

# カーネル法によるノンパラメトリックな因果推論

大阪大学大学院 基礎工学研究科 松岡 佑知

大阪大学大学院 基礎工学研究科 濱田 悦生

## 研究背景

Rubin の因果モデルは、ある介入が結果に与える影響をデータから推定する統計学の枠組みである (e.g. [4]). 2 値変数  $T \in \{0, 1\}$  を介入有無のインディケーター変数, **Potential outcomes** を  $Y(1), Y(0) \in \mathbb{R}$ , 共変量を  $X \in \mathbb{R}^d$  で表すとする. データ  $(X_i, Y_i, T_i)_{i=1}^N$  が得られたとき, 平均処置効果 (**Average Treatment Effect, ATE**)  $E[Y(1)] - E[Y(0)]$  を求めたい. 平均処置効果を求めるためには, 強い意味での無視可能性 (**strong ignorability**), すなわち  $\mathbb{P}(Y(1), Y(0)|T, X) = \mathbb{P}(Y(1), Y(0)|X)$  を仮定して, 傾向スコア  $e(x_i) = \mathbb{P}(T_i = 1|X_i = x_i)$  を用いた **IPW** 推定量, **Doubly Robust** 推定量, 傾向スコアマッチングなどが提案されており, 傾向スコアをロジスティック回帰分析などで推定した場合, これらはセミパラメトリックな方法となる. モデルを完全に誤特定した場合 (**misspecified**) これらの推定精度は著しく低下する ([2]). カーネル回帰モデルなどのノンパラメトリック回帰モデルで推定することも可能だが, 次元の呪いを強く受けてしまい多くの場面で有用でない. 本発表では, カーネル平均埋め込みにより, Rubin の因果モデルを再生核ヒルベルト空間においてノンパラメトリックに表現し, 平均処置効果を推定する枠組みを提案し, 数値実験による検証を行う.

## 提案手法

カーネル平均埋め込みは, 確率測度を再生核ヒルベルト空間 (**RKHS**) 上の特徴写像として表現する枠組みである.  $\mathcal{X}$  上の確率変数  $X \sim \mathbb{P}_X$  に対して, 特性的な正定値カーネル  $k: \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$  と対応する **RKHS** を  $\mathcal{H}_X$  としたとき,  $X$  のカーネル平均埋め込みは,  $\mu_X^k := \int_{\mathcal{X}} k(\cdot, x) d\mathbb{P}_X(x) \in \mathcal{H}_X$  により定義される (e.g. [3]). これは, 元の確率測度と 1 対 1 対応しており, 確率測度の **RKHS** での表現となっている. また, 周辺分布だけでなく, 条件付き分布や事後分布の **RKHS** 埋め込みを定義し, 推定することも可能である. これは完全にノンパラメトリックな方法であり, 推定量の収束レートは次元によらない ([3]). 本講演では, カーネル平均による, 傾向スコアや Rubin の因果モデルの **RKHS** 埋め込みの枠組みを提案する. これらはモデルに一切の仮定を置くことなく, 行列演算のみで推定することができる. さらに, 共変量  $X$  が高次元である場合や, 既存のセミパラメトリックモデルが **misspecified** である場合 ([2]) にも, 高い精度で推定が可能であることを数値実験において確認する. 推定の際の正則化パラメータの選択において, 単純な予測精度に基づく **LOOCV** による選択は実は適切ではない. そこで本発表では, 条件付き従属性のカーネル尺度 ([1]) を用いた正則化パラメータ選択規準を提案する.

## 参考文献

- [1] K. Fukumizu, A. Gretton, X. Sun, and B. Schölkopf. Kernel measures of conditional dependence. In *Advances in neural information processing systems 20* : 489-496, 2008.
- [2] J. D. Kang, J. L. Schafer. Demystifying double robustness: A comparison of alternative strategies for estimating a population mean from incomplete data. *Statistical science* : 523-539, 2007.
- [3] L. Song, J. Huang, A. Smola, and K. Fukumizu. Hilbert space embeddings of conditional distributions with applications to dynamical systems. In *Proceedings of the 26th International Conference on Machine Learning (ICML)*, June 2009.
- [4] 岩崎学. 統計的因果推論. 朝倉書店 2015.