Efficient Influence Function in Parametric Model

Yi-Cheng, Tai

National Yang Ming Chiao Tung University

March 2, 2023

Introduction

We review the efficiency theory based on influence function under parametric setting following Tsiatis (2006)[1].

- Introduce the property of influence function under the special case in M-estimation.
- Characterize the influence function by the properties derived.
- Derive the efficient influence function.
- The connection between efficient influence function and efficient score function.

M-estimation

Consider the parametric setting. Let $\theta \in \Re^p$ be the parameter of interest. Given an estimating equation $m(Z;\theta)$, the empirical solution, $\hat{\theta}_n$, satisfies, under some regularity conditions,

$$
\begin{aligned}
0_p &= n^{-1}\sum_{i=1}^{n} m(Z_i;\hat{\theta}_n) \\
&= n^{-1}\sum_{i=1}^{n} m(Z_i,\theta_0) + n^{-1}\sum_{i=1}^{n} \frac{\partial}{\partial\theta} m(Z_i;\theta)\Big|_{\theta=\theta_0} (\hat{\theta}_n - \theta_0) + R_2 \\
\Rightarrow \hat{\theta}_n - \theta_0 &= -n^{-1}\sum_{i=1}^{n} E\left[\frac{\partial}{\partial\theta} m(Z;\theta_0)\right]^{-1} m(Z_i;\theta_0) + o_P(\frac{1}{\sqrt{n}})
\end{aligned}
$$

It shows that the influence function for M-estimator $\hat{\theta}_n$ is

$$
\phi(z;\theta) = -E\left[\frac{\partial}{\partial\theta} m(z;\theta)\right]^{-1} m(z;\theta)
$$

M-estimation (Cont.)

Since, the estimating equation satisfies,

$$E_\theta[m(Z;\theta)] = 0, \ \forall\theta$$

it can be shown that

$$\begin{aligned}
0 &= \frac{\partial}{\partial\theta}\int m(z;\theta)p(z;\theta)dz \\
&= \int\left[\frac{\partial}{\partial\theta}m(z;\theta)\right]p(z;\theta)dz + \int m(z;\theta)\frac{\partial}{\partial\theta}p(z;\theta)dz \\
&= E_\theta\left[\frac{\partial}{\partial\theta}m(Z;\theta)\right] + \int m(z;\theta)\frac{\frac{\partial}{\partial\theta}p(z;\theta)}{p(z;\theta)}p(z;\theta)dz \\
&= E_\theta\left[\frac{\partial}{\partial\theta}m(Z;\theta)\right] + E_\theta\left[m(Z;\theta)S(Z;\theta)^T\right]
\end{aligned}$$

, where S is the score function.

Properties of Influence Function

The conclusion in the last slide can be written as

$$E_\theta\left[\phi(Z;\theta)S(Z;\theta)^T\right] = 1_p$$

If $\theta = (\beta, \eta)$ with $\beta \in \Re^q$ and $\eta \in \Re^r$, then

$$E_\theta\left[\phi_\beta(Z;\theta)S_\beta(Z;\theta)^T\right] = 1_q$$
$$E_\theta\left[\phi_\beta(Z;\theta)S_\eta(Z;\theta)^T\right] = 0_{q\times r}$$

These two conclusions will be characterized as sufficient conditions to identify the collection of influence functions for β.

Constructing Estimator with given Influence Function

Here is a converse problem:

Is a q-dimensional random function, $\phi(Z)$, satisfying

$$E_\theta\left[\phi(Z)S_\beta(Z;\theta)^T\right] = 1_q$$
$$E_\theta\left[\phi(Z)S_\eta(Z;\theta)^T\right] = 0_{q\times r}$$

an influence function for some estimators of β? (Yes!)

Define the estimating equation

$$m(Z;\beta,\eta) = \phi(Z) - E_\theta\{\phi(Z)\}$$

If η could be "nicely" estimated by $\hat{\eta}_n(\beta)$, given β, we claim that the solution $\hat{\beta}_n$ to

$$\sum_{i=1}^{n} m(Z_i;\beta,\hat{\eta}_n(\beta)) = 0$$

is asymptotically linear with influence function ϕ.

Constructing Estimator with given Influence Function (Cont. 1)

$$
\begin{aligned}
0_q &= n^{-1}\sum_{i=1}^{n} m(Z_i;\hat{\beta}_n,\hat{\eta}_n(\hat{\beta}_n)) \\
&= n^{-1}\sum_{i=1}^{n} m(Z_i;\beta_0,\hat{\eta}_n) + n^{-1}\sum_{i=1}^{n}\frac{\partial}{\partial\beta}m(Z_i;\beta_n^*,\hat{\eta}_n)(\hat{\beta}_n-\beta_0) \\
&= n^{-1}\sum_{i=1}^{n}\left[m(Z_i;\beta_0,\eta_0) + \frac{\partial}{\partial\eta}m(Z_i;\beta_0,\eta_n^*)(\hat{\eta}_n-\eta_0)\right] \\
&\qquad + n^{-1}\sum_{i=1}^{n}\frac{\partial}{\partial\beta}m(Z_i;\beta_n^*,\hat{\eta}_n)(\hat{\beta}_n-\beta_0)
\end{aligned}
$$

Constructing Estimator with Given Influence Function (Cont. 2)

Since, ϕ is mean-zero under the truth, $m(Z;\beta_0,\eta_0)=\phi(Z)$,

$$\begin{aligned} n^{-1}\sum_{i=1}^{n}&\left[m(Z_i;\beta_0,\eta_0)+\frac{\partial}{\partial\eta}m(Z_i;\beta_0,\eta_n^*)(\hat{\eta}_n-\eta_0)\right] \\ &\approx n^{-1}\sum_{i=1}^{n}\phi(Z_i)+E_0\left[\frac{\partial}{\partial\eta}m(Z;\beta_0,\eta_0)\right](\hat{\eta}_n-\eta_0) \end{aligned}$$

It can be shown that $E_0\left[\frac{\partial}{\partial\eta}m(Z;\beta_0,\eta_0)\right]=0$. Hence, this term would converge to the average of $\phi(Z_i)$.

Prove $E_0\left[\frac{\partial}{\partial \eta} m(Z;\beta_0,\eta_0)\right] = 0$

Since, $E_{\beta_0,\eta}[m(Z;\beta_0,\eta)] = 0_q$, $\forall \eta$, it implies that

$$
\begin{aligned}
0_{q\times r} &= \frac{\partial}{\partial \eta} E_0\left[m(Z;\beta_0,\eta_0)\right] \\
&= E_0\left[\frac{\partial}{\partial \eta} m(Z;\beta_0,\eta_0)\right] + E_0\left[m(Z;\beta_0,\eta_0) S_\eta(Z;\beta_0,\eta_0)^T\right] \\
&= E_0\left[\frac{\partial}{\partial \eta} m(Z;\beta_0,\eta_0)\right] + E_0\left[\phi(Z) S_\eta(Z;\beta_0,\eta_0)^T\right] \\
&= E_0\left[\frac{\partial}{\partial \eta} m(Z;\beta_0,\eta_0)\right]
\end{aligned}
$$

, by using the condition on ϕ.

Constructing Estimator with Given Influence Function (Cont. 3)

For the remaining term,

$$n^{-1}\sum_{i=1}^{n}\frac{\partial}{\partial\beta}m(Z_i;\beta_n^*,\hat{\eta}_n)(\hat{\beta}_n-\beta_0)\approx E_0\left[\frac{\partial}{\partial\beta}m(Z;\beta_0,\eta_0)\right](\hat{\beta}_n-\beta_0)$$

By the similar derivation in last slide, it can be shown that

$$E_0\left[\frac{\partial}{\partial\beta}m(Z;\beta_0,\eta_0)\right]=-1_q$$

Consequently, we construct an asymptotic linear estimator with influence function ϕ.

$$\hat{\beta}_n-\beta_0=n^{-1}\sum_{i=1}^{n}\phi(Z_i)+o_P(\frac{1}{\sqrt{n}})$$

Efficient Influence Function

We have shown that the collection of influence functions could be described as random functions satisfies the condition

$$E\left[\phi(Z;\theta)S_\beta(Z;\theta)\right] = 1$$
$$E\left[\phi(Z;\theta)S_\eta(Z;\theta)\right] = 0$$

Here is an alternative description for the collection function:
Given an influence function, ϕ^*, the collection of influence function is $\{\phi^* + \ell : \ell \in \mathcal{T}^\perp\}$, where $\mathcal{T}$ is the tangent space.
Proof:

$$E[\phi^*(Z)S_\beta(Z)] + E[\ell(Z)S_\beta(Z)] = 1$$
$$E[\phi^*(Z)S_\eta(Z)] + E[\ell(Z)S_\eta(Z)] = 0$$

It shows that $\phi^* + \ell$ is an influence function. For the opposite direction, it can be shown that $\phi = \phi^* + (\phi - \phi^*)$, where $\phi - \phi^* \in \mathcal{T}^\perp$.

Efficient Influence Function (Cont.)

How to find the influence function with the smallest variance?

The space of random functions with zero-mean and finite variance equipped with inner-product $E_{\theta_0}[g_1 g_2]$ can be represented as a direct sum of tangent space, $\mathcal{T}$, and its orthogonal complement, $\mathcal{T}^\perp$. Then any influence function, ϕ, can be represented as

$$\phi = \text{proj}(\phi|\mathcal{T}) + \text{proj}(\phi|\mathcal{T}^\perp)$$

Since, $\text{proj}(\phi|\mathcal{T}) = \phi - \text{proj}(\phi|\mathcal{T}^\perp)$, it is an influence function. Any other influence function can be denoted as $\text{proj}(\phi|\mathcal{T}) + \ell$, where $\ell \in \mathcal{T}^\perp$. Consequently,

$$Var[\text{proj}(\phi|\mathcal{T}) + \ell] = Var[\text{proj}(\phi|\mathcal{T})] + Var[\ell] \geq Var[\text{proj}(\phi|\mathcal{T})]$$

$\text{proj}(\phi|\mathcal{T})$ is the efficient influence function.

Efficient Score Function

Let $S_e = S_\beta - \text{proj}(S_\beta|\Lambda)$, where Λ is the nuisance tangent space. Then, S_e is perpendicular to the nuisance tangent space.

$$E[S_e S_\eta^T] = 0$$

It is possible to re-scale S_e to be an influence function.

$$\begin{aligned} E[S_e S_\beta^T] &= E[S_e S_\beta^T] - E[S_e \text{proj}(S_\beta|\Lambda)^T] + E[S_e \text{proj}(S_\beta|\Lambda)^T] \\ &= E[S_e S_e^T] + E[S_e \text{proj}(S_\beta|\Lambda)^T] \end{aligned}$$

Note that $E[S_e \text{proj}(S_\beta|\Lambda)^T] = 0$, since $S_e \perp \Lambda$.

Therefore, $E[S_e S_e^T]^{-1} E[S_e S_\beta^T] = 1$ and $E[S_e S_e^T]^{-1} S_e$ is an influence function. Since it is in the tangent space, it is the efficient influence function.

Tangent Space

Let $\theta = (\beta, \eta) \in \Re^p$, $\beta \in \Re^q$ and $\eta \in \Re^r$. The tangent space is defined as the collection of $BS(Z;\theta_0)$ for all possible matrix $B : 2 \times 2$. Then the elements in tangent space can be decomposed as $B_1 S_\beta(Z;\theta_0) + B_2 S_\eta(Z;\theta_0)$. The collection of $B_2 S_\eta(Z;\theta_0)$ is called nuisance tangent space.

The property of the influence function,

$$E\left[\phi_\beta(Z;\theta) S_\eta(Z;\theta)\right] = 0$$

tells us that the influence function for the parameter of interest is perpendicular to the nuisance tangent space. That is, it is in the orthogonal complement of the nuisance tangent space.

References

A. A. Tsiatis.
Semiparametric theory and missing data.
Springer, 2006.