where the sample is treated as training data drawn from an unknown probability distribution, and the SAA solution emerges from training (i.e. optimizing) the SAA model. In the last decade there have been many papers that study distributionally robust versions of the SAA problem. These seek to minimize the worst-case expected cost when evaluated using probability distributions lying in a given uncertainty set that represents plausible distributions that might have generated the sample. It has been observed that this approach often produces decisions that outperform the SAA solutions when simulated out of sample. My talk discusses some of our recent work on this phenomenon in cases where the uncertainty set is defined using Wasserstein metrics. Different choices of these give rise to different formulations of the distributionally robust optimization problem. These are illustrated with some simple quadratic optimization examples.