



Variational Bayesian Methods (in Neuroscience)

Tyler Benster & Aaron Andalman

Deisseroth (Tyler & Aaron) and Druckmann (Tyler) Labs



Computational Neuroscience Journal Club
Stanford University

July 10, 2019



2019-07-10

└ Roadmap

Introduction: Why care about the distribution of data?

Problem: Analyzing high dimensional data is hard

Solution: Variational Methods

Example: variational autoencoder

Discussion: Neuroscience applications

Introduction: Why care about the distribution of data?

Problem: Analyzing high dimensional data is hard

Solution: Variational Methods

Example: variational autoencoder

Discussion: Neuroscience applications

Introduction: Why care about the distribution of data?



Introduction: Why care about the distribution of data?

Problem: Analyzing high dimensional data is hard

Solution: Variational Methods

Example: variational autoencoder

Discussion: Neuroscience applications

2019-07-10

Introduction: Why care about the distribution of data?

└ Introduction: Why care about the distribution of data?

└ Introduction: Why care about the distribution of data?

Introduction: Why care about the distribution of data?

Problem: Analyzing high dimensional data is hard

Solution: Variational Methods

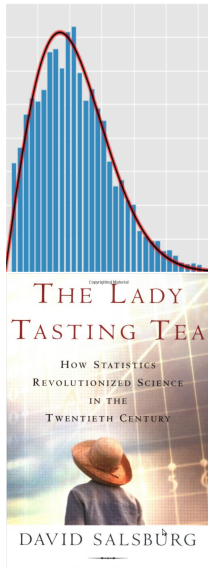
Example: variational autoencoder

Discussion: Neuroscience applications

The probability distribution revolution



- ▶ Karl Pearson (1857-1936) came with the idea that scientific measurements should be conceived as coming from probability distributions.
- ▶ Scientific measurements are just random reflections of the underlying truth that is the distribution.
- ▶ “A great book on the history of statistics” → Aaron



2019-07-10

- Introduction: Why care about the distribution of data?
 - Introduction: Why care about the distribution of data?
 - The probability distribution revolution

- ▶ Karl Pearson (1857-1936) came with the idea that scientific measurements should be conceived as coming from probability distributions.
- ▶ Scientific measurements are just random reflections of the underlying truth that is the distribution.
- ▶ “A great book on the history of statistics” → Aaron



The power of probability distributions



Distributions allow scientists to:

- ▶ Understand scientific measurement
- ▶ Predict the probability of specific data
- ▶ Test specific hypothesis (p-values)
- ▶ Produce generative models
- ▶ Better conceptual understanding data.

2019-07-10

Introduction: Why care about the distribution of data?
└ Introduction: Why care about the distribution of data?
└ The power of probability distributions

Distributions allow scientists to:

- ▶ Understand scientific measurement
- ▶ Predict the probability of specific data
- ▶ Test specific hypothesis (p-values)
- ▶ Produce generative models
- ▶ Better conceptual understanding data.



Low-Dimensional:

- ▶ Great tools to fit and understand the underlying probability distribution of data.

High-Dimensional:

- ▶ In some cases, classical statistical tools are insufficient.
- ▶ Problematic for modern neuroscience:
- ▶ Thousands of electrodes.
- ▶ Millions of voxels.

2019-07-10

Introduction: Why care about the distribution of data?
└ Introduction: Why care about the distribution of data?
└ Estimating distributions from data

Low-Dimensional:

- ▶ Great tools to fit and understand the underlying probability distribution of data.

High-Dimensional:

- ▶ In some cases, classical statistical tools are insufficient.
- ▶ Problematic for modern neuroscience:
- ▶ Thousands of electrodes.
- ▶ Millions of voxels.



Estimate the probability distribution of high-dimensional neural data.

- ▶ Compute the probability of observing a particular neural state.
- ▶ Sample neural states/trajectory from the estimate.
- ▶ Generate statistics / test hypotheses
- ▶ Estimate latent factors/states that drive the observations.
- ▶ Reduce dimensionality (with advantages over other methods, e.g. PCA, T-SNE)

2019-07-10

Introduction: Why care about the distribution of data?
└ Introduction: Why care about the distribution of data?
└ Variational Bayesian Methods

Estimate the probability distribution of high-dimensional neural data.

- ▶ Compute the probability of observing a particular neural state.
- ▶ Sample neural states/trajectory from the estimate.
- ▶ Generate statistics / test hypotheses
- ▶ Estimate latent factors/states that drive the observations.
- ▶ Reduce dimensionality (with advantages over other methods, e.g. PCA, T-SNE)

Problem: Analyzing high dimensional data is hard



Introduction: Why care about the distribution of data?

Problem: Analyzing high dimensional data is hard

Solution: Variational Methods

Example: variational autoencoder

Discussion: Neuroscience applications

2019-07-10

Problem: Analyzing high dimensional data is hard

└ Problem: Analyzing high dimensional data is hard

└ Problem: Analyzing high dimensional data is hard

Introduction: Why care about the distribution of data?

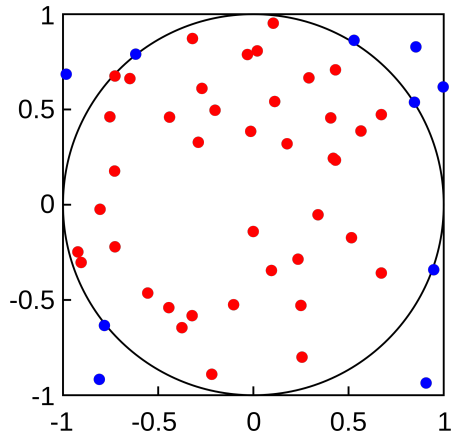
Problem: Analyzing high dimensional data is hard

Solution: Variational Methods

Example: variational autoencoder

Discussion: Neuroscience applications

Approximating the area of a circle with Monte Carlo



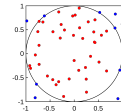
40/50 samples in circle gives area of $4 * 0.8 = 3.2 \approx \pi$. $\sim 2\%$ error

CCO 1.0, wikimedia, MonteCarloIntegrationCircle.svg

2019-07-10

Problem: Analyzing high dimensional data is hard
└ Problem: Analyzing high dimensional data is hard

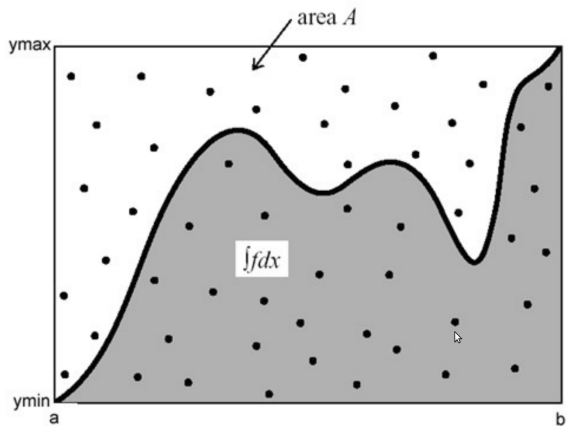
└ Approximating the area of a circle with Monte Carlo



40/50 samples in circle gives area of $4 * 0.8 = 3.2 \approx \pi$. $\sim 2\%$ error

CCO 1.0, wikimedia, MonteCarloIntegrationCircle.svg

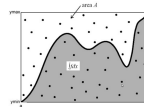
Useful for analytically intractable integrals



Robert Lin - Monte Carlo Integration

2019-07-10

- Problem: Analyzing high dimensional data is hard
 - Problem: Analyzing high dimensional data is hard
 - Useful for analytically intractable integrals



Robert Lin - Monte Carlo Integration

Extension to unit hypercube



We can approximate a high-dimensional integral using a Monte Carlo approximation:

$$\int_0^1 \cdots \int_0^1 g(x_1, \dots, x_n) dx_1, \dots, dx_n \approx \frac{1}{N} \sum_{j=1}^N g(\bar{x}_j)$$

where $\bar{x}_1, \dots, \bar{x}_N \sim \mathcal{U}(0, 1)$ is the i^{th} random sample

2019-07-10

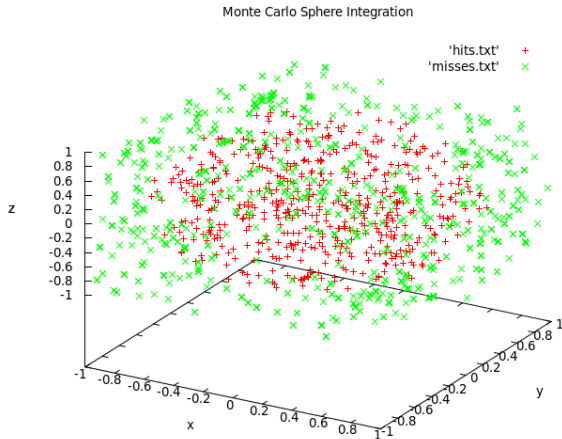
Problem: Analyzing high dimensional data is hard
└ Problem: Analyzing high dimensional data is hard
└ Extension to unit hypercube

We can approximate a high-dimensional integral using a Monte Carlo approximation:

$$\int_0^1 \cdots \int_0^1 g(x_1, \dots, x_n) dx_1, \dots, dx_n \approx \frac{1}{N} \sum_{j=1}^N g(\bar{x}_j)$$

where $\bar{x}_1, \dots, \bar{x}_N \sim \mathcal{U}(0, 1)$ is the i^{th} random sample

Performance scales poorly with number of dimensions

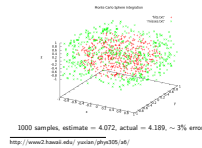


1000 samples, estimate = 4.072, actual = 4.189, $\sim 3\%$ error

<http://www2.hawaii.edu/~yuxian/phys305/a6/>

2019-07-10

Problem: Analyzing high dimensional data is hard
└ Problem: Analyzing high dimensional data is hard
└ Performance scales poorly with number of dimensions



1000 samples, estimate = 4.072, actual = 4.189, $\sim 3\%$ error
<http://www2.hawaii.edu/~yuxian/phys305/a6/>

Solution: Variational Methods



Introduction: Why care about the distribution of data?

Problem: Analyzing high dimensional data is hard

Solution: Variational Methods

Example: variational autoencoder

Discussion: Neuroscience applications

2019-07-10

Solution: Variational Methods

└ Solution: Variational Methods

└ Solution: Variational Methods

Introduction: Why care about the distribution of data?

Problem: Analyzing high dimensional data is hard

Solution: Variational Methods

Example: variational autoencoder

Discussion: Neuroscience applications

Alternate approach: Variational Bayes

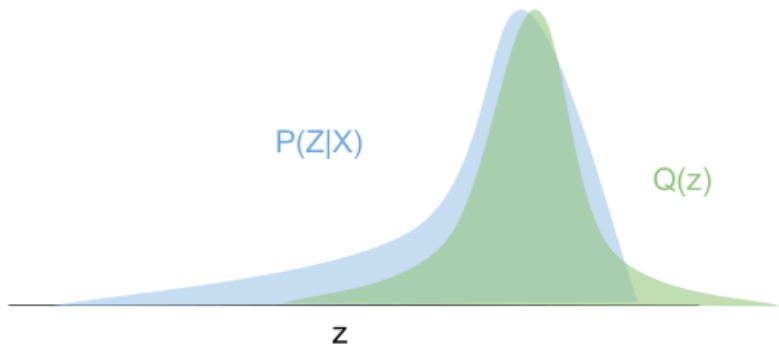


2019-07-10

Solution: Variational Methods

└ Solution: Variational Methods

└ Alternate approach: Variational Bayes



Example: variational autoencoder



Introduction: Why care about the distribution of data?

Problem: Analyzing high dimensional data is hard

Solution: Variational Methods

Example: variational autoencoder

Discussion: Neuroscience applications

2019-07-10

Example: variational autoencoder

└ Example: variational autoencoder

└ Example: variational autoencoder

Introduction: Why care about the distribution of data?

Problem: Analyzing high dimensional data is hard

Solution: Variational Methods

Example: variational autoencoder

Discussion: Neuroscience applications

Summary



- ▶ Problem: Analyzing high dimensional data is hard
 - ▶
- ▶ Solution: Variational Methods
 - ▶
- ▶ Example: variational autoencoder
 - ▶

2019-07-10

Example: variational autoencoder
└ Example: variational autoencoder
└ Summary

- ▶ Problem: Analyzing high dimensional data is hard
 - ▶
- ▶ Solution: Variational Methods
 - ▶
- ▶ Example: variational autoencoder
 - ▶

Discussion: Neuroscience applications



Introduction: Why care about the distribution of data?

Problem: Analyzing high dimensional data is hard

Solution: Variational Methods

Example: variational autoencoder

Discussion: Neuroscience applications

2019-07-10

Discussion: Neuroscience applications

└ Discussion: Neuroscience applications

└ Discussion: Neuroscience applications

Introduction: Why care about the distribution of data?

Problem: Analyzing high dimensional data is hard

Solution: Variational Methods

Example: variational autoencoder

Discussion: Neuroscience applications



body

2019-07-10

- Discussion: Neuroscience applications
 - └ Discussion: Neuroscience applications
 - └ (backup slide) LFADS

body