Topics in Probability (STAT 547C)
Fall 2019
https://github.com/GiankDiluvi/gd-stat547c

Gian Carlo Diluvi
gian.diluvi@stat.ubc.ca

Exchangeability in Gaussian Process Regression

Project Outline

1 Title

The working title of my project is *Exchangeability in Gaussian Process Regression*.

2 Background

Exchangeability is a fundamental—although oftentimes overlooked—concept in probability which has elegant theoretical implications, namely, de Finetti's integral representation theorems [de 30]. Furthermore, the number of potential uses of exchangeability in contemporary statistics seems to be growing. For instance, [Kin78; Ald10; OR15] discuss applications of exchangeability in population genetics and general random structures.

Another area in Bayesian computational statistics that has enjoyed a surge in popularity is the use of Gaussian Process (GP) regression for modelling complex phenomena. For example, [Fra18] surveys applications of GPs in Bayesian Optimization, [PD16] in disease spreading modeling, and [Woo+17] in Bayesian design of experiments.

However, these two concepts have so far been treated separately. In this line of thought, in [McC05] McCullagh discusses exchangeability in the context of regression models. I am interested in giving a more detailed treatment of these ideas in the case of Gaussian Process regression. Specifically, I want to study if a result analogous to de Finetti's can be conjectured in this setting, as well as the potential practical implications of studying GP regression from an exchangeability point of view.

3 Technical aspects

The project will draw on technical aspects of the following areas: independence, conditioning and disintegration, stochastic processes, optimization theory.

4 Literature

The key references for this project are:

- [McC05] discusses how to generalize the concept of exchangeability to regression settings.
- [Cam+19] explore a weaker form of exchangeability and its theoretical implications.
- [Kin78; Ber96] give a probabilistic treatment of exchangeability; [de 30] is the original paper in which de Finetti proves his integral representation theorem.
- [RW06] is the go-to reference for Gaussian Process regression.

5 Plan

I will carry out this project with the following sequence of steps:

1. Study the concept of exchangeability, following [Kin78] and [Ber96], and (give a sketch of the derivation of) the de Finetti representation theorem.
2. Review Gaussian Process regression and survey some of its applications (see items in Background section).
3. Discuss the concept of exchangeability in a regression setting as per [McC05] and do an in-depth study of its implications for Gaussian Process regression.
4. If possible, conjecture a representation theorem in this setting, or at least study how such a theorem would look like.
5. Determine if this approach has any practical implications, and if so which ones.

6 Why I'm interested in this topic

I am interested in both Gaussian Process regression and its applications and the concept of exchangeability, and this project seems like a nice way of joining them.

References

[de 30] B. de Finetti. "Funzione caratteristica di un fenomeno aleatorio". In: *R. Academia Nazionale dei Lince* 4.1 (1930), pp. 86–133.

[Kin78] J. F. C. Kingman. "Uses of Exchangeability". In: *The Annals of Probability* 6.2 (1978), pp. 183–197. URL: `http://dx.doi.org/10.1214/aop/1176995566`.

[Ald10] D. J. Aldous. "More uses of exchangeability: representations of complex random structures". In: *Probability and Mathematical Genetics: Papers in Honour of Sir John Kingman*. Ed. by N. H. Bingham and C. M. Goldie. London Mathematical Society Lecture Note Series. Cambridge University Press, 2010, pp. 35–63.

[OR15] P. Orbanz and D. M. Roy. "Bayesian Models of Graphs, Arrays and Other Exchangeable Random Structures". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 37.2 (2015), pp. 437–461.

[Fra18] P. I. Frazier. "A Tutorial on Bayesian Optimization". In: *arXiv e-prints*, arXiv:1807.02811 (2018), arXiv:1807.02811. arXiv: `1807.02811 [stat.ML]`.

[PD16] G. Pokharel and R. Deardon. "Gaussian process emulators for spatial individual-level models of infectious disease". In: *The Canadian Journal of Statistics* 44.4 (2016), pp. 480–501.

[Woo+17] D. C. Woods et al. "Bayesian design of experiments for generalized linear models and dimensional analysis with industrial and scientific application". In: *Quality Engineering* 29.1 (2017), pp. 91–103.

[McC05] P. McCullah. "Exchangeability and regression models". In: *Celebrating Statistics: Papers in honour of Sir David Cox on his 80th birthday*. Oxford": Oxford University Press, 2005. Chap. 4, pp. 89–114.

[Cam+19] T. Campbell et al. "Local Exchangeability". In: *arXiv e-prints*, arXiv:1906.09507 (2019), arXiv:1906.09507. arXiv: `1906.09507 [math.ST]`.

[Ber96] J. M. Bernardo. "The concept of exchangeability and its applications". In: *Far East Journal of Mathematical Sciences* 4 (1996), pp. 111–122.

[RW06] C. E. Rasmussen and C. K. I. Williams. *Gaussian Processes for Machine Learning*. 1st ed. The MIT Press, 2006.