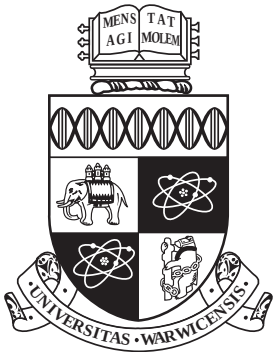
Dissertation in Interdisciplinary Mathematics

Learning parameters of discrete determinantal point processes

Johannes Müller

June 12, 2018



University of Warwick

Department of Mathematics

Acknowledgements

Determinantal point processes are random subsets that exhibit a diversifying behaviour in the sense that the randomly selected points tend to be not similar in some way. This repellent structure first arrose in theortical physics and pure mathematics, but they have recently been used to model a variety of many real world scenarios in a machine learning setup. We aim to give an overview over the main ideas of this approach which is easily accessible even without prior knowledge in the area of machine learning and sometimes omit technical calculations in order to keep the focus on the concepts.

Abstract

Determinantal point processes are random subsets that exhibit a diversifying behaviour in the sense that the randomly selected points tend to be not similar in some way. This repellent structure first arrose in theortical physics and pure mathematics, but they have recently been used to model a variety of many real world scenarios in a machine learning setup. We aim to give an overview over the main ideas of this approach which is easily accessible even without prior knowledge in the area of machine learning and sometimes omit technical calculations in order to keep the focus on the concepts.

Contents

I Introduction and motivating examples **1**
I.1 Motivation . . . 1
I.2 Previous work . . . 1
I.3 Aim and outline of the dissertation . . . 1

II Determinantal points processes: Basic notions and properties **2**
II.1 Historical remarks . . . 2
II.2 Definitions and properties . . . 2
II.3 L-ensembles . . . 2
II.4 Related structures: conditional DPPs, structured DPPs and k-DPPs . . . 2
II.5 The magic properties of DPPs . . . 2
II.6 The mode problem . . . 2

III Learning setups **3**
III.1 What does learning mean? . . . 3
III.2 Motivation for learning DPPs . . . 3
III.3 Direct reconstruction of the kernel . . . 3
III.4 A learning approach to estimating the kernel . . . 3
III.5 Learning the quality functions . . . 3
III.6 Estimating the mixture coefficients of k-DPPs . . . 3

IV Toy examples and experiments **4**
IV.1 Minimal example? . . . 4
IV.2 Points on the line . . . 4
IV.3 Points in the square . . . 4
IV.4 Toy example for quality learning . . . 4

V Summary and conclusion **5**

A Generated code **6**

Bibliography **7**

Chapter I

Introduction and motivating examples

I.1 Motivation

I.2 Previous work

I.3 Aim and outline of the dissertation

Chapter II

Determinantal points processes: Basic notions and properties

II.1 Historical remarks

II.2 Definitions and properties

II.3 L-ensembles

II.4 Related structures: conditional DPPs, structured DPPs and k-DPPs

II.5 The magic properties of DPPs

II.6 The mode problem

Chapter III

Learning setups

III.1 What does learning mean?

III.2 Motivation for learning DPPs

III.3 Direct reconstruction of the kernel

III.4 A learning approach to estimating the kernel

III.5 Learning the quality functions

III.6 Estimating the mixture coefficients of k-DPPs

Chapter IV

Toy examples and experiments

IV.1 Minimal example?

IV.2 Points on the line

IV.3 Points in the square

IV.4 Toy example for quality learning

Chapter V

Summary and conclusion

Chapter A

Generated code

Bibliography

[Affandi et al., 2014] Affandi, R. H., Fox, E., Adams, R., and Taskar, B. (2014). Learning the parameters of determinantal point process kernels. In *International Conference on Machine Learning*, pages 1224–1232.

[Benard and Macchi, 1973] Benard, C. and Macchi, O. (1973). Detection and“emission”processes of quantum particles in a“chaotic state”. *Journal of Mathematical Physics*, 14(2):155–167.

[Borodin, 2009] Borodin, A. (2009). Determinantal point processes. *arXiv preprint arXiv:0911.1153*.

[Boyd and Vandenberghe, 2004] Boyd, S. and Vandenberghe, L. (2004). *Convex optimization*. Cambridge university press.

[Griffin and Tsatsomeros, 2006] Griffin, K. and Tsatsomeros, M. J. (2006). Principal minors, part ii: The principal minor assignment problem. *Linear Algebra and its applications*, 419(1):125–171.

[Higham, 1990] Higham, N. J. (1990). Exploiting fast matrix multiplication within the level 3 blas. *ACM Transactions on Mathematical Software (TOMS)*, 16(4):352–368.

[Kulesza et al., 2012] Kulesza, A., Taskar, B., et al. (2012). Determinantal point processes for machine learning. *Foundations and Trends® in Machine Learning*, 5(2–3):123–286.

[Magen and Zouzias, 2008] Magen, A. and Zouzias, A. (2008). Near optimal dimensionality reductions that preserve volumes. In *Approximation, Randomization and Combinatorial Optimization. Algorithms and Techniques*, pages 523–534. Springer.

[Rising et al., 2015] Rising, J., Kulesza, A., and Taskar, B. (2015). An efficient algorithm for the symmetric principal minor assignment problem. *Linear Algebra and its Applications*, 473:126–144.

[Samuel, 1959] Samuel, A. L. (1959). Some studies in machine learning using the game of checkers. *IBM Journal of research and development*, 3(3):210–229.

[Urschel et al., 2017] Urschel, J., Brunel, V.-E., Moitra, A., and Rigollet, P. (2017). Learning determinantal point processes with moments and cycles. *arXiv preprint arXiv:1703.00539*.