Learning Deep Features in Instrumental Variable Regression

Liyuan Xu[1], Yutian Chen[2], Siddarth Srinivasan[3], Nando de Freitas[2], Arnaud Doucet[2], Arthur Gretton[1]

[1]Gatsby Unit [2]DeepMind [3]University of Washington

Abstract

- Consider Instrumental Variable (IV) method to correct additive confounding bias.
- Develop a novel method that can learn a complex structural function using neural networks.
- Observe the superiority of the proposed method in empirical studies.

Preliminaries

Problem Settings

- Consider the additive confounding ε between treatment X and outcome Y.

$$Y = f_{\mathrm{struct}}(X) + \varepsilon, \quad \mathbb{E}[\varepsilon] = 0 \quad \mathbb{E}[\varepsilon|X] \neq 0$$

Regression causes bias since $f_{\mathrm{struct}}(X) \neq \mathbb{E}[Y|X]$

- Instrumental variable Z that satisfies
  - The conditional distribution $P(X|Z)$ is not constant in Z
  - $\mathbb{E}[\varepsilon|Z] = 0$.

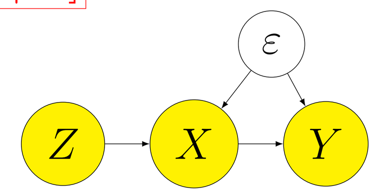

Figure 1: Causal Graph

Instrumental variable can correct the additive confounding bias

Examples

Consider you are evaluating the effect of a new tutoring program

- Treatment X: Participation in the tutoring program
- Outcome Y: Marks made in the exam

→ One's motivation can be the confounder

- Instrumental Z: Proximity of the student's dorm to the tutoring

Two Stage Regression Method

Proposition 1 [Newey & Powell (2013)]

Under regularity conditions, we have

$$f_{\mathrm{struct}} = \arg\min_f \mathbb{E}_{YZ}\left[(Y - \mathbb{E}_{X|Z}\left[f(X)\right])^2\right]$$

Two-stage regression [Newey & Powell (2003); Singh et al. (2019)]

- Model

$$f_{\mathrm{struct}}(x) = \boldsymbol{u}^\top \boldsymbol{\psi}(x) \text{ and } \mathbb{E}_{X|z}[\boldsymbol{\psi}(X)] = \boldsymbol{V}\boldsymbol{\phi}(z),$$

where $\boldsymbol{\psi}, \boldsymbol{\phi}$ are static feature maps and $\boldsymbol{u}, \boldsymbol{V}$ are the parameters.

- (Stage 1 Regression) Learn $\hat{\boldsymbol{V}}$ by minimizing

$$\mathcal{L}_{\mathrm{stage1}}(\boldsymbol{V}) = \mathbb{E}_{X,Z}\left[\|\boldsymbol{\psi}(X) - \boldsymbol{V}\boldsymbol{\phi}(Z)\|^2\right] + \lambda_1 \|\boldsymbol{V}\|^2$$

- (Stage 2 Regression) Learn $\hat{\boldsymbol{u}}$ by minimizing

$$\mathcal{L}_{\mathrm{stage2}}(\boldsymbol{u}) = \mathbb{E}_{Y,Z}[\|Y - \underbrace{\boldsymbol{u}^\top \hat{\boldsymbol{V}}\boldsymbol{\phi}(Z)}_{\simeq \mathbb{E}[f(X)|Z]}\|^2] + \lambda_2 \|\boldsymbol{u}\|^2$$

☺ Has closed-form solution, Consistency Proof, Sample Efficient
☹ Limited flexibility

- Difficult to determine basis functions for images, words, ...

Proposed method learns features adaptively

Proposed Method

Proposed Method: Deep Feature Instrumental Variable (DFIV)

- Model

$$f_{\mathrm{struct}}(x) = \boldsymbol{u}^\top \boldsymbol{\psi}_{\theta_X}(x) \text{ and } \mathbb{E}_{X|z}[\boldsymbol{\psi}_{\theta_X}(X)] = \boldsymbol{V}\boldsymbol{\phi}_{\theta_Z}(z),$$

where $\boldsymbol{\psi}, \boldsymbol{\phi}$ are adaptive feature maps parameterized by θ_X, θ_Z.

- Solve two-stage regression with fixed θ_X, θ_Z

$$\hat{\boldsymbol{V}}_{\theta_X,\theta_Z} = \mathbb{E}\left[\boldsymbol{\psi}_{\theta_X}(X)\boldsymbol{\phi}_{\theta_Z}^\top(Z)\right]\left(\mathbb{E}\left[\boldsymbol{\phi}_{\theta_Z}(Z)\boldsymbol{\phi}_{\theta_X}^\top(Z)\right] + \lambda_1 I\right)^{-1}$$

$$\hat{\boldsymbol{u}}_{\theta_X,\theta_Z} = \left(\mathbb{E}\left[\hat{\boldsymbol{V}}\boldsymbol{\phi}_{\theta_Z}(Z)\boldsymbol{\phi}_{\theta_Z}^\top(Z)\hat{\boldsymbol{V}}^\top\right] + \lambda_2 I\right)^{-1}\mathbb{E}\left[Y\hat{\boldsymbol{V}}\boldsymbol{\phi}_{\theta_Z}(Z)\right]$$

- Update parameter θ_X, θ_Z

$$\theta_Z \leftarrow \theta_Z - \alpha\nabla_{\theta_Z}\mathcal{L}_{\mathrm{stage1}}(\hat{\boldsymbol{V}}_{\theta_X,\theta_Z}), \quad \theta_X \leftarrow \theta_X - \alpha\nabla_{\theta_X}\mathcal{L}_{\mathrm{stage2}}(\hat{\boldsymbol{u}}_{\theta_X,\theta_Z}, \hat{\boldsymbol{V}}_{\theta_X,\theta_Z})$$

where

$$\mathcal{L}_{\mathrm{stage1}}(\hat{\boldsymbol{V}}_{\theta_X,\theta_Z}) = \mathbb{E}_{X,Z}\left[\|\boldsymbol{\psi}_{\theta_X}(X) - \hat{\boldsymbol{V}}_{\theta_X,\theta_Z}\boldsymbol{\phi}_{\theta_Z}(Z)\|^2\right] + \lambda_1\|\hat{\boldsymbol{V}}_{\theta_X,\theta_Z}\|^2$$

$$\mathcal{L}_{\mathrm{stage2}}(\hat{\boldsymbol{u}}_{\theta_X,\theta_Z}, \hat{\boldsymbol{V}}_{\theta_X,\theta_Z}) = \mathbb{E}_{Y,Z}[\|Y - \hat{\boldsymbol{u}}_{\theta_X,\theta_Z}^\top \hat{\boldsymbol{V}}_{\theta_X,\theta_Z}\boldsymbol{\phi}_{\theta_Z}(Z)\|^2] + \lambda_2\|\hat{\boldsymbol{u}}_{\theta_X,\theta_Z}\|^2$$

- Repeat updating $(\hat{\boldsymbol{u}}, \hat{\boldsymbol{V}})$ and (θ_X, θ_Z)

Note:

- $(\hat{\boldsymbol{u}}, \hat{\boldsymbol{V}})$ are functions of (θ_X, θ_Z) and can be backproped.
- Preferable to update θ_Z more frequently than θ_X.

Causal Experiments

Experiment based on dSprite dataset [Mattheyet al., 2017]

- Image dataset generated by four latent parameters {scale, rotation, posX, posY}
- Treatment X is the image generated (with Gaussian noise)
- Outcome Y is

$$Y = \underbrace{\frac{\|AX\|_2^2 - 5000}{1000}}_{\text{Structural function}} + \underbrace{32(\mathtt{posY} - 0.5) + \varepsilon}_{\text{Additive confounding}}, \quad \varepsilon \sim \mathcal{N}(0, 0.5),$$

where A is random matrix.

- Instrumental $Z = (\mathtt{scale}, \mathtt{rotation}, \mathtt{posX})$

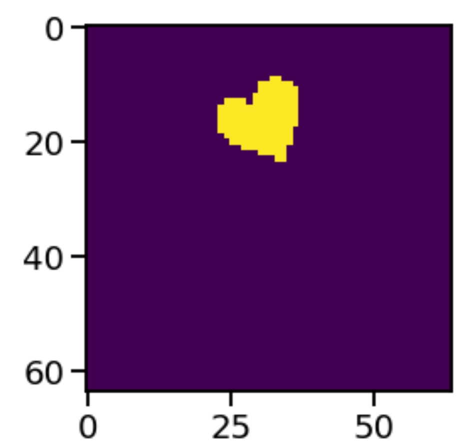

Figure 2: dSprite Image

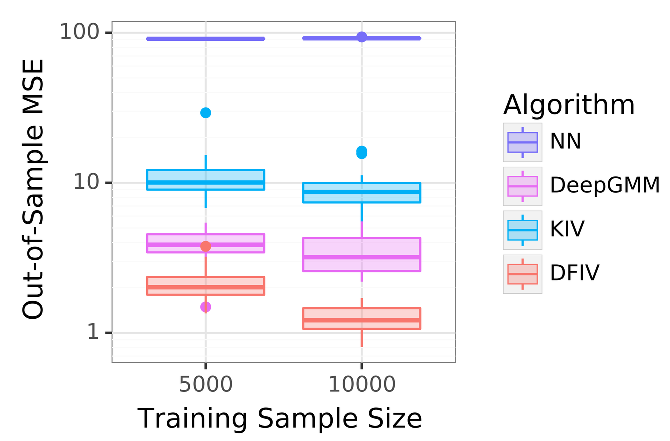

Figure 3: Result of dSprite Experiment

Performs best among nonlinear IV methods

Related Work

Here, we review the nonlinear IV methods we compare in experiments:

- Kernel IV (KIV):
  - Two stage regression where feature maps are in RKHS.
- DeepGMM
  - Use the moment condition of $\mathbb{E}[Y - f_{\mathrm{struct}}(X)|Z] = 0$
  - Estimate f_{struct} by solving a minimax objective with deep nets.

Application to RL

Off-policy Policy Evaluation and IV method

In policy evaluation, we are interested in Q-value:

$$Q^\pi(s, a) = \mathbb{E}\left[\sum_{t=0}^\infty \gamma^t r_t \middle| s_0 = s, a_0 = a\right],$$

which is a minimizer of Bellman loss

$$\mathcal{L}_{\mathrm{Bellman}} = \mathbb{E}_{s,a,r}\left[\left(r + \gamma\mathbb{E}_{s'\sim P(\cdot|s,a),a'\sim\pi(\cdot|s')}\left[Q^\pi(s',a')\right] - Q^\pi(s,a)\right)^2\right].$$

Corresponds to IV method loss $\mathcal{L} = \mathbb{E}_{YZ}\left[(Y - \mathbb{E}\left[f(X)|Z\right])^2\right]$ with

$$X = (s', a', s, a), Y = r, Z = (s, a), f_{\mathrm{struct}}(X) = Q^\pi(s, a) - \gamma Q^\pi(s', a')$$

IV methods can be applied to policy evaluation

Off-policy Policy Evaluation Experiments

- Test on three environments: catch, mountain car, and cartpole
- Replace the action by a random action with probability $p \in [0, 0.5]$
- Evaluate Q-value for trained DQN policies [Mnih et al., 2015]

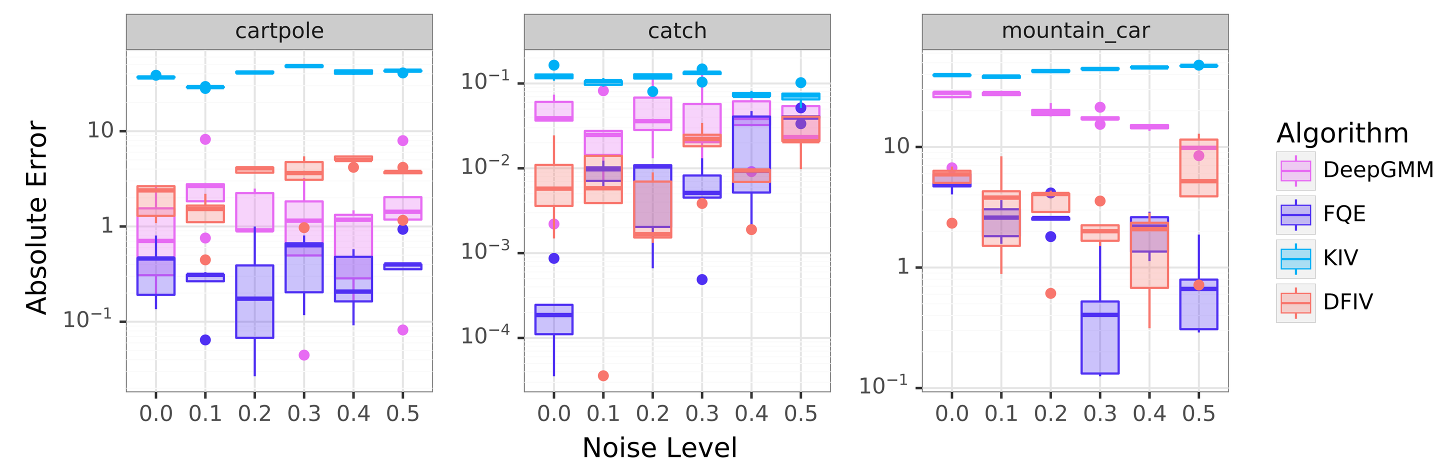

Figure 4: Result of OPE Experiments

Performs best among IV methods, Comparable to the SOTA OPE method

Reference

- A. Bennett, N. Kallus, and T. Schnabel. Deep generalized method of moments for instrumentalvariable analysis. NeurIPS 2019.
- V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves,M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik,I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis. Human-levelcontrol through deep reinforcement learning. Nature 2015.
- L. Matthey, I. Higgins, D. Hassabis, and A. Lerchner. dSprites: Disentanglement testingsprites dataset, 2017. URL:https://github.com/deepmind/dsprites-dataset/.
- W. K. Newey and J. L. Powell. Instrumental variable estimation of nonparametric models. Econometrica 2003.
- R. Singh, M. Sahani, and A. Gretton. Kernel instrumental variable regression. NeurIPS 2019.

GATSBY

DeepMind