Chair of Aerodynamics and Fluid Mechanics
Department of Mechanical Engineering
Technical University of Munich

Probabilistic Programming for Scientific Discovery

Lecture 2

Ludger Paehler
Lviv Data Science Summer School

July 29, 2020

Table of Contents

Outline

Approaches to Inference - the Inference Engines

- Monte-Carlo
- Variational Inference

Probabilistic Programming Frameworks

- Stan
- Venture
- Anglican
- PyMC3
- TensorFlow Probability
- Pyro & NumPyro
- Edward2
- Gen
- PyProb
- Turing

Practical Introduction to a Probabilistic Programming Framework

Extending the ideas to a more complex examples

Approaches to Inference - the Inference Engines

- A typical probabilistic programming system consists of:
  - A domain-specific language (DSL), which enables the user to express his model using the language-specific primitives
  - Provides a library of inference algorithms, which enable inference on probabilistic models definable in the DSL.
  - Prevalent Monte-Carlo and variational inference approaches have their own specific sets of strength, it is hence important to understand the inference algorithms one utilizes

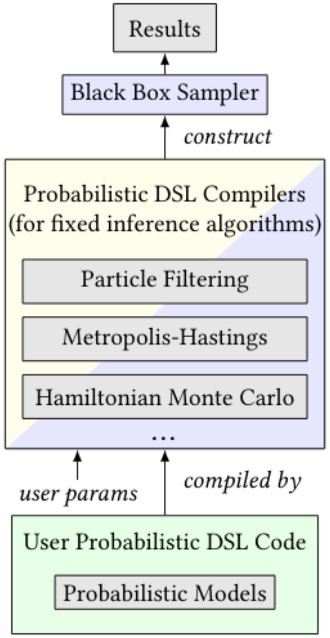

Figure: Structure of a typical probabilistic programming system. Source: *Gen: A General-Purpose Probabilistic Programming Systems with Programmable Inference*

Monte-Carlo Approaches to Inference

Hamiltonian Monte-Carlo [1]

- Hamiltonian Monte-Carlo 1

[1] Neal, R.M., 2011. MCMC using Hamiltonian dynamics. Handbook of markov chain monte carlo, 2(11), p.2.

Monte-Carlo Approaches to Inference

Hamiltonian Monte-Carlo

- Hamiltonian Monte-Carlo 2

Chair of Aerodynamics and Fluid Mechanics
Department of Mechanical Engineering
Technical University of Munich

Monte-Carlo Approaches to Inference

Hamiltonian Monte-Carlo

- Hamiltonian Monte-Carlo 3

Monte-Carlo Approaches to Inference

Hamiltonian Monte-Carlo

- Hamiltonian Monte-Carlo 4

Monte-Carlo Approaches to Inference

Random-Walk Metropolis Hastings [2] [3]

- Random-Walk Metropolis Hastings 1

[2] Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A. and Rubin, D.B., 2013. Bayesian data analysis. CRC press.

[3] Gilks, W.R. and Richardson, S., S. and Spiegelhalter, D.(1996). Markov chain Monte Carlo in practice. London, UK: Chapman k Hall/CRC.

Monte-Carlo Approaches to Inference

Random-Walk Metropolis Hastings

- Random-Walk Metropolis Hastings 2

Chair of Aerodynamics and Fluid Mechanics
Department of Mechanical Engineering
Technical University of Munich

Monte-Carlo Approaches to Inference

Random-Walk Metropolis Hastings

- Random-Walk Metropolis Hastings 3

Monte-Carlo Approaches to Inference

Random-Walk Metropolis Hastings

- Random-Walk Metropolis Hastings 4

Monte-Carlo Approaches to Inference

Stochastic-Gradient Langevin Dynamics [4] [5]

- Stochastic-Gradient Langevin Dynamics 1

[4] Welling, M. and Teh, Y.W., 2011. Bayesian learning via stochastic gradient Langevin dynamics. In Proceedings of the 28th international conference on machine learning (ICML-11) (pp. 681-688).

[5] Brosse, N., Durmus, A. and Moulines, E., 2018. The promises and pitfalls of stochastic gradient Langevin dynamics. In Advances in Neural Information Processing Systems (pp. 8268-8278).

Monte-Carlo Approaches to Inference

Stochastic-Gradient Langevin Dynamics

- Stochastic-Gradient Langevin Dynamics 2

Chair of Aerodynamics and Fluid Mechanics
Department of Mechanical Engineering
Technical University of Munich

Monte-Carlo Approaches to Inference

Stochastic-Gradient Langevin Dynamics

- Stochastic-Gradient Langevin Dynamics 3

Monte-Carlo Approaches to Inference

Stochastic-Gradient Langevin Dynamics

- Stochastic-Gradient Langevin Dynamics 4

Variational Approaches to Inference

Variational Inference [6] [7]

- Variational Inference 1

[6]Blei, D.M., Kucukelbir, A. and McAuliffe, J.D., 2017. Variational inference: A review for statisticians. Journal of the American statistical Association, 112(518), pp.859-877.

[7]Zhang, C., Bütepage, J., Kjellström, H. and Mandt, S., 2018. Advances in variational inference. IEEE transactions on pattern analysis and machine intelligence, 41(8), pp.2008-2026.

Variational Approaches to Inference

Variational Inference

- Variational Inference 2

Chair of Aerodynamics and Fluid Mechanics
Department of Mechanical Engineering
Technical University of Munich

Variational Approaches to Inference

Variational Inference

- Variational Inference 3

Variational Approaches to Inference

Variational Inference

- Variational Inference 4

Variational Approaches to Inference

Stochastic Variational Inference [8] [9]

- Stochastic Variational Inference 1

[8] Hoffman, M.D., Blei, D.M., Wang, C. and Paisley, J., 2013. Stochastic variational inference. The Journal of Machine Learning Research, 14(1), pp.1303-1347.

[9] Robbins, H. and Monro, S., 1951. A stochastic approximation method. The annals of mathematical statistics, pp.400-407.

Variational Approaches to Inference

Stochastic Variational Inference

- Stochastic Variational Inference 2

Chair of Aerodynamics and Fluid Mechanics
Department of Mechanical Engineering
Technical University of Munich

Variational Approaches to Inference

Stochastic Variational Inference

- Stochastic Variational Inference 3

Variational Approaches to Inference

Stochastic Variational Inference

- Stochastic Variational Inference 4

Variational Approaches to Inference

Automatic Differentiation Variational Inference [10] [11]

- Automatic Differentiation Variational Inference 1

[10] Kucukelbir, A., Tran, D., Ranganath, R., Gelman, A. and Blei, D.M., 2017. Automatic differentiation variational inference. The Journal of Machine Learning Research, 18(1), pp.430-474.

[11] Kucukelbir, A., Ranganath, R., Gelman, A. and Blei, D., 2015. Automatic variational inference in Stan. In Advances in neural information processing systems (pp. 568-576).

Chair of Aerodynamics and Fluid Mechanics
Department of Mechanical Engineering
Technical University of Munich

Variational Approaches to Inference

Automatic Differentiation Variational Inference

- Automatic Differentiation Variational Inference 2

Variational Approaches to Inference

Automatic Differentiation Variational Inference

- Automatic Differentiation Variational Inference 3

Chair of Aerodynamics and Fluid Mechanics
Department of Mechanical Engineering
Technical University of Munich

Variational Approaches to Inference

Automatic Differentiation Variational Inference

- Automatic Differentiation Variational Inference 4

Variational Approaches to Inference

Black Box Variational Inference [12] [13]

- Black Box Variational Inference 1

[12] Ranganath, R., Gerrish, S. and Blei, D., 2014, April. Black box variational inference. In Artificial Intelligence and Statistics (pp. 814-822).

[13] Chu, C., Minami, K. and Fukumizu, K., 2020. The equivalence between Stein variational gradient descent and black-box variational inference. arXiv preprint arXiv:2004.01822.

Chair of Aerodynamics and Fluid Mechanics
Department of Mechanical Engineering
Technical University of Munich

Variational Approaches to Inference

Black Box Variational Inference

- Black Box Variational Inference 2

Variational Approaches to Inference

Black Box Variational Inference

- Black Box Variational Inference 3

Variational Approaches to Inference

Black Box Variational Inference

- Black Box Variational Inference 4

Outline

Approaches to Inference - the Inference Engines
- Monte-Carlo
- Variational Inference

Probabilistic Programming Frameworks
- Stan
- Venture
- Anglican
- PyMC3
- TensorFlow Probability
- Pyro & NumPyro
- Edward2
- Gen
- PyProb
- Turing

Practical Introduction to a Probabilistic Programming Framework

Extending the ideas to a more complex examples

Stan [14]

Overview

- General overview of the purpose behind Stan

[14] Carpenter, B., Gelman, A., Hoffman, M.D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P. and Riddell, A., 2017. Stan: A probabilistic programming language. Journal of statistical software, 76(1).

Stan

Syntax

- Example code to get a grasp for the syntax

Stan

Application Performance

- Example applications

Chair of Aerodynamics and Fluid Mechanics
Department of Mechanical Engineering
Technical University of Munich

Venture [15] [16]

Overview

- General overview of the purpose behind venture

[15] Mansinghka, V., Selsam, D. and Perov, Y., 2014. Venture: a higher-order probabilistic programming platform with programmable inference. arXiv preprint arXiv:1404.0099.

[16] Goodman, N., Mansinghka, V., Roy, D.M., Bonawitz, K. and Tenenbaum, J.B., 2012. Church: a language for generative models. arXiv preprint arXiv:1206.3255.

Venture

Syntax

- Example code to get a gauge for the syntax

Venture

Application Performance

- Application performance

Anglican [17]

Overview

- General overview of the purpose behind Anglican

[17] Tolpin, D., van de Meent, J.W., Yang, H. and Wood, F., 2016, August. Design and implementation of probabilistic programming language anglican. In Proceedings of the 28th Symposium on the Implementation and Application of Functional programming Languages (pp. 1-12).

Anglican

Syntax

- Example code

Anglican

Application Performance

- Application performance

PyMC3 [18]

Overview

- General overview of the purpose behind PcMC3

[18] Salvatier, J., Wiecki, T.V. and Fonnesbeck, C., 2016. Probabilistic programming in Python using PyMC3. PeerJ Computer Science, 2, p.e55.

PyMC3

Syntax

- Example code

PyMC3

Application Performance

- Application performance

TensorFlow Probability [19]

Overview

- General overview of the purpose behind Tensorflow Probability

[19] Dillon, J.V., Langmore, I., Tran, D., Brevdo, E., Vasudevan, S., Moore, D., Patton, B., Alemi, A., Hoffman, M. and Saurous, R.A., 2017. Tensorflow distributions. arXiv preprint arXiv:1711.10604.

TensorFlow Probability

Syntax

- Example code

TensorFlow Probability

Application Performance

- Application performance

Pyro [20] & NumPyro [21]

Overview

- General overview of the purpose behind Pyro & NumPyro

[20] Bingham, E., Chen, J.P., Jankowiak, M., Obermeyer, F., Pradhan, N., Karaletsos, T., Singh, R., Szerlip, P., Horsfall, P. and Goodman, N.D., 2019. Pyro: Deep universal probabilistic programming. The Journal of Machine Learning Research, 20(1), pp.973-978.

[21] Phan, D., Pradhan, N. and Jankowiak, M., 2019. Composable effects for flexible and accelerated probabilistic programming in NumPyro. arXiv preprint arXiv:1912.11554.

Pyro & NumPyro

Syntax

- Example code of Pyro & NumPyro

Pyro & NumPyro

Application Performance

- Application performance

Edward2 [22] [23]

Overview

- General overview of the purpose behind Edward2

[22] Tran, D., Hoffman, M.W., Moore, D., Suter, C., Vasudevan, S. and Radul, A., 2018. Simple, distributed, and accelerated probabilistic programming. In Advances in Neural Information Processing Systems (pp. 7598-7609).

[23] Tran, D., Dusenberry, M., van der Wilk, M. and Hafner, D., 2019. Bayesian layers: A module for neural network uncertainty. In Advances in Neural Information Processing Systems (pp. 14660-14672).

Chair of Aerodynamics and Fluid Mechanics
Department of Mechanical Engineering
Technical University of Munich

Edward2

Syntax

- Example code of Edward2

Chair of Aerodynamics and Fluid Mechanics
Department of Mechanical Engineering
Technical University of Munich

Edward2

Application Performance

- Application performance

Gen [24] [25]

Overview

- General overview of the purpose behind Gen

[24] Cusumano-Towner, M.F., Saad, F.A., Lew, A.K. and Mansinghka, V.K., 2019, June. Gen: a general-purpose probabilistic programming system with programmable inference. In Proceedings of the 40th ACM SIGPLAN Conference on Programming Language Design and Implementation (pp. 221-236).

[25] Cusumano-Towner, M., Lew, A.K. and Mansinghka, V.K., 2020. Automating Involutive MCMC using Probabilistic and Differentiable Programming. arXiv preprint arXiv:2007.09871.

Gen

Programmable Inference

- Illustration of Gen's slightly different construction compared to the other frameworks

Chair of Aerodynamics and Fluid Mechanics
Department of Mechanical Engineering
Technical University of Munich

Gen

Syntax

- Example code of Gen

Gen

Application Performance

- Application performance

PyProb [26]

Overview

- General overview of the purpose behind PyProb

[26] Baydin, A.G., Shao, L., Bhimji, W., Heinrich, L., Naderiparizi, S., Munk, A., Liu, J., Gram-Hansen, B., Louppe, G., Meadows, L. and Torr, P., 2019. Efficient probabilistic inference in the quest for physics beyond the standard model. In Advances in neural information processing systems (pp. 5459-5472).

PyProb

Syntax

- Example code of PyProb

PyProb

Application Performance

- Application performance

Turing [27]

Overview

- General overview of the purpose behind Turing

[27] Ge, H., Xu, K. and Ghahramani, Z., 2018, March. Turing: A Language for Flexible Probabilistic Inference. In International Conference on Artificial Intelligence and Statistics (pp. 1682-1690).

Turing

Syntax

- Example code of Turing

Application performance

- Application performance

Probabilistic Programming Frameworks

Summary

- Summary of all probabilistic programming frameworks in a single table

Chair of Aerodynamics and Fluid Mechanics
Department of Mechanical Engineering
Technical University of Munich

Outline

- Approaches to Inference - the Inference Engines
  - Monte-Carlo
  - Variational Inference
- Probabilistic Programming Frameworks
  - Stan
  - Venture
  - Anglican
  - PyMC3
  - TensorFlow Probability
  - Pyro & NumPyro
  - Edward2
  - Gen
  - PyProb
  - Turing
- Practical Introduction to a Probabilistic Programming Framework
- Extending the ideas to a more complex examples

Introduction to Turing

- We will do our first steps in a probabilistic programming framework with Turing covering
  - The modelling syntax
  - Sampling
  - Accessing the trace
  - Automatic differentiation
  - Working with dynamic Hamiltonian Monte-Carlo
- All content can be accessed in the Jupyter notebook **IntrotoTuring.ipynb**

Outline

- Approaches to Inference - the Inference Engines
  - Monte-Carlo
  - Variational Inference
- Probabilistic Programming Frameworks
  - Stan
  - Venture
  - Anglican
  - PyMC3
  - TensorFlow Probability
  - Pyro & NumPyro
  - Edward2
  - Gen
  - PyProb
  - Turing
- Practical Introduction to a Probabilistic Programming Framework
- Extending the ideas to a more complex examples

More Complex Example in Turing

Model-based inference for causal effects in completely randomized experiments

- Expanding on the simple syntax we will now move to a more complex case:
  - Starting with a Bayesian perspective on causal inference
  - Assignment mechanisms
  - Posterior inference
- All content can be accessed in the Jupyter notebook **MBInferenceforCausalEffects.ipynb**