Variational Inference for Inverse Reinforcement Learning with Gaussian Processes

Paulius Dilkas

School of Computing Science
Sir Alwyn Williams Building
University of Glasgow
G12 8QQ

Masters project proposal

Date of submission placed here

Contents

1 Introduction

briefly explain the context of the project problem

1.1 A subsection

Please note your proposal need not follow the included section headings - this is only a suggested structure. Also add subsections etc as required

2 Statement of Problem

clearly state the problem to be addressed in your forthcoming project. Explain why it would be worthwhile to solve this problem.

3 Background Survey

present an overview of relevant previous work including articles, books, and existing software products. Critically evaluate the strengths and weaknesses of the previous work.

4 Proposed Approach

state how you propose to solve the software development problem. Show that your proposed approach is feasible, but identify any risks.

5 Work Plan

show how you plan to organize your work, identifying intermediate deliverables and dates.

6 Notes on papers (to be removed)

6.1 Miscellaneous

(Directed) similarity between MDPs using restricted Boltzmann machines [6]

Chapter 6 on distance measures [23]

The PhD thesis behind maximum causal entropy [35]

6.2 Gaussian Processes

Your recommended book [28]

Simple introduction to GPs for time-series modelling [30]

Spectral kernels [33]

GPs over graphs instead of vectors (haven't actually read) [32]

Another introduction from physics (skimmed through) [18]

Learning a GP from very little data [25]

One GP for multiple correlated output variables [3]

Kernels for categorical and count data [31]

6.3 Interpretability

Learning latent factors [21]

The behaviour of Reddit users [12]

6.4 Inverse Reinforcement Learning

One of the first papers on the topic [24]

Follow-up on the previous paper [1]

Bayesian setting [27]

Learning optimal composite features [11]

The main paper [20]

A different take on IRL with GPs [26]

The paper that introduced maximum entropy into IRL [36]

IRL for large state spaces (haven't read) [7]

Multiple reward functions [10]

6.4.1 Multiple Strategies

EM clustering [2]

Structured priors [13]

There are more, but I haven't gotten to them yet.

6.5 Variational Inference

Chapter 10 on approximate inference [4]

Part IV on probabilities and inference [22]

A recent review [5]

Normalizing flows [29]

Linear VI for GPs [9]

Stochastic VI [17]

Structured stochastic VI (haven't read) [16]

Another review of recent advances [34]

Sparse VI for GP [15]

Sparse GPs [8]

IRL via deep GP [19]

References

[1] Pieter Abbeel and Andrew Y. Ng. Apprenticeship learning via inverse reinforcement learning. In Carla E. Brodley, editor, *Machine Learning, Proceedings of the Twenty-first International Conference (ICML 2004), Banff, Alberta, Canada, July 4-8, 2004*, volume 69 of *ACM International Conference Proceeding Series*. ACM, 2004.

[2] Monica Babes, Vukosi N. Marivate, Kaushik Subramanian, and Michael L. Littman. Apprenticeship learning about multiple intentions. In Lise Getoor and Tobias Scheffer, editors, *Proceedings of the 28th International Conference on Machine Learning, ICML*

2011, Bellevue, Washington, USA, June 28 - July 2, 2011, pages 897–904. Omnipress, 2011.

[3] Ilias Bilionis, Nicholas Zabaras, Bledar A. Konomi, and Guang Lin. Multi-output separable Gaussian process: Towards an efficient, fully Bayesian paradigm for uncertainty quantification. *J. Comput. Physics*, 241:212–239, 2013.

[4] Christopher M. Bishop. *Pattern recognition and machine learning, 5th Edition*. Information science and statistics. Springer, 2007.

[5] David M Blei, Alp Kucukelbir, and Jon D McAuliffe. Variational inference: A review for statisticians. *Journal of the American Statistical Association*, 112(518):859–877, 2017.

[6] H. Bou Ammar, E. Eaton, M.E. Taylor, D.C. Mocanu, K. Driessens, G. Weiss, and K.P. Tuyls. An automated measure of MDP similarity for transfer in reinforcement learning. In *Proceedings of the MLIS-2014 collocated with The Twenty-Eighth AAAI Conference on Artificial Intelligence, 27-28 July 2014, Quebec City, Canada*, pages 31–37, 2014.

[7] Abdeslam Boularias, Jens Kober, and Jan Peters. Relative entropy inverse reinforcement learning. In Geoffrey J. Gordon, David B. Dunson, and Miroslav Dudík, editors, *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics, AISTATS 2011, Fort Lauderdale, USA, April 11-13, 2011*, volume 15 of *JMLR Proceedings*, pages 182–189. JMLR.org, 2011.

[8] Joaquin Quiñonero Candela and Carl Edward Rasmussen. A unifying view of sparse approximate Gaussian process regression. *Journal of Machine Learning Research*, 6:1939–1959, 2005.

[9] Ching-An Cheng and Byron Boots. Variational inference for Gaussian process models with linear complexity. In Guyon et al. [14], pages 5190–5200.

[10] Jaedeug Choi and Kee-Eung Kim. Nonparametric Bayesian inverse reinforcement learning for multiple reward functions. In Peter L. Bartlett, Fernando C. N. Pereira, Christopher J. C. Burges, Léon Bottou, and Kilian Q. Weinberger, editors, *Advances in Neural Information Processing Systems 25: 26th Annual Conference on Neural Information Processing Systems 2012. Proceedings of a meeting held December 3-6, 2012, Lake Tahoe, Nevada, United States.*, pages 314–322, 2012.

[11] Jaedeug Choi and Kee-Eung Kim. Bayesian nonparametric feature construction for inverse reinforcement learning. In Francesca Rossi, editor, *IJCAI 2013, Proceedings of the 23rd International Joint Conference on Artificial Intelligence, Beijing, China, August 3-9, 2013*, pages 1287–1293. IJCAI/AAAI, 2013.

[12] Sanmay Das and Allen Lavoie. The effects of feedback on human behavior in social media: an inverse reinforcement learning model. In Ana L. C. Bazzan, Michael N. Huhns, Alessio Lomuscio, and Paul Scerri, editors, *International conference on Autonomous Agents and Multi-Agent Systems, AAMAS '14, Paris, France, May 5-9, 2014*, pages 653–660. IFAAMAS/ACM, 2014.

[13] Christos Dimitrakakis and Constantin A. Rothkopf. Bayesian multitask inverse reinforcement learning. In Scott Sanner and Marcus Hutter, editors, *Recent Advances in Reinforcement Learning - 9th European Workshop, EWRL 2011, Athens, Greece, September 9-11, 2011, Revised Selected Papers*, volume 7188 of *Lecture Notes in Computer Science*, pages 273–284. Springer, 2011.

[14] Isabelle Guyon, Ulrike von Luxburg, Samy Bengio, Hanna M. Wallach, Rob Fergus, S. V. N. Vishwanathan, and Roman Garnett, editors. *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, 4-9 December 2017, Long Beach, CA, USA*, 2017.

[15] James Hensman, Nicolas Durrande, and Arno Solin. Variational Fourier features for Gaussian processes. *Journal of Machine Learning Research*, 18:151:1–151:52, 2017.

[16] Matthew D. Hoffman and David M. Blei. Structured stochastic variational inference. In Guy Lebanon and S. V. N. Vishwanathan, editors, *Proceedings of the Eighteenth International Conference on Artificial Intelligence and Statistics, AISTATS 2015, San Diego, California, USA, May 9-12, 2015*, volume 38 of *JMLR Workshop and Conference Proceedings*. JMLR.org, 2015.

[17] Matthew D. Hoffman, David M. Blei, Chong Wang, and John William Paisley. Stochastic variational inference. *Journal of Machine Learning Research*, 14(1):1303–1347, 2013.

[18] David J.C. MacKay. Introduction to Gaussian processes. 168, 01 1998.

[19] Ming Jin, Andreas C. Damianou, Pieter Abbeel, and Costas J. Spanos. Inverse reinforcement learning via deep Gaussian process. In Gal Elidan, Kristian Kersting, and Alexander T. Ihler, editors, *Proceedings of the Thirty-Third Conference on Uncertainty in Artificial Intelligence, UAI 2017, Sydney, Australia, August 11-15, 2017*. AUAI Press, 2017.

[20] Sergey Levine, Zoran Popovic, and Vladlen Koltun. Nonlinear inverse reinforcement learning with Gaussian processes. In John Shawe-Taylor, Richard S. Zemel, Peter L. Bartlett, Fernando C. N. Pereira, and Kilian Q. Weinberger, editors, *Advances in Neural Information Processing Systems 24: 25th Annual Conference on Neural Information Processing Systems 2011. Proceedings of a meeting held 12-14 December 2011, Granada, Spain.*, pages 19–27, 2011.

[21] Yunzhu Li, Jiaming Song, and Stefano Ermon. InfoGAIL: Interpretable imitation learning from visual demonstrations. In Guyon et al. [14], pages 3815–3825.

[22] David J. C. MacKay. *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press, New York, NY, USA, 2003.

[23] B. McCune, J.B. Grace, and D.L. Urban. *Analysis of Ecological Communities*. MjM Software Design, 2002.

[24] Andrew Y. Ng and Stuart J. Russell. Algorithms for inverse reinforcement learning. In Pat Langley, editor, *Proceedings of the Seventeenth International Conference on Machine Learning (ICML 2000), Stanford University, Stanford, CA, USA, June 29 - July 2, 2000*, pages 663–670. Morgan Kaufmann, 2000.

[25] John C. Platt, Christopher J. C. Burges, S. Swenson, C. Weare, and A. Zheng. Learning a Gaussian process prior for automatically generating music playlists. In Thomas G. Dietterich, Suzanna Becker, and Zoubin Ghahramani, editors, *Advances in Neural Information Processing Systems 14 [Neural Information Processing Systems: Natural and Synthetic, NIPS 2001, December 3-8, 2001, Vancouver, British Columbia, Canada]*, pages 1425–1432. MIT Press, 2001.

[26] Qifeng Qiao and Peter A. Beling. Inverse reinforcement learning with gaussian process. *CoRR*, abs/1208.2112, 2012.

[27] Deepak Ramachandran and Eyal Amir. Bayesian inverse reinforcement learning. In Manuela M. Veloso, editor, *IJCAI 2007, Proceedings of the 20th International Joint Conference on Artificial Intelligence, Hyderabad, India, January 6-12, 2007*, pages 2586–2591, 2007.

[28] Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian processes for machine learning*. Adaptive computation and machine learning. MIT Press, 2006.

[29] Danilo Jimenez Rezende and Shakir Mohamed. Variational inference with normalizing flows. In Francis R. Bach and David M. Blei, editors, *Proceedings of the 32nd International Conference on Machine Learning, ICML 2015, Lille, France, 6-11 July 2015*, volume 37 of *JMLR Workshop and Conference Proceedings*, pages 1530–1538. JMLR.org, 2015.

[30] Stephen J. Roberts, Matt Osborne, Mark Ebden, Steve Reece, Neale Gibson, and Suzanne Aigrain. Gaussian processes for time-series modelling. *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences*, 371 1984:20110550, 2013.

[31] Terrance Savitsky, Marina Vannucci, and Naijun Sha. Variable selection for nonparametric Gaussian process priors: Models and computational strategies. *Statistical science: a review journal of the Institute of Mathematical Statistics*, 26(1):130, 2011.

[32] Arun Venkitaraman, Saikat Chatterjee, and Peter Händel. Gaussian processes over graphs. *CoRR*, abs/1803.05776, 2018.

[33] Andrew Wilson and Ryan Adams. Gaussian process kernels for pattern discovery and extrapolation. In Sanjoy Dasgupta and David McAllester, editors, *Proceedings of the 30th International Conference on Machine Learning*, volume 28 of *Proceedings of Machine Learning Research*, pages 1067–1075, Atlanta, Georgia, USA, 17–19 Jun 2013. PMLR.

[34] Cheng Zhang, Judith Bütepage, Hedvig Kjellström, and Stephan Mandt. Advances in variational inference. *CoRR*, abs/1711.05597, 2017.

[35] Brian D. Ziebart. *Modeling Purposeful Adaptive Behavior with the Principle of Maximum Causal Entropy*. PhD thesis, Pittsburgh, PA, USA, 2010. AAI3438449.

[36] Brian D Ziebart, Andrew L Maas, J Andrew Bagnell, and Anind K Dey. Maximum entropy inverse reinforcement learning. In *AAAI*, volume 8, pages 1433–1438. Chicago, IL, USA, 2008.