# Course Proposal: SDS7xx Advanced Statistical Computing

**A. Course Credits:** 9 (3-0-0-0)

**B. Objective**: This course is intended for PhD students and advanced undergraduate students in statistics, data science, machine learning, and related fields who want a deeper understanding of computational methods for statistical inference. Students will learn the theoretical foundations of modern optimization and Bayesian computation, along with the practical challenges of implementing these algorithms efficiently and reliably. The course will also highlight applications to large-scale statistical models and contemporary AI/ML methods, with an emphasis on understanding computational accuracy, scalability, and uncertainty. Regular implementation and coding practice will be done in R/Python/Julia. A student in the course can expect to get

1. exposure to a range of applications of optimization and sampling in statistical modeling and modern algorithms. The course will attempt to motivate topics via motivating examples from the real world;
2. a thorough theoretical understanding of convergence guarantees for optimization and sampling;
3. an insight into the practical challenges of implementing popular algorithms and how tuning parameters affect performance;
4. exposure to fundamental concepts in Monte Carlo sampling and an appreciation for when certain MCMC methods might work better than others;
5. an appreciation for when approximate sampling methods are required and how modern algorithms utilize them for scalable inference.

**C. Contents**

| **S.No.** | **Broad Title** | **Topics** | **No. of lectures** |
|---|---|---|---|
| | | Part 1: Optimization | |
| 1. | Numerical foundations | Fundamentals of numerical linear algebra<br>Computational complexity, matrix decompositions, automatic differentiation | 2 |
| 2. | Optimization fundamentals | Unconstrained optimization overview<br>Optimality conditions, convexity Quasi-convexity, strong convexity, L-smoothness, weak convexity | 2 |
| 3. | Gradient-based methods | *Motivating Examples: Logistic regression, ridge regression.*<br><br>Steepest descent, gradient descent, Example, accelerated gradient descent, Newton's method, Quasi-Newton methods | 4 |
| 4. | Beyond gradient methods | *Motivating Examples: Mixture models, LASSO, Nonnegative least squares, positive definite covariance matrix estimation, convex clustering.*<br><br>Majorize-minimize principle Expectation-maximization algorithm, Proximal gradient descent, coordinate descent. | 6 |
| 5. | Stochastic optimization methods | *Motivating Examples: Big data regression, and non-convex optimization.*<br><br>Stochastic gradient descent, accelerated stochastic gradient methods, ADAM | 2 |
| 6. | Constrained optimization | *Motivating Examples: Least absolute deviation regression, Low rank sparse matrix decomposition.*<br><br>Constrained optimization, KKT conditions, Duality, Barrier method, Quadratic penalty method, ADMM, | 4 |

| | | | |
|---|---|---|---|
| | | Distance-to-set penalty method | |
| Part 2: Bayesian Computation | | | |
| 7. | Sampling fundamentals | Inverse transform, rejection sampling, Bayesian models. | 2 |
| 8. | Metropolis-Hastings | *Examples: Bayesian regression, random effects models, Bayesian Bradley Terry Model.*<br><br>MH algorithm, random walk proposals, independence proposals, MALA proposal, Hamiltonian Monte Carlo | 5 |
| 9. | Gibbs samplers | *Motivating Example: Latent Dirichlet Allocation*<br><br>Gibbs samplers as variable-at-a-time MH, Metropolis-within-Gibbs, blocked and collapsed Gibbs. | 3 |
| 10. | Practical MCMC | MCMC convergence, rates of convergence, estimating Monte Carlo error, parallel chains | 2 |
| 11. | Approximate and scalable inference | *Motivating Example: Big data Bayesian regression, Reinforcement learning.*<br><br>Stochastic gradient Langevin dynamics, stochastic MCMC, INLA, Variational Bayes | 5 |
| 12. | Special topics and other applications (based on novel and new applications) | Next word predictions (Gumbel max trick), diffusion models and Langevin dynamics, non-convex optimization and neural networks. | 3 |
| **Total** | | | **40** |

**D. Pre-requisites:** MSO205 or equivalent, SDS210 or equivalent, SDS441 or equivalent, instructor's consent.

**E. Short summary:** This course covers advanced computational methods for statistical inference, with a focus on optimization and Bayesian computation. Topics include numerical linear algebra, gradient and stochastic optimization, other common unconstrained and constrained optimization algorithms, MCMC, Monte Carlo methods, variational inference, and scalable computational methods. Emphasis is placed on both the theoretical properties of algorithms and the practical challenges of computation, including convergence, numerical stability, computational complexity, and uncertainty quantification. Throughout the course, contemporary and foundational statistical applications will motivate the need for various algorithmic developments.

**Recommended Books and reference material**


a) J. Nocedal and S. J. Wright, Numerical Optimization, Springer, 2nd ed., 2006.
b) S. Boyd and L. Vandenberghe, Convex Optimization, Cambridge University Press, 2004.
c) S. Sra, S. Nowozin, and S. J. Wright (eds.), Optimization for Machine Learning, MIT Press, 2011.
d) Brooks, Steve, Andrew Gelman, Galin Jones, and Xiao-Li Meng, eds. *Handbook of Markov chain Monte Carlo*. CRC press, 2011.
e) Robert, Christian, and George Casella. *Monte Carlo statistical methods*. Springer Science & Business Media, 2013.


**Other Departments/IDPs which may be interested in the course:** Computer Science, Electrical Engineering, Intelligent Systems.


**Proposing Instructor:** Dootika Vats (dootika@iitk.ac.in)
**Proposing Department:** Department of Statistics and Data Science