CS281 Midterm

Discrete Distributions

- Understand mean parameter form and function of discrete distributions including Bernoulli, Binomial, Categorical, Multinomial, and Poisson.
- Construct and manipulate conjugate priors for discrete distributions focusing on Beta and Dirichlet, and understand role of hyperparameters.
- Compare and contrast parameter estimation methods of MLE, MAP, and full posterior estimation, as well as corresponding predictive distributions and understand the computational and modeling trade-offs inherent in each.

Continuous Distributions

- Understand the mean parameter form and function of multivariate Gaussians and the underlying geometric interpretation.
- Interpret the use of multivariate Gaussian as a prior distribution.
- Manipulate correlated Gaussian random variables through the use of marginalization, conditioning, and application of linear Gaussian updates.
- (Not Tested) We will not ask you to rederive properties of Gaussian algebra on the exame.

Linear Regression and Classification

- Full mastery is expected of Bayesian linear regression and Bayesian Naive Bayes. Understand each component of the models, the conditioning properties, role of common priors, difference in MLE, MAP, full posterior, use of the term "linear".
- Compare and contrast different variants of parameterization of Naive Bayes.
- Compare and contrast Bayesian and non-Bayesian linear regression. Derive variants of Bayesian linear regression utilizing Gaussians.

Exponential Families

- Recall and use of the exponential family mathematical form and the role of each component.
- Ability to transform a distribution into this form in order to label and interpret each component.
- Application of the major properties of exponential families including usage of the log-partition function, computation of general form of maximum likelihood, and derivation of conjugate prior form.

General Linear Models

- Grasp the high-level the definition of a general linear model, relationship to exponential families, and conditioning properties.
- Understand the commonly used general linear models including linear, logistic, and softmax regression and their derivations.
- Be able to construct algorithms for computing the MLE of an arbitrary GLM.

Neural Networks

- Comfort with interpreting basis functions, both static basis and adaptive basis functions, within the general (non)-linear model framework
- Intuitive grasp of the benefits / downsides of non-linear modeling with adaptive basis functions.
- High-level understanding of the multivariate chain rule, and the computational process of autograd. For example, given a homework coding problem, you should be able to describe the intermediate values necessary for pytorch to store for one step of stochastic gradient descent.
- (Not tested) We will not ask about convolutional networks, word embeddings, or specific neural networks on the exam.

Directed Graphical Models

- Construction of directed graphical models directly from factored joint distributions and vice-versa, including models with plates.
- Ability to answer questions about conditional independence directly from directed graphical models.
- Familiarity with common directed graphical models including naive bayes, hidden markov models, and likelihood distributions over many variables.
- Understand special case of linear Gaussian DGMs and how it differs from discrete case.

Undirected Graphical Models

- Construction of undirected graphical model from factored joint distributions and from directed graphical models.
- For discrete UGMs, the ability to move seamlessly from UGM to exponential family form.
- Ability to answer questions about conditional independence directly from undirected graphical models.
- (Not Tested) Linear Gaussian UGMs.

Time-Series Models

- Familiarity with the family of Markov-style models, ability to construct a joint time-series distribution from the specification of it graphical model and parameterization.
- High-level understanding of a UGM (and to a lesser extent DGMs) as an "abstraction", how a graph can represent many different distributions through choice of parameterization. Why this might be useful.
- (Not Tested) Specific names of specialized time-series models.

Exact Inference

- High-level understanding of computational difficulty of inference and computation of log-partition in general, and the difficulty of brute-force computation (as in the homework).
- Step-by-step comprehension of the forward-backward and sum-product algorithm. Including flow of information and computational and memory efficiency.
- Ability to infer worst-case complexity from the structure of the graph using tree-width.
- (Not Tested) We will not ask about NP-hardness proof of sum-product using 3-SAT.