Lecture 17: Wrap-up

Modern Methods in Applied Statistics

STAT 34800 (Spring 2023)

THE UNIVERSITY OF CHICAGO

Overview

- **Part I**: Foundations + building blocks
  - Decision theory
    - *loss / likelihood / prior*
    - *risk minimization*
  - Probabilistic reasoning
    - *prior, likelihood, posterior, kernel, marginal, predictive distributions*
    - *interpretations of probability*
  - Exponential families
    - *conjugacy*
  - Information theory
    - *entropy (i.e., "uncertainty")*
    - *KL, mutual info*
  - Graphical models
    - *semantics (conditional indep.)*
    - *directed vs undirected*
- **Part II**: Inference + algorithms
  - Exact inference:
    - *conjugate updating*
    - *variable elimination*
  - Point estimation (MLE / MAP):
    - *empirical Bayes*
    - *EM*
  - Approximate inference:
    - *MCMC / Gibbs (samples)*
    - *CAVI (approximate density)*

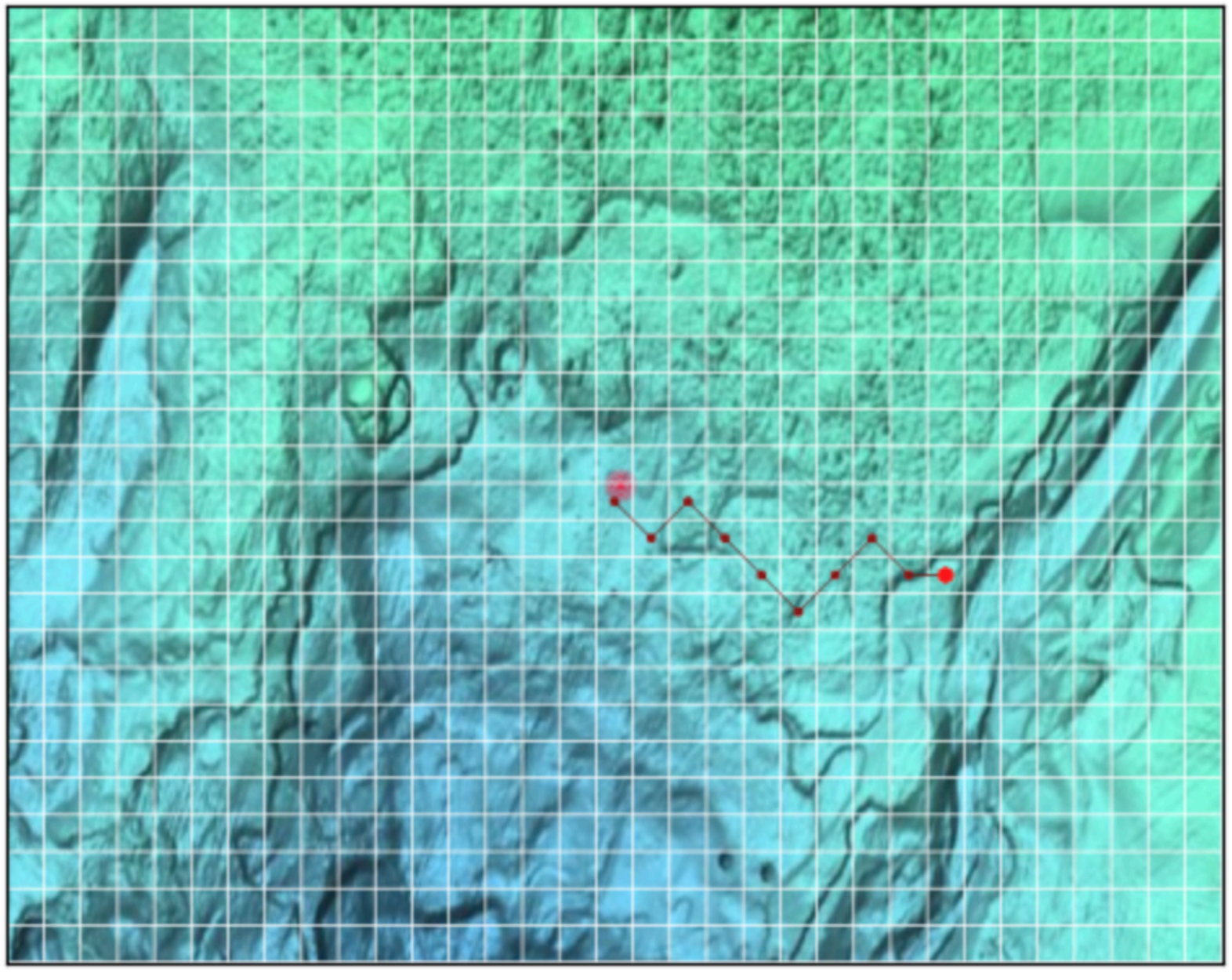
Inference + algorithms

Algorithm	Target	Property	Model type
Conjugate updating	Posterior	Exact (analytically)	Conjugate
Empirical Bayes (Type II MLE)	MLE / MAP	Exact (analytically)	Conjugate
Variable elimination	Posterior	Exact (algorithmically)	Discrete graphical model
Belief prop. / sum-product	Posterior	Exact (algorithmically)	Discrete tree
Forwards-backwards	Posterior	Exact (algorithmically)	HMM
Expectation-maximization (EM)	MLE / MAP	Exact (computationally)	Conditionally conjugate
Gibbs sampling	Posterior	Approximate (Monte Carlo)	Conditionally conjugate
Variational inference (CAVI)	Posterior	Approximate (density)	Conditionally conjugate

Motivations for Bayes

- Decision theoretic motivation
  - *Necessity of prior odds*
  - *Lady drinking tea vs. drunk friend*
- Frequentist motivations
  - *"Optimality" of Bayes for risk minimization*
  - *Admissibility of shrinkage estimators*
- Supervised learning motivation
  - *Regularization to avoid overfitting*
  - *Regularizers ≈ inductive biases ≈ priors*
- Orthodox motivations
  - *De Finetti (exchangeability → prior-likelihood)*
  - *Dutch book argument (coherent beliefs → probabilities)*
- Practical motivation
  - *Unified/coherent language for latent variables and uncertainty*
  - *Modularity: few algorithms for many models*

Thank you for your service



The final voyage of the USS Scorpion

Part III…∞

- **Gaussian models**
  - HMMs → LDS / Kalman filters
  - VE → Gaussian VE
  - MF → factor analysis, PCA

- **More complex canonical models**
  - Hierarchical & nested models
  - Tensor factorization & multi-view models

- **Stochastic process models**
  - Point process and continuous-time models *(stochastic processes as likelihoods)*
  - Infinite mixtures & Bayesian nonparametrics *(stochastic processes as priors)*

- **Advanced / scalable inference**
  - Augmentation schemes
  - Gradient-based stochastic inference
    - Langevin diffusion (SG-MCMC)
    - Gradient descent on the ELBO (VI)
    - Variational autoencoders (VAEs)

- **Model-checking**
  - Marginal likelihood
  - Posterior predictive checks

- **Causality**
  - Causal graphical models
  - Synthetic controls