Bocconi
Università Commerciale Luigi Bocconi

Leveraging Learned Structure for Better Inference: Two Vignettes

Subhabrata Sen
Harvard University

Friday October 3rd, 2025
12:00 pm **Room 3-E4-SR03** Via Roentgen 1 Milano

Abstract

Accurate inference in high-dimensional problems relies crucially on exploiting the structure of the latent signal. When that structure can itself be learned from the data, a two-stage pipeline—"learn the structure, then perform inference"— can yield substantial statistical dividends. This talk substantiates that claim through two theoretically grounded examples.

In the first part of the talk, we will investigate Empirical Bayes inference in the high-dimensional linear model. We will introduce a new method which learns the latent prior using the NPMLE and then implements Bayesian inference using the estimated prior. To perform Bayesian inference in this setting, we will use the Naive Mean Field approximation to the marginal likelihood. We will establish the consistency of our method, and illustrate its empirical performance.

In the second part of the talk, we will study a specific formulation of the general pre-training/fine-tuning paradigm, which is commonly used to train modern language models. The pre-training phase uses unlabeled data to learn a latent representation, which is leveraged in the downstream fine-tuning phase.

We will rigorously establish the benefits of unsupervised pre-training in our setup.

The talk will be based on joint works with Sumit Mukherjee, Bodhisattva Sen, Aukosh Jagannath and Taj Jones-McCormick.

Department of Decision Sciences
Via Röntgen 1 – 20136 Milano

Tel. 02 5836.5632
Fax 02 5836.5630