Nonparametric Instrumental Regression

He Li, ***hl3420***

This paper considers the nonparametric estimation of an instrumental regression problem. Generally, for any function $\varphi \in L_F^2(Z)$, we call it an instrumental regression if,

$$Y = \varphi(Z) + U,\ \mathbb{E}[U|W] = 0, \tag{1}$$

where F is the joint c.d.f. of Y, Z, W. Now we introduce the following two notations:

(i) $T : L_F^2(Z) \to L_F^2(W), \varphi \to T\varphi = \mathbb{E}[\varphi(Z)|W]$.
(ii) $T^* : L_F^2(W) \to L_F^2(Z), \psi \to T^*\psi = \mathbb{E}[\psi(W)|Z]$.

Using these notations, we can now translate equation (1) into,

$$\begin{aligned} \int \varphi(z) \frac{f_{Z,W}(z,w)}{f_W(w)} \, \mathrm{d}z &= \int y \frac{f_{Y,W}(y,w)}{f_W(w)} \, \mathrm{d}y \\ \iff \quad T\varphi = r \end{aligned} \tag{2}$$

where $f(y,z,w)$ is the density function of F w.r.t. the Lebesgue measure and $r(w) = f_{Y,W}(y,w)/f_W(w)$.

The inverse problem (2) can be seen as solving a infinite-dimension linear system to obtain $\varphi(\cdot)$, if we treat the density function $f(y,z,w)$ as known.

Given some mild condition on the joint distribution F, there exists a singular values decomposition of the adjoint Hilbert–Schmidt operators T and T^* such that for orthonormal sequences of $L_F^2(Z)$ $(\varphi_i, i \geq 0)$ and $L_F^2(W)$ $(\psi_i, i \geq 0)$, and real numbers $\lambda_0 = 1 \geq \lambda_1 \geq \ldots$,

$$\begin{aligned} T\varphi_i &= \lambda_i \psi_i,\ i \geq 0 \\ T^*\psi_i &= \lambda_i \varphi_i,\ i \geq 0. \end{aligned}$$

Then the solution φ of the inverse problem (2) is identifiable and the existence of φ is guaranteed by,

$$\varphi = \sum_{i \geq 0} \frac{1}{\lambda_i} \langle r, \psi_i \rangle \varphi_i. \tag{3}$$

However, since we do not know the density f of F, we use some density estimation method to approximate f in practice. Therefore, as λ_i can be arbitrarily small ($\lambda_i \to 0$ as $i \to \infty$), a noisy measurement from r to $r + \delta\psi_i$ leads to a perturbed solution $\varphi + \frac{\delta}{\lambda_i}\varphi_i$, which can be infinitely far from the true solution φ. This problem is said to be ill-posed under this circumstance.

One way to deal with the ill-posed inverse problem is looking for a regularized solution (same as the common linear regression case). A Tikhonov regularized solution is defined as,

$$\varphi^\alpha = (\alpha I + T^*T)^{-1} T^* r = \sum_{i \geq 0} \frac{\lambda_i}{\alpha + \lambda_i^2} \langle r, \psi_i \rangle \varphi_i. \tag{4}$$

or, equivalently,

$$\varphi^\alpha = \arg\min_{\varphi} \left[\|r - T\varphi\|^2 + \alpha \|\varphi\|^2 \right]. \tag{5}$$

Equation (5) can be seen as a functional form of linear regression. It can be shown that

$$\lim_{\alpha \to 0} \|\varphi - \varphi^\alpha\|^2 = \mathcal{O}\left(\alpha^{\beta \wedge 2}\right),$$

where β depends on the degrees of smoothness of the solution as well as the degree of ill-posedness of the inverse problem. And a iterated regularization schemes,

$$\begin{aligned} \varphi^{\alpha}_{(1)} &= (\alpha I + T^*T)^{-1} T^*T\varphi \\ \varphi^{\alpha}_{(k)} &= (\alpha I + T^*T)^{-1} \left[T^*T\varphi + \alpha\varphi^{\alpha}_{(k-1)}\right] \ , \\ \cdots \end{aligned} \tag{6}$$

admits a solution of

$$\varphi^{\alpha}_{(k)} = \sum_{i\geq 0} \frac{\left(\lambda_i^2 + \alpha\right)^k - \alpha^k}{\lambda_i \left(\alpha + \lambda_i^2\right)^k} \left\langle \varphi, \varphi_i \right\rangle \varphi_i, \tag{7}$$

and satisfies

$$\lim_{\alpha\to 0} \|\varphi - \varphi^{\alpha}_{(k)}\|^2 = \mathcal{O}\left(\alpha^{\beta\wedge 2k}\right),$$

which can take advantage of a degree of smoothness β for φ larger than 2.

Now we turn to the statistical estimation problem. Let $h \equiv h_N \to 0$ denote a bandwidth [8] and let $K_h(\cdot,\cdot)$ denote a univariate generalized kernel function with the properties $K_h(u,t) = 0$ if $u > t$ or $u < t-1$; for all $t \in [0,1]$,

$$h^{-(j+1)} \int_{t-1}^{t} u^j K_h(u,t)du = \begin{cases} 1, & \text{if } j = 0 \\ 0, & \text{if } 1 \leq j \leq l-1 \end{cases}$$

We call $K_h(\cdot,\cdot)$ a univariate generalized kernel function of order l. And multivariate generalized kernel function of order l can be given by products of univariate generalized kernel functions of order l. Given two multivariate generalized kernel functions $K_{Z,h}$ and $K_{W,h}$, with dimension p and q respectively, we can estimate the density function of F as,

$$\begin{aligned} \widehat{f_{z,W}}(z,w) &= \frac{1}{Nh^{p+q}} \sum_{n=1}^{N} K_{Z,h}\left(z - z_n, z\right) K_{W,h}\left(w - w_n, w\right), \\ \widehat{f}_W(w) &= \frac{1}{Nh^q} \sum_{n=1}^{N} K_{W,h}\left(w - w_n, w\right), \\ \widehat{f}_Z(z) &= \frac{1}{Nh^p} \sum_{n=1}^{N} K_{Z,h}\left(z - z_n, z\right). \end{aligned}$$

This can give us the estimators of T, T^* as well as r,

$$\begin{aligned} (\hat{T}\varphi)(w) &= \int \varphi(z) \frac{\widehat{f}_{Z,W}(z,w)}{\widehat{f}_W(w)} \mathrm{d}z, \\ \left(\hat{T}^*\psi\right)(z) &= \int \psi(w) \frac{\widehat{f}_{Z,W}(z,w)}{\widehat{f}_Z(z)} \mathrm{d}w, \\ \hat{r}(w) &= \frac{\sum_{n=1}^{N} y_n K_{W,h}\left(w - w_n, w\right)}{\sum_{n=1}^{N} K_{W,h}\left(w - w_n, w\right)}. \end{aligned}$$

Therefore, the final estimation of our solution to the inverse problem (2) is then given by,

$$\hat{\varphi}^{\alpha_N} = \left(\alpha_N I + \hat{T}^*\hat{T}\right)^{-1} \hat{T}^*\hat{r}, \tag{8}$$

where the regularization parameter α_N is a positive number depending on N.

It can be shown that our estimator $\hat{\varphi}^{\alpha_N}$ also converges to the true solution φ with the cost of the nonparametric estimation of conditional expectations negligible. For some bandwith h_N and α_N,

$$\|\hat{\varphi}^{\alpha_N} - \varphi\|^2 = O_P\left[N^{-(\beta\wedge 2)/((\beta\wedge 2)+2)}\right]. \tag{9}$$

To conclude, this paper considers the nonparametric estimation of a regression function. The core of estimation is using a kernel-based nonparametric density estimation method. Under general mild conditions, the consistency of the nonparametric instrumental variables estimator is ensured. And the convergence rate of this estimator is strongly related to the degree of ill-posedness of the inverse problem (2).

References

Darolles, S., Y. Fan, J.-P. Florens, and E. Renault (2011). Nonparametric instrumental regression. *Econometrica 79*(5), 1541–1565.