

Model selection for estimating CEF derivatives

sasha petrov

May 3, 2022

Fixing objects

From the *all causes* framework, the object of interest:

$$\left\{ \frac{\partial}{\partial x} \phi(X, u) \right\}_{u \in \text{supp } U} \quad (1)$$

The average that's considered to be a more feasible target:

$$\frac{\partial}{\partial x} \mathbb{E} [\phi(X, U) | X = x] \quad (2)$$

Weight derivatives at different values of the domain with the density of those values:

$$\mathbb{E} \left[\frac{\partial}{\partial x} \mathbb{E} [\phi(X, U) | X = x] \right] \quad (3)$$

Can we do model selection in this context – 1

Even if you have a decent predictor for levels $\hat{f}(x)$, it might not work as well for predicting $f(x') - f(x)$. Think of the average treatment effect as a result of moving from x to x' .

- Recall Yitzhaki weights – β_{BLP} estimates a ‘weird’ average of derivatives with weights
- Banerjee (2007): Directly target the average derivative – local OLS is offered as a solution

To me this looks like a b/v issue but with an interesting twist – both supervised and unsupervised ?? How much can we gain in terms of variance by imposing the bias of the sort that β_{BLP} does?

What could do for the proj with reasonable prob: Take a context, simulate (wiemann style) and see if BLP can do better for some dgps?

Can we do model selection in this context – 2

- What is the bias like for β_{BLP} in this setting? Can we obtain the parameter that β_{BLP} estimates as a result of some penalisation? So it would be a penalisation that leads to the bandwidth being the complete domain, right?
- When we do lasso: we chose a particular 'space' with respect to which to penalise the model, right? What can we choose as a 'space' in this case of estimating derivatives?
- It seems weird to impose a minimum distance between the set of weights and the pdf of X , right? Maybe penalise directly the distance between the set of Yitzhaki weights and the true pdf of X ?

Set of weights $\omega : \text{supp } X \rightarrow \mathbf{R}_+$.

$$\min d(\omega, f_X) \tag{4}$$

$$\text{s.t. } d(\omega, \omega_z) \leq \lambda \tag{5}$$