

生成模型

变分推断、流模型、GAN、扩散模型与统一视角 · 博士进阶课程

MarkZZZ WeChat: MarkZZZ20XX

课程简介

生成模型 (Generative Models) 是现代深度学习的核心研究领域之一，旨在学习数据的底层概率分布并从中生成新样本。本课程从**密度估计**的统计学基础出发，系统深入地讲解四类主流生成框架：**变分自编码器 (VAE)**、**标准化流 (Normalizing Flows)**、**生成对抗网络 (GAN)** 与 **扩散模型 (Diffusion Models)**。全程以**严格数学推导**为主线：ELBO 的完整变分推导、重参数梯度的无偏性证明、Wasserstein 距离的 Kantorovich 对偶定理完整证明、DDPM 正向/逆向过程推导、随机微分方程 (SDE) 框架及逆时 SDE 推导，均给出博士水平的严格处理。课程还涵盖**高效采样** (DDIM、DPM-Solver、一致性模型)、**条件生成** (Classifier/Classifier-Free Guidance、ControlNet)、**Flow Matching** 与**最优传输**，以及**能量模型**与**Score Matching** 的统一视角。全部 12 讲以博士研究生水平呈现，使学生掌握生成模型的数学本质并具备开展前沿研究的理论工具。

适合人群

- 已具备深度学习基础，希望深入理解生成模型理论的博士生与研究者
- 从事图像合成、语音生成、分子设计、科学计算等生成任务的工程师
- 对概率推断、随机微分方程、最优传输等数学方向感兴趣的研究者
- 希望追踪扩散模型、Flow Matching 等前沿方向的 AI 从业者

前置知识

- 概率论基础：贝叶斯定理、条件概率、KL 散度、期望/方差
- 深度学习基础：神经网络、反向传播、常见优化器 (Adam/SGD)
- 变分推断基础：Evidence Lower Bound (ELBO) 概念有帮助
- 线性代数：矩阵运算、行列式、雅可比矩阵
- 随机微分方程基础 (伊藤积分) 对第 7-9 讲有帮助但非必须

1 课程内容

讲次	主题	内容概要
1	生成建模框架	生成模型的目标: 密度估计与采样; 显式/隐式密度模型的分类; KL 散度与逆 KL 散度 (模式覆盖 vs 模式搜索); 最大似然估计的统计基础; 评估指标全景: FID、IS、LPIPS、Precision/Recall; 四类主流框架概览与比较 (VAE/Flow/GAN/Diffusion)
2	变分自编码器 (VAE)	隐变量模型的概率图表示; ELBO 的完整变分推导 (两种等价形式); 重参数技巧 (Reparameterization Trick): 梯度估计的无偏性证明; 高斯 VAE 的 ELBO 闭合形式; 后验坍塌 (Posterior Collapse) 的诊断与修复
3	变分推断进阶	层次 VAE (HVAE) 与 NVAE: 自上而下推断过程; β -VAE 的信息论分析与解缠表示 (Disentanglement); VQ-VAE: 向量量化、承诺损失、直通估计器 (Straight-Through Estimator); VQ-VAE-2 与层次量化; 离散潜变量模型的 ELBO 分析
4	标准化流	Change-of-Variables 公式的完整推导与雅可比行列式计算; RealNVP: 仿射耦合层的设计与可逆性证明; Glow: 1×1 可逆卷积的 LU 分解技巧; FFJORD: 连续时间流的瞬时变量替换公式; 流模型的表达能力与计算复杂度分析
5	GAN 基础	生成对抗网络的 min-max 博弈框架; 原始 GAN 的最优判别器推导与 JS 散度等价性定理 (完整证明); 纳什均衡分析与全局最优解的存在性; 训练不稳定性的理论分析 (梯度消失/饱和); 模式崩塌 (Mode Collapse) 的成因与实验现象
6	GAN 训练理论	Wasserstein-1 距离 (Earth Mover's Distance) 的定义与直觉; Kantorovich-Rubinstein 对偶定理的完整证明; WGAN 训练算法与 Lipschitz 约束 (权重裁剪 vs 梯度惩罚); 梯度惩罚 (WGAN-GP) 的推导与 DRAGAN; 谱归一化 (SN-GAN) 与 Lipschitz 条件的关系
7	扩散模型基础	去噪扩散概率模型 (DDPM) 的动机; 正向过程: 马尔可夫加噪链与 $q(x_t x_0)$ 的闭合形式推导; 逆向过程: $p_\theta(x_{t-1} x_t)$ 的参数化与 ELBO 推导; 噪声预测目标与 x_0 预测目标的等价性证明; DDPM 训练算法与采样算法

讲次	主题	内容概要
8	扩散模型理论	SDE 框架 (Song et al. 2021): VP-SDE、VE-SDE、Sub-VP-SDE; 正向 SDE 的 Fokker-Planck 方程; 逆时 SDE (Anderson 1982): 完整推导与评分函数 $\nabla_x \log p_t(x)$; Tweedie 公式及其在扩散模型中的应用; 概率流 ODE 的推导与确定性采样
9	高效采样	DDIM (Denoising Diffusion Implicit Models): 非马尔可夫过程的构造与隐式采样推导; DPM-Solver: 扩散 ODE 的高阶数值求解; 一致性模型 (Consistency Models): 一致性函数定义、一致性蒸馏与训练; 采样加速的统一分析框架; NFE (Function Evaluations) 的权衡
10	条件生成	条件生成的一般框架; Classifier Guidance: 条件评分函数的分解与引导强度; Classifier-Free Guidance (CFG): 联合训练框架与插值系数; ControlNet: 编码器克隆架构与 Zero Convolution; 文本条件生成 (CLIP 嵌入与交叉注意力); 引导强度与多样性-保真度权衡
11	Flow Matching 与最优传输	连续归一化流 (CNF) 与向量场学习; Flow Matching 目标函数的推导; OT-CFM (最优传输条件流匹配): 正则化最优传输的 Brenier 定理; Stochastic Interpolants 框架; 与扩散模型的对比: 轨迹直线性与采样效率; Minibatch OT 的实现技巧
12	前沿与统一视角	能量模型 (EBM) 与 Langevin 动力学采样; Score Matching (DSM/SSM) 的推导与与扩散模型的联系; 四类框架的统一视角: 评分函数、ELBO、最优传输、信息几何; 研究前沿: 视频生成、3D 生成、科学生成 (蛋白质/分子); 开放问题与未来方向

2 参考资料

- Kingma, D.P. & Welling, M. "Auto-Encoding Variational Bayes." *ICLR*, 2014.
- Rezende, D.J., Mohamed, S. & Wierstra, D. "Stochastic Backpropagation and Approximate Inference in Deep Generative Models." *ICML*, 2014.
- Goodfellow, I. et al. "Generative Adversarial Nets." *NeurIPS*, 2014.
- Arjovsky, M., Chintala, S. & Bottou, L. "Wasserstein GAN." *ICML*, 2017.
- Gulrajani, I. et al. "Improved Training of Wasserstein GANs." *NeurIPS*, 2017.
- Ho, J., Jain, A. & Abbeel, P. "Denoising Diffusion Probabilistic Models." *NeurIPS*, 2020.

7. Song, Y. et al. “Score-Based Generative Modeling through Stochastic Differential Equations.” *ICLR*, 2021.
8. Song, J. et al. “Denoising Diffusion Implicit Models.” *ICLR*, 2021.
9. Dinh, L., Sohl-Dickstein, J. & Bengio, S. “Density Estimation Using Real-Valued Non-Volume Preserving Transformations.” *ICLR*, 2017.
10. Kingma, D.P. et al. “Glow: Generative Flow with Invertible 1×1 Convolutions.” *NeurIPS*, 2018.
11. Lipman, Y. et al. “Flow Matching for Generative Modeling.” *ICLR*, 2023.
12. Albergo, M.S. & Vanden-Eijnden, E. “Building Normalizing Flows with Stochastic Interpolants.” *ICLR*, 2023.
13. Rombach, R. et al. “High-Resolution Image Synthesis with Latent Diffusion Models.” *CVPR*, 2022.
14. Song, Y. et al. “Consistency Models.” *ICML*, 2023.
15. Ho, J. & Salimans, T. “Classifier-Free Diffusion Guidance.” *NeurIPS Workshop*, 2021.