

Modelo Matemático

Anzony Quispe

En esta entrega del seminario les comentaré acerca del modelo matemático que seguiré en la tesis. Este modelo está basado en el trabajo acerca de Automatic Debiased Machine Learning (Chernozhukov et al., 2022). Este trabajo tiene por objetivo la estimación del parámetro θ_0 , en un contexto de alta dimensionalidad.

$$\theta_0 = E(m(W; g_0))$$

En esta ecuación tenemos que $W := (Y, Z)$ y $g_0(Z) := E[Y|Z]$. Tenemos que Y es el outcome de interés y $Z = (T, X)$ contiene a las covariables y el tratamiento de interés. Como lo expresamos en el modelo descriptivo, no podemos obtener la estimación directamente de θ_0 a través de un modelo ML debido al sesgo generado por regularización y overfitting. Para ello, necesitamos de procesos de debiasing como el explicado por Chernozhukov et al. (2018). Sin embargo, en trabajos más recientes (Chernozhukov et al. (2021)) esta primera ecuación puede ser expresada como la multiplicación de dos términos bajo el teorema de representación de Riesz.

$$\theta_0 = E(m(W; g_0)) = E(\alpha_0(Z)g_0(Z))$$

Donde $\alpha_0(Z)$ existe si y solo si $E(m(W; g_0))$ es una función lineal continua de g_0 . Se denomina a $\alpha_0(Z)$ como Riez representer (RR). Luego, este término es útil para generar términos que quiten el sesgo de la función de momento $m(W; g_0)$. Para el caso mostrado arriba, se muestra que podemos estimar el parámetro de la siguiente ecuación.

$$\hat{\theta} = E_n[m(W; \hat{g}) + \hat{\alpha}(Z)(Y - \hat{g}(Z))]$$

Entonces, cuando el producto de los ratios de convergencia de las nuisance functions es mayor que $n^{0.5}$, el estimador converge en distribución. Finalmente, si g_0 está correctamente estimado, el requerimiento del ratio de $\hat{\alpha}$ es menos estricto.

Estos términos permiten quitar el sesgo de la estimación de $\hat{\theta}$. La literatura ha ido construyendo estos términos de forma independiente para diferentes casos. Sin embargo, Chernozhukov et al. (2021) muestra que este término puede ser estimado como una función de minimización que evita la derivación de una forma específica del término, especificación del término.

$$\alpha_0 = \arg \min_{\alpha} E[(\alpha(Z) - \alpha_0(Z))^2]$$

$$\alpha_0 = \arg \min_{\alpha} E[\alpha(Z)^2 - 2m(W; \alpha)]$$

$$\hat{\alpha} = \arg \min_{\alpha \in A_n} E[\alpha(Z)^2 - 2m(W; \alpha)]$$

Entonces, siguiendo el teorema representación de Riesz Podemos obtener el término debiasing sin derivar su forma analítica, y por el proceso de minimización Podemos tener un intercambio entre sesgo y varianza.

En este trabajo de investigación se buscará implementar esta metodología que aún no está disponible para Python ni R. Luego, se aplicará para un ejemplo empírico y compararlo con otras metodologías (rlassoATE, DoubleMLIRM).

Referencias

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., & Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters.

Chernozhukov, V., Newey, W. K., Quintas-Martinez, V., & Syrgkanis, V. (2021). Automatic debiased machine learning via neural nets for generalized linear regression. arXiv preprint arXiv:2104.14737.

Chernozhukov, V., Newey, W. K., & Singh, R. (2022). Automatic debiased machine learning of causal and structural effects. *Econometrica*, 90(3), 967-1027.

Chernozhukov, V., Newey, W., Quintas-Martínez, V. M., & Syrgkanis, V. (2022, June). Riesznet and forestriesz: Automatic debiased machine learning with neural nets and random forests. In International Conference on Machine Learning (pp. 3901-3914). PMLR.