

Computational Statistics

The course has two main components. One is a thorough introduction to what we call mixed models, a large family of models extensively used across applied sciences that includes as special cases random effect models, generalized linear mixed models, item response theory, factor models, probabilistic matrix factorization models, additive models (including splines etc), ideal point models, Carter-Lee type models used in demography, and many more. The second is a thorough introduction to variational inference. This is an approach to probabilistic machine learning that has had a huge impact over the past decade, especially for being able to carry out large scale inference where large models are trained on large datasets. The material is based on a book that Papaspiliopoulos is currently writing on variational inference for mixed models, jointly with the computational social scientist, Dr Max Goplerud.

Part I:

- + Motivating datasets: voter turnout, political ideology, movie recommendation
- + GLMs
- + Varying coefficients, interactions and aliasing and identifiability
- + From varying to random coefficient models
- + Random factors and low-rank interactions
- + From fixed effects, to random effects to random factors

Part II:

- + Intro to Variational Inference
- + Objective function and properties
- + Fixed and free form VI
- + Structured VI
- + Laplace approximation, gradient flows and ADVI
- + Coordinate Ascent VI
- + Variational EM
- + Uncertainty Quantification for VI