Kernel reconstruction of DPPs

J. Müller

University of Warwick

September 13, 2018

Table of Contents

Table of Contents

Table of Contents

Setting and basic definitions

Setting

Let $\mathcal{Y}$ be a finite set, which we call the *ground set* and $N := |\mathcal{Y}|$ its cardinality. We call the elements of $\mathcal{Y}$ *items* and assume for the sake of easy notation $\mathcal{Y} = \{1, \dots, N\}$ unless specified differently.

Point process

A *point process* on $\mathcal{Y}$ is a random subset of $\mathcal{Y}$, i.e. a random variable with values in the powerset $2^{\mathcal{Y}}$. We identify this random variable with its law $\mathbb{P}$ and thus refer to probability measures $\mathbb{P}$ on $2^{\mathcal{Y}}$ as point processes. Further, $\mathbf{Y}$ will always denote a random subset distributed according to $\mathbb{P}$.

Definition of DPP and repulsive structure

Determinantal point process

We call $\mathbb{P}$ a *determinantal point process*, or in short a *DPP*, if we have

$$\mathbb{P}(A \subseteq \mathbf{Y}) = \det(K_A) \quad \text{for all } A \subseteq \mathcal{Y} \tag{1}$$

where K is a symmetric matrix indexed by the elements in $\mathcal{Y}$ and K_A denotes the submatrix $(K_{ij})_{ij\in A}$ of K indexed by the elements of A. We call K the *marginal kernel* of the DPP. If the marginal kernel K is diagonal, we call $\mathbb{P}$ a *Poisson point process*.

Choosing $A = \{i\}$ and $A = \{i, j\}$ for $i, j \in \mathcal{Y}$ in (1) yields

$$\begin{aligned} \mathbb{P}(i \in \mathbf{Y}) &= K_{ii} \quad \text{and} \\ \mathbb{P}(i, j \in \mathbf{Y}) &= K_{ii}K_{jj} - K_{ij}^2 = \mathbb{P}(i \in \mathbf{Y}) \cdot \mathbb{P}(j \in \mathbf{Y}) - K_{ij}^2. \end{aligned} \tag{2}$$

Properties of the marginal kernel and existence

Positivity

The marginal kernel is always positive semi-definite. Further the complement of a DPP is also a DPP with marginal kernel $I - K$ and hence $0 \leq K \leq I$.

Existence

Let K be a symmetric $N \times N$ matrix. Then K is the marginal kernel of a DPP if and only if $0 \leq K \leq I$.

Table of Contents

Main idea

Setting

Let $\mathcal{Y}$ be a finite set of cardinality N and let $K \in \mathbb{R}^{N\times N}_{\text{sym}}$ satisfy $0 \leq K \leq I$. Let further $\mathbf{Y}_1, \mathbf{Y}_2, \dots$ be independent and distributed according to a DPP with marginal kernel K. The goal is to estimate the kernel K based on the observations $(\mathbf{Y}_n)_{n\in\mathbb{N}}$.

Consider now the empirical measures

$$\hat{\mathbb{P}}_n := \frac{1}{n}\sum_{i=1}^{n} \delta_{\mathbf{Y}_i}$$

and assume that they are determinantal with marginal kernel $\hat{K}_n$. Then $\hat{K}_n$ would be a natural estimate for K since by the SLLN we have

$$\hat{\mathbb{P}}_n \xrightarrow{n\to\infty} \mathbb{P}.$$

The principal minor assignment problem

The principal minor assignment problem (PMAB)

Let $K \in \mathbb{R}^{N \times N}$ be a symmetric matrix. We want to investigate whether K is uniquely specified by its principal minors

$$\Delta_S := \det(K_S) \quad \text{where } S \subseteq \{1, \ldots, N\}$$

and if so how it can be reconstructed from those.

Determinantal equivalence

Two symmetric matrices $A, B \in \mathbb{R}^{N \times N}$ are called *determinantally equivalent* if they have the same principal minors and we write $A \sim B$.

Reconstruction for 3×3 matrix

The diagonal and absolut values of the off diagonal can be obtained by

$$K_{ii} = \Delta_{\{i\}} \quad \text{and}$$
$$K_{ij}^2 = K_{ii}K_{jj} - \Delta_{\{i,j\}}.$$

In order to reconstruct the signs we need to consider the determinant

$$\Delta_{\{1,2,3\}} = K_{11}K_{22}K_{33} + 2K_{12}K_{13}K_{23} - K_{11}K_{23}^2 - K_{22}K_{13}^2 - K_{33}K_{12}^2.$$

Any assignment of the signs that satisfies this, i.e. such that

$$K_{12}K_{13}K_{23} = \frac{1}{2}\left(\Delta_{\{1,2,3\}} + K_{11}K_{23}^2 + K_{22}K_{13}^2 + K_{33}K_{12}^2 - K_{11}K_{22}K_{33}\right)$$

yields a matrix K with the prediscribed principles minors.

Table of Contents

Graph theoretical preliminaries

Let $G = (V, E)$ be a finite graph. We will need the following notions:

(*i*) *Degree:* Number of edges that contains $v \in V$.

(*ii*) *Subgraph:* A graph $\tilde{G} = (\tilde{V}, \tilde{E}) \subseteq (V, E)$.

(*iii*) *Induced graph:* For $S \subseteq V$ the *induced graph* $G(S) = (S, E(S))$ is formed of all edges $e \in E$ of G that are subsets of S.

(*iv*) *Path:* A sequence $v_0 v_1 \cdots v_k$ of vertices such that $\{v_{i-1}, v_i\} \in E$.

(*v*) *Connected graph:* A graph where every two vertices $v, w \in V$ there is a path from v to w.

(*vi*) *Cycle:* A *cycle* C is a connected subgraph such that every vertex has even degree in C.

Graph theoretical preliminaries II

(*vii*) *Cycle space:* Identify a cycle C with $x = x(C) \in \mathbb{F}_2^E$ such that

$$x_e := \begin{cases} 1 & \text{if } e \in C \\ 0 & \text{if } e \notin C. \end{cases}$$

The *cycle space* $\mathcal{C}$ is the span of $\{x(C) \mid C \text{ is a cycle}\}$ in $\mathbb{F}_2^E$.

(*viii*) *Simple cycle:* A cycle C such that every vertex has degree 2 in C.

(*ix*) *Chordless cycle:* A cycle C such that two vertices $v, w \in C$ form an edge in G if and only they form an edge in C.

(*x*) *Cycle sparsity:* The minimal number l such that a basis of the cycle space consisting of chordless simple cycles of length at most l exists which we call *shortest maximal cycle basis* or short *SMCB*.

(*xi*) *Pairings:* A *pairing* P of $S \subseteq V$ is a subset of edges of $G(S)$ such that two different edges of P are disjoint. Vertices contained in the edges of P are denoted by $V(P)$, the set of all pairings by $\mathcal{P}(S)$.

Graph theoretical preliminaries III

Existence of SMCBs

There always exists a basis $\{x(C_1), \dots, x(C_k)\}$ of the cycle space where $C_1, \dots, C_k$ are chordless simple cycles.

Proof.

We proceed in the two following steps:

(*i*) Decompose a cycle into disjoint simple cycles.

(*ii*) Write a simple cycle as the symmetric difference of simple chordless cycles.

Those two steps show that the set of simple chordless cycles generates the cycle space. □

Table of Contents

Table of Contents