41005 - Advanced Computational Statistics

PhD in Statistics & Computer Science - Bocconi University

A.Y. 2026-2027

Instructor: Giacomo Zanella

Course description

The course introduces students to advanced methodologies in Computational Statistics and Machine Learning, and their theoretical analysis. It is divided in three parts:

1. **Variational Inference**: the first part reviews popular computational methodologies based on deterministic approximations, illustrating them through applications to some canonical high-dimensional models. Possible topics include:
    - Laplace approximation: derivation, error bounds in the large data regime, INLA and applications to latent Gaussian models
    - Variational Inferences (VI): mean-field, CAVI, application to Latent Dirichlet Allocation models
    - Stochastic optimization methods for VI: stochastic variational inference, fixed-form VI, variance-reduction techniques for gradient estimators
2. **Sampling and Optimization** the second part is more mathematical, and it deals with log-concave sampling theory, highlighting parallels with classical convex optimization results. Possible topics include:
    - Basics of convex optimization: convergence of gradient descent and stochastic gradient descent
    - Log-concave sampling: couplings and convergence of Langevin MCMC
    - Functional Inequalities and convergence of Markov processes
    - Non-convex distributions and multimodality
3. **Diffusion generative models:** the third part introduces the topic of generative diffusion models, describing recent advances in the area. Possible topics include:
    - Diffusion models: energy-based models, reversed SDEs, score-matching, score-based diffusions
    - Theoretical analysis and extensions: bounds for non-convex distributions, flow-matching, discrete diffusions