引力と斥力を制御可能なべき集合上の分布族

川島貴大 (ZOZO Research), 日野英逸（統数研・理研 AIP）

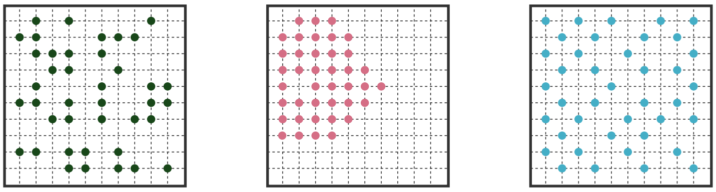
arXiv:2408.01022

背景

台集合 $\mathcal{Y} = \{1, ..., N\}$ からランダムに部分集合をとりだすための確率モデルを考えたい

▷ 実験計画，推薦システム，確率的最適化，……

引力と斥力が重要！



Independent　Positive dependence　Negative dependence

計算機的に扱いやすく，引力と斥力を制御可能な確率モデルはこれまでなかった

▷ これらを実現する分布族 **DKPP** を提案！

提案モデル：DKPP

DKPP（離散カーネル点過程）

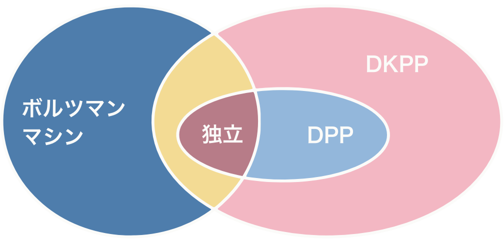
$$P_\varphi(\mathcal{A}; \boldsymbol{L}) = \frac{\mathrm{etr}(\varphi(\boldsymbol{L}[\mathcal{A}]))}{Z_\varphi(\boldsymbol{L})}$$

φ：連続関数

$\boldsymbol{L}$：$N \times N$の正定値カーネル行列

▷ $\boldsymbol{L}$ がアイテム間類似度を，φ が引力・斥力を制御

モデルの包含関係



結合係数がすべて正 or 負のボルツマンマシン，行列式点過程 (DPP) を内包

DKPP の引力・斥力

集合関数 $\log P(\mathcal{A})$ が

- 優モジュラ → 引力
- 劣モジュラ → 斥力

作用素単調増加性：
$\boldsymbol{A} \preccurlyeq \boldsymbol{B} \Rightarrow f(\boldsymbol{A}) \preccurlyeq f(\boldsymbol{B})$

DKPP では，φ の導関数 φ' が

- 作用素単調増加 → 優モジュラ → 引力
- 作用素単調減少 → 劣モジュラ → 斥力

DKPP の計算機上での扱い

モード探索

$\max_{\mathcal{A} \subseteq \mathcal{Y}} \log P_\varphi(\mathcal{A}; \boldsymbol{L})$

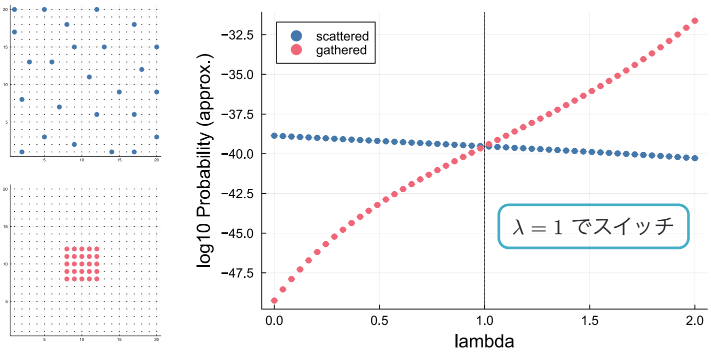
- $\log P_\varphi$ が優モジュラ → 劣モジュラ最小化問題
  ▷ 多項式時間で解ける
- $\log P_\varphi$ が劣モジュラ → 劣モジュラ最大化問題
  ▷ 貪欲法ベースのアルゴリズムで近似解が求まる

周辺確率・条件付き確率

$\mathbb{P}(\mathcal{A}_{\mathrm{in}} \subseteq \mathcal{A} \subseteq \mathcal{A}_{\mathrm{out}})$ や $\mathbb{P}(|\mathcal{A}| = k)$ のような周辺確率は **Rao–Blackwell 化重点サンプリング**で効率的に推定

Box–Cox 変換 $\varphi_\lambda(x)$ の $\lambda \in [0, 2]$ を動かし条件付き確率 $P_{\varphi_\lambda}(\mathcal{A} \mid |\mathcal{A}| = k; \boldsymbol{L})$ を評価

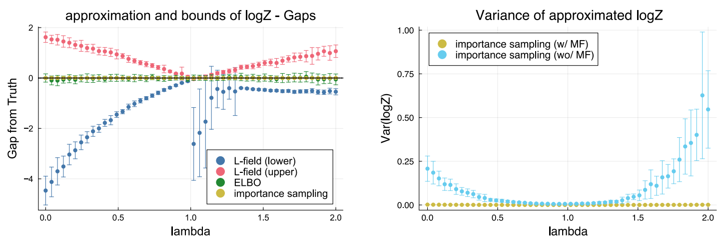
▷ $\lambda \in [0, 1]$ で斥力，$\lambda \in [1, 2]$ で引力



▷ たしかに引力・斥力が制御されている！

正規化定数

$Z_\varphi(\boldsymbol{L})$ は**平均場近似 + 重点サンプリング**で推定



▷ 不偏性 & 低バリアンス性を実現

パラメータの学習

カーネル行列の学習には **Ratio Matching** が有効

Ratio Matching による損失関数

$$J(\boldsymbol{L}) = \frac{1}{M} \sum_{n=1}^{N} \sum_{m=1}^{M} \frac{\mathrm{etr}(\varphi(\boldsymbol{L}[\mathcal{A}_m^{\overline{n}}]))}{\mathrm{etr}(\varphi(\boldsymbol{L}[\mathcal{A}_m])) + \mathrm{etr}(\varphi(\boldsymbol{L}[\mathcal{A}_m^{\overline{n}}]))}$$

▷ 正規化定数の計算を回避 & ミニバッチ化