

贝叶斯方法

推断框架、计算方法与现代贝叶斯学习

MarkZZZ WeChat: MarkZZZ20XX

课程简介

本课程系统介绍贝叶斯统计推断的理论基础与现代计算方法，从贝叶斯定理出发，严格建立先验-似然-后验框架与决策理论基础，深入剖析共轭先验族、无信息先验（Jeffreys 先验）与层次贝叶斯模型，随后进入贝叶斯计算的核心算法：Metropolis-Hastings 及其收敛性（总变差距离、混合时间界）、Gibbs 采样与数据增广、哈密顿蒙特卡洛（HMC）及 NUTS 的辛积分推导与理论接受率，继而系统讲授变分推断（ELBO 完整推导、坐标上升 VI）及高级变分方法（重参数技巧、VAE、正则化流）。课程后半部分覆盖贝叶斯非参数方法（Dirichlet 过程、高斯过程回归）、贝叶斯优化（EI/UCB 采集函数与 GP-UCB 遗憾界）以及贝叶斯深度学习（Laplace 近似、MC Dropout、不确定性量化与校准）。所有定理给出完整证明，课程标准对齐 MIT/Stanford/CMU 博士级贝叶斯方法课程。

适合人群

- 统计学、计算机科学、机器学习、运筹学等方向的博士生及高年级硕士生
- 希望掌握现代贝叶斯计算（MCMC、VI、HMC）严格理论的研究人员
- 从事概率机器学习、不确定性量化或贝叶斯优化研究的学者
- 已有频率主义统计基础，寻求系统贝叶斯视角的理论工作者

前置知识

- **概率论**：条件概率、期望、常见分布族（正态、指数族）、收敛概念
- **数理统计**：参数估计（MLE、矩估计）、充分统计量、Fisher 信息
- **线性代数**：矩阵运算、正定矩阵、特征分解、Cholesky 分解
- **微积分与优化**：多元微积分、Lagrange 乘子法、KL 散度基础
- **机器学习基础（有益）**：概率图模型、神经网络基础

1 课程大纲

讲次 主题

内容概要

1 贝叶斯推断框架

贝叶斯定理、先验/后验/似然/边缘似然、后验预测分布、决策理论（Bayes 风险、最小化后验期望损失）、MAP 估计与 MLE 的关系、可信区间（Credible Interval）与置信区间的对比

讲次	主题	内容概要
2	共轭先验族	指数族的自然参数与充分统计量、共轭先验的构造原理、Beta-Binomial、Dirichlet-Multinomial、Normal-Normal、Normal-InvGamma、Poisson-Gamma 共轭对完整推导、先验选择的信息准则
3	无信息先验	Laplace 均匀先验及其局限、Jeffreys 先验（Fisher 信息推导与不变性定理完整证明）、参考先验（Reference Prior）理论、位置族和尺度族先验、Bernardo-Berger 框架
4	层次贝叶斯	层次模型的动机与构建、超先验与可交换性（de Finetti 定理）、经验贝叶斯（EB）估计、James-Stein 收缩估计器及其无偏风险估计（SURE）、条件独立结构与混合模型
5	MCMC 基础	Metropolis-Hastings 算法详解、细致平衡条件与平稳分布、Markov 链收敛性理论（不可约性、非周期性）、总变差距离、谱间隙与混合时间界、Metropolis-within-Gibbs
6	Gibbs 采样与数据增广	Gibbs 采样的条件分布推导、充分条件与几何收敛速度、数据增广（Data Augmentation）策略、隐变量模型（混合高斯、LDA）、EM 算法与 Gibbs 采样的关系
7	哈密顿蒙特卡洛	哈密顿动力学与辛积分（Leapfrog 方案）、体积保持性与时间可逆性证明、理论接受率公式推导、No-U-Turn Sampler (NUTS)、预条件 HMC (Riemannian HMC 思想)
8	变分推断	变分推断动机与 ELBO 完整推导（含 KL 散度分解）、Mean-Field 近似（全因子近似）、坐标上升 VI (CAVI) 算法与收敛性、Evidence Lower Bound 与边缘似然的关系、Variational EM
9	高级变分方法	重参数技巧（Reparameterization Trick）与随机梯度 VI、变分自编码器（VAE）理论推导、正则化流（Normalizing Flows）理论（Change-of-Variables 公式、RealNVP/Planar Flows）、IWAE 重要性加权自编码器
10	贝叶斯非参数	Dirichlet 过程（DP）的构造（Ferguson 1973）、中国餐厅过程（CRP）表示、折棍过程（Stick-Breaking, Sethuraman 定理完整证明）、高斯过程（GP）回归：后验推断与预测分布推导

讲次	主题	内容概要
11	贝叶斯优化	高斯过程替代模型、采集函数 (Expected Improvement、UCB、Thompson Sampling) 的推导与性质、Gaussian Process UCB (GP-UCB) 遗憾界定理及完整证明、批量贝叶斯优化
12	贝叶斯深度学习	神经网络权重的贝叶斯先验、Laplace 近似 (Hessian 计算与近似方法)、MC Dropout 作为近似贝叶斯推断、深度集成方法、预测不确定性量化、校准 (Reliability Diagram、ECE) 与温度缩放

2 参考书目

1. Andrew Gelman, John B. Carlin, Hal S. Stern, David B. Dunson, Aki Vehtari, and Donald B. Rubin, *Bayesian Data Analysis*, 3rd ed., CRC Press, 2013. (贝叶斯数据分析权威教材)
2. Christopher M. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006. (第 2-4、9-10 章: 贝叶斯方法与图模型)
3. Kevin P. Murphy, *Machine Learning: A Probabilistic Perspective*, MIT Press, 2012. (第 3-5、11、24 章: 贝叶斯推断与非参数方法)
4. Kevin P. Murphy, *Probabilistic Machine Learning: An Introduction*, MIT Press, 2022. (现代概率机器学习的全面处理)
5. Kevin P. Murphy, *Probabilistic Machine Learning: Advanced Topics*, MIT Press, 2023. (变分推断、深度生成模型、贝叶斯优化前沿)
6. David J. C. MacKay, *Information Theory, Inference, and Learning Algorithms*, Cambridge University Press, 2003. (免费在线版; 第 20-28 章: 贝叶斯推断)
7. Radford M. Neal, "MCMC Using Hamiltonian Dynamics," in *Handbook of Markov Chain Monte Carlo*, CRC Press, 2011. (HMC 理论的权威综述)
8. Martin J. Wainwright and Michael I. Jordan, "Graphical Models, Exponential Families, and Variational Inference," *Foundations and Trends in Machine Learning*, 1(1-2):1-305, 2008. (变分推断的奠基性综述)
9. Carl Edward Rasmussen and Christopher K. I. Williams, *Gaussian Processes for Machine Learning*, MIT Press, 2006. (GP 方法的标准参考, 免费在线版)
10. Nils Lid Hjort, Chris Holmes, Peter Müller, and Stephen G. Walker (eds.), *Bayesian Nonparametrics*, Cambridge University Press, 2010. (贝叶斯非参数的系统处理)