Introduction to Algorithms for Data Science and Physics Spring/2026
(数据科学及物理问题中的算法导论)

@IMP/Fudan

Syllabus

Bao-Jun Cai, bjcai@fudan.edu.cn

—Course Description—

This course offers an introduction to the fundamental concepts of algorithms and several widely used methods in machine learning, state estimation/inference, and modern data science. A key objective is to help students understand the underlying principles and theoretical foundations of machine learning algorithms. Topics covered include basic probability and statistical techniques, simulation methods, optimization strategies, and essential ideas from high-dimensional problems and reconstruction algorithms. To keep pace with the rapid evolution of big data and modern analytics, the course also explores a selection of contemporary algorithms developed for large-scale data processing and high-dimensional statistical analysis.

—Target Audience—

This course is intended for advanced undergraduate and graduate students in physics, computer science, data science, or related fields who are interested in the foundations and practical applications of machine learning and modern data analysis. It is also suitable for researchers and professionals seeking to strengthen their understanding of algorithmic methods in state estimation, optimization, computer vision and physics. A basic background in calculus, linear algebra, and probability is recommended, though key concepts will be reviewed as needed throughout the course. You are all very welcome to attend the course!

—Exercises and Problems—

Each class will last approximately 45x3=135 minutes. There will be two types of exercises: analytical problems (such as estimations or derivations) and programming assignments. The analytical problems are relatively straightforward, while the programming tasks require more thoughtful design and implementation. Each week, a few programming/theoretical exercises will be assigned to reinforce the key concepts covered in class. These exercises are designed to closely reflect practical techniques in areas such as physics, state estimation, geometric reasoning under uncertainty, and optimization.

—Course Grading Policy—

(a) *Homework*: 30%, except Module F.

(b) *Quiz*: 30%, except Module F.

(c) *Final Exam*: 40%, except Module F.

—Detailed Schedules (Preliminary)—

Module–A *Basis of Calculus, Probability and Statistics* (微积分、概率与统计基础)

Lecture 1 — *Course Introduction and Some Examples*

(a) method of guess and estimation, period of a simple pendulum system
(b) master theorem, divide-and-conquer, Fibonacci series
(c) fast algorithm for matrix multiplication
(d) solution of the algebraic equation $x^n(t) = \Omega + tx(t)/\Lambda$, $x(t) \in \mathrm{R}^+$, $n \geq 5$, $n \in \mathrm{N}^+$

LECTURE 2 — *Calculus and Concept of Sampling*

(a) Taylor's expansion of a function
(b) five-point algorithm
(c) gradient and Hessian of a matrix
(d) Monte Carlo method for integration
(e) estimating π by hit-or-miss and by Markov chain sampling

LECTURE 3 — *Elementary Introduction to Statistics and Probability*

(a) mean and variance of a distribution
(b) Bayes' theorem
(c) moment-generating function
(d) random variable generation, Box–Muller method

Module–B *Optimization, Learning from Data and Bias–variance Trade–off* (基本优化算法、方差-偏差分解)

LECTURE 4 — *First-order Optimization Algorithms*

(a) gradient descent search, exact-line search
(b) Jensen's inequality for convex function
(c) regularization, singular behavior of Hessian
(d) concept of momentum
(e) AdaGrad, RMSProp, Adam, hypergradient

LECTURE 5 — *Trade–off between Bias and Variance*

(a) linear fitting, loss function
(b) nonlinear curve fitting, learning model
(c) no–free–lunch theorem, trade–off between bias and variance
(d) stochastic gradient descent
(e) geometrical error learning model

LECTURE 6 — *Annealing and Metropolis Algorithm*

(a) thermal motion, Boltzmann's distribution
(b) simulated (thermal) annealing
(c) rejection sampling, importance sampling
(d) fast annealing on finding the minimum of a multi-dimensional function
(e) 2D Ising model: simulations

LECTURE 7 — *Convergent Analysis and Large-scale Sparse Matrix*

(a) some facts related to the Hessian
(b) search algorithm directly using Hessian
(c) Newton–Raphson method, order of convergence
(d) Gauss–Newton and Levenberg–Marquardt schemes, applicable conditions
(e) conjugate gradient search

Module–C *State Estimation/Inference and Data-driven Algorithms* (状态估计/推断与数据驱动的算法)

LECTURE 8 — *Kalman Filter as a Data-driven Optimization Method*

(a) Kalman filter (without analytical derivations)
(b) nonlinear and non-Gaussian extensions
(c) bias estimation for the learning model $f_w(x) = e^{-wx}$

LECTURE 9 — *Bayesian Curve Fitting as a State Estimation Problem*

(a) high-dimensional Gaussian, decomposition of Gaussian
(b) Bayesian consideration on coin drawing experiment
(c) linear curve fitting using Bayesian calculation

LECTURE 10 — *3D Reconstruction from 2D Images for Computer Vision Problems*

(a) depth determination from stereo pairs
(b) picture of 3d reconstruction, motion determination, SLAM

(c) least squares, weighted least squares
(d) bundle adjustment (conceptual introduction)

Module–D *High-dimensional Problems, Randomized Algorithms and Fast Computing* (高维问题导论、随机算法与快速计算)

LECTURE 11 — *Clustering, Robustness and Sparsity*

(a) concept of classification, effects of dimensions
(b) k-means, k-center, k-median
(c) robustness and sparsity, LASSO
(d) Huber loss with ℓ-1 constraint, the optimization algorithm

LECTURE 12 — *High Dimensions, Singular-value Decomposition and Best-fit Subspace*

(a) law of large numbers, random data in high dimensions
(b) properties of a high-dimensional ball
(c) singular value decomposition, a greedy algorithm
(d) basic principle component analysis
(e) analysis on the MNIST dataset and whitening of noise data

LECTURE 13 — *Randomized Matrix Multiplication, Integer Partition and Phase Transition*

(a) massive data problem
(b) distinct elements of a set
(c) CUR decomposition
(d) saddle-point approximation
(e) Laplace's method for integration

Module–E *Fast Fourier Transform and Solvers for Partial Differential Equations* (快速 Fourier 变换与偏微分方程的离散化计算)

LECTURE 14 — *Fast Fourier Transform for Data Processing*

(a) product of polynomials
(b) fast Fourier transform for signal processing
(c) convolution of signals

LECTURE 15 — *Differencing Schemes for Partial Differential Equations*

(a) Euler search, Runge–Kutta algorithm
(b) convection equation, differencing schemes and stability conditions
(c) up-wind method
(d) random walk for Laplaces' equation

Module–F *A Primer for Quantum Algorithms* (量子算法基础)

LECTURE 16 — *Review of Basis of Quantum Mechanics*

(a) wave nature of particles and complex amplitude
(b) uncertainty relation between momentum and position
(c) operator and wave function
(d) amplitude as sum over histories, example: harmonic oscillator
(e) path integral simulations

LECTURE 17 — *Algorithmic Interference*

(a) quantum state, qubit, coherence and decoherence
(b) collapse of wave function and entanglement
(c) Shor algorithm and Deutsch's problem
(d) Gover's searching algorithm, square-root acceleration

SPECIAL LECTURE ***Renormalization Group, Effective Field Theory and Deep Neural Network*
If time permits, I will introduce some modern idea on understanding deep neural network, via the methods of renormalization group and effective field theory.

—Lecture Notes and Selected References—

There will be no designated textbook for the course, but lecture notes will be provided progressively as the course develops. However, several reference books that offer broad coverage of relevant topics may be helpful for those seeking a deeper understanding (books with "♠" are recommended). You do not need to read everything.

(1) I. Jolliffe, *Principal Component Analysis*, Springer, 1986.
(2) R. Neal, *Bayesian Learning for Neural Networks*, Springer, 1996.
(3) G. Cowan, *Statistical Data Analysis*, Oxford, 1998.
(4) E. Lehmann and G. Casella, *Theory of Point Estimation*, Springer, 1998.
(5) M. Newman and G. Barkema, *Monte Carlo Methods in Statistical Physics*, Clarendon, 1999.
(6) B. Schölkopf and A. Smola, *Learning with Kernels*, Cambridge, 2002.
(7) ♠D. MacKay, *Information Theory, Inference, and Learning Algorithms*, Cambridge, 2003.
(8) S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge, 2004.
(9) R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge, 2004.
(10) ♠J. Kleinberg and E. Tardos, *Algorithm Design*, Pearson, 2005.
(11) C. Rasmussen and C. Williams, *Gaussian Processes for Machine Learning*, MIT, 2005.
(12) S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*, MIT, 2005.
(13) ♠C. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006.
(14) G. Casella and R. Berger, *Statistical Inference*, Thompson, 2006.
(15) T. Cover and J. Thomas, *Elements of Information Theory*, Wiley, 2006.
(16) ♠W. Krauth, *Statistical Physics: Algorithms and Computations*, Oxford, 2006.
(17) ♠J. Nocedal and S. Wright, *Numerical Optimization*, Springer, 2006.
(18) D. Sivia and J. Skilling, *Data Analysis: A Bayesian Tutorial*, Oxford, 2006.
(19) ♠♠W. Press, S. Teukolsky, W. Vetterling, and B. Flannery, *Numerical Recipes*, Cambridge, 2007.
(20) T. Hastie, R. Tibshirani, and J. Friedmann, *The Elements of Statistical Learning*, Springer, 2008.
(21) D. Koller and N. Friedman, *Probabilistic Graphical Models*, MIT, 2009.
(22) L. Wasserman, *All of Statistics*, Springer, 2010.
(23) C. Moore and S. Mertens, *The Nature of Computations*, Springer, 2011.
(24) D. Barber, *Bayesian Reasoning and Machine Learning*, Cambridge, 2012.
(25) R. Horn and C. Johnson, *Matrix Analysis*, Cambridge, 2012.
(26) Y. Mostafa, M. Ismail, and H.T. Lin, *Learning from Data*, AMLBook, 2012.
(27) ♠S. Prince, *Computer Vision: Models, Learning, and Inference*, Cambridge, 2012.
(28) S. Aaronson, *Quantum Computing Since Democritus*, Cambridge, 2013.
(29) A. Gelman, J. Carlin, H. Stern, D. Dunson, A. Vehtari, and D. Rubin, *Bayesian Data Analysis*, CRC, 2013.
(30) ♠♠G. Golub and Van C. Loan, *Matrix Computations*, John Hopkins, 2013.
(31) I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT, 2016.
(32) R. Vidal, Y. Ma, and S. Sastry, *Generalized Principal Component Analysis*, Springer, 2016.
(33) M. Mitzenmacher and E. Upfal, *Probability and Computing*, Cambridge, 2017.
(34) M. Wilde, *The Quantum Information Theory*, Cambridge, 2017.
(35) R. Vershynin, *High-dimensional Probability*, Cambridge, 2018.
(36) M. Wainwright, *High-dimensional Statistics*, Cambridge, 2018.
(37) G. Strange, *Linear Algebra and Learning from Data*, Cambridge, 2019.
(38) ♠A. Blum, J. Hopcroft, and R. Kannan, *Foundations of Data Science*, Cambridge, 2020.
(39) M. Deisenroth,*Mathematics for Machine Learning*, Cambridge, 2020.
(40) T. Hastie, R. Tibshirani, and M. Wainwright, *Statistical Learning with Sparsity*, Chapman & Hall, 2020.
(41) S. Skiena, *The Algorithm Design Manual*, Springer, 2020.
(42) L. Böttcher and H. Herrmann, *Computational Statistical Physics*, Cambridge, 2021.
(43) S. Brunton and J. Kutz, *Data-Driven Science and Engineering*, Cambridge, 2022.
(44) ♠♠T. Cormen, C. Leiserson, R. Rivest, and C. Stein, *Introduction to Algorithms*, MIT, 2022.
(45) J. Wright and Y. Ma, *High-dimensional Data Analysis with Low-dimensional Models*, Cambridge, 2022.
(46) ♠S. Prince, *Understanding Deep Learning*, MIT, 2023.
(47) A. Torralba, P. Isola, and W. Freeman, *Foundations of Computer Vision*, MIT, 2024.
(48) S. Dorogovtsev and J. Mendes, *The Nature of Complex Network*, Oxford, 2025.