Learning Optimal Treatment Regimes with Survival Data under Imperfect Compliance: An Instrumental Variable Approach

Zhejiang University, Beijing Technology and Business University,

East China Normal University, Stellenbosch University and Africa Health Research Institute

Supplementary Material

The online Supplementary Material contains detailed proofs of Theorems 1-2 and Proposition 1 as well as the theoretical results of Fisher consistency, excess risk bound and universal consistency of the estimated treatment regimes.

S1 Proofs

S1.1 Proof of Theorem 1

By Cui and Tchetgen Tchetgen (2021), we have

$$\arg\max_d \mathbb{E}[T_{d(X)}] = \arg\max_d \mathbb{E}\left[\frac{ZAT\mathbb{1}\{A = d(X)\}}{\delta(X)\Pr(Z|X)}\right].$$

Furthermore, we have

$$\begin{aligned}
&\mathbb{E}\left[\frac{\Delta Y}{S_C(Y|X,A,Z)}\frac{ZA\mathbb{1}\{A=d(X)\}}{\delta(X)\Pr(Z|X)}\right]\\
=&\mathbb{E}\left[\mathbb{E}\left[\frac{\Delta Y}{S_C(Y|X,A,Z)}|Z,A,X\right]\frac{ZA\mathbb{1}\{A=d(X)\}}{\delta(X)\Pr(Z|X)}\right]\\
=&\mathbb{E}\left[\mathbb{E}[T|Z,A,X]\frac{ZA\mathbb{1}\{A=d(X)\}}{\delta(X)\Pr(Z|X)}\right]\\
=&\mathbb{E}\left[\mathbb{E}\left[\frac{TZA\mathbb{1}\{A=d(X)\}}{\delta(X)\Pr(Z|X)}|Z,A,X\right]\right]\\
=&\mathbb{E}\left[\frac{TZA\mathbb{1}\{A=d(X)\}}{\delta(X)\Pr(Z|X)}\right].
\end{aligned}$$

Therefore, we have

$$\arg\max_{d\in\mathcal{D}}\mathbb{E}[T_{d(X)}]=\arg\max_{d\in\mathcal{D}}\mathbb{E}\left[\frac{ZA\Delta Y\,\mathbb{1}\{A=d(X)\}}{S_C(Y|X,A,Z)\delta(X)\Pr(Z|X)}\right].$$

S1.2 Proof of Theorem 2

Note that we have

$$\begin{aligned}
&\mathbb{E}\left[Q(X,\,Z,\,A,\,Y,\,\Delta,\,S_C,\,S_T)\frac{ZA\mathbb{1}\{A=d(X)\}}{\delta(X)\Pr(Z|X)}\right]\\
=&\mathbb{E}\left[\mathbb{E}\left[Q(X,\,Z,\,A,\,Y,\,\Delta,\,S_C,\,S_T)|Z,A,X\right]\frac{ZA\mathbb{1}\{A=d(X)\}}{\delta(X)\Pr(Z|X)}\right]\\
=&\mathbb{E}\left[\mathbb{E}[T|Z,A,X]\frac{ZA\mathbb{1}\{A=d(X)\}}{\delta(X)\Pr(Z|X)}\right]\\
=&\mathbb{E}\left[\mathbb{E}\left[\frac{TZA\mathbb{1}\{A=d(X)\}}{\delta(X)\Pr(Z|X)}|Z,A,X\right]\right]\\
=&\mathbb{E}\left[\frac{TZA\mathbb{1}\{A=d(X)\}}{\delta(X)\Pr(Z|X)}\right],
\end{aligned}$$

where the second equality holds because

$$E[Q(X,\, Z,\, A,\, Y,\, \Delta,\, S_C,\, S_T)|X, A, Z] = E[T|X, A, Z],$$

following the proof as shown in Van der Laan and Robins (2003); Rubin and van der Laan (2012). Therefore, we have

$$\arg\max_{d\in\mathcal{D}} \mathbb{E}[T_{d(X)}] = \arg\max_{d\in\mathcal{D}} \mathbb{E}\left[Q(X, A, Z, \Delta, Y, S_C, S_T)\frac{ZA\mathbb{1}\{A = d(X)\}}{\delta(X)\Pr(Z|X)}\right].$$

S1.3 Proof of Proposition 1

Under Assumptions 2-9, by Equation (1), $\tilde{d}$ is the regime which maximizes $\mathbb{E}[T_{d(X)}]$ over the policy class $\mathcal{D}$. Then $\mathbb{E}[T_{\tilde{d}(X)}] \geq \mathbb{E}[T_{d^{ITT}(X)}]$ follows by the definition of d^{ITT}.

S2 Asymptotic Properties of The Proposed OTR

In this section, we provide asymptotic properties of estimated treatment regime with $W = W^{IW}$. A similar result holds for $W = W^{DR}$.

We define the following risk

$$R(g) \triangleq \mathbb{E}[W\mathbb{1}\{A \neq \text{sign}(g(X))\}],$$

The optimal decision function associated with the OTR $\bar{d}$ is defined as $g^* \triangleq \arg\min_{g\in\mathcal{G}} R(g)$ and corresponding Bayes risk is $R^* \triangleq R(g^*)$, where $\mathcal{G}$

is the class of all measurable functions. We also define the ϕ-risk

$$R_\phi(g) \triangleq \mathbb{E}[|W|\phi(\text{sign}(W)Ag(X))],$$

where ϕ is the hinge loss function. The minimal ϕ-risk is $R_\phi^* \triangleq \inf_{g\in\mathcal{G}} R_\phi(g)$, and $g_\phi^* \triangleq \arg\min_{g\in\mathcal{G}} R_\phi(g)$. The following Proposition 2 on Fisher consistency and Proposition 3 on upper bound of the excess risk under 0-1 loss follows from Theorems 2.1-2.2 of Zhou and Kosorok (2017).

Proposition 2. Under Assumptions 2-9, $R^* = R(g_\phi^*)$.

Proposition 3. Under Assumptions 2-9, for any measurable decision function g, we have

$$R(g) - R^* \leq R_\phi(g) - R_\phi^*.$$

It is worth noting that Proposition 3 implies that the regret of the regime associated with the decision function g can be bounded by the corresponding regret under the hinge loss. This excess bound also serves as an intermediate step for the universal consistency of the estimated treatment regime given below.

Before establishing the consistency of the estimated treatment regime with a universal kernel (e.g., Gaussian kernel), we introduce the concept of universal kernels (Steinwart and Christmann, 2008). A continuous kernel K on a compact metric space $\mathcal{X}$ is called universal if its associated reproducing

kernel Hilbert space (RKHS) $\mathcal{H}_K$ is dense in $C(\mathcal{X})$, where $C(\mathcal{X})$ is the space of all continuous functions on the compact metric space $\mathcal{X}$ endowed with the usual supremum norm. Let K be a universal kernel, and $\mathcal{H}_K$ be the associated RKHS. We assume the following assumptions about λ_n, g_ϕ^*, and the weight W.

Assumption 10. Consider a sequence of tuning parameters $\lambda_n \to 0$ and $n\lambda_n \to \infty$ as $n \to \infty$. Suppose that g_ϕ^* is measurable and bounded, $|g_\phi^*| \leq M_g$, and $|\sqrt{\lambda}_n b_n| \leq M_b$ almost surely for some constants M_g and M_b.

Assumption 11. $M_\delta < |\delta(X)|$, $M_z < \Pr(Z = 1|X) < 1 - M_z$ for some $0 < M_\delta < 1$, $0 < M_z < 1/2$ almost surely.

Assumption 12.

$$\sup_{x\in\mathcal{X}} |\hat{\delta}(x) - \delta(x)| \xrightarrow{p} 0, \quad \sup_{x\in\mathcal{X}} |\hat{\Pr}(Z = 1|x) - \Pr(Z = 1|x)| \xrightarrow{p} 0,$$

$$\sup_{t\leq h, x\in\mathcal{X}} |\hat{S}_C(t|x) - S_C(t|x)| \xrightarrow{p} 0 \ \text{ as } \ n \to \infty.$$

Then, we have the following result.

Proposition 4. Under Assumptions 2-12, we have

$$\lim_{n\to\infty} R(\hat{g}) = R^*,$$

where $\hat{g} = \hat{h} + \hat{b}$ is the estimated decision function from

$$\min_{g=h+b\in\mathcal{H}_K+\{1\}} \frac{1}{n}\sum_{i=1}^{n} |\hat{w}_i|\phi(\text{sign}(\hat{w}_i)a_i g(x_i)) + \frac{\lambda}{2}||h||_K^2.$$

The proof of the above theorems is akin to Zhao et al. (2012) and Zhou and Kosorok (2017), so we omit it here.

Bibliography

Cui, Y. and E. Tchetgen Tchetgen (2021). A semiparametric instrumental variable approach to optimal treatment regimes under endogeneity. *Journal of the American Statistical Association 116*(533), 162–173.

Rubin, D. B. and M. J. van der Laan (2012). Statistical issues and limitations in personalized medicine research with clinical trials. *The International Journal of Biostatistics 8*(1), 18.

Steinwart, I. and A. Christmann (2008). *Support vector machines*. Springer Science & Business Media.

Van der Laan, M. J. and J. M. Robins (2003). *Unified methods for censored longitudinal data and causality*, Volume 5. Springer.

Zhao, Y., D. Zeng, A. J. Rush, and M. R. Kosorok (2012). Estimating individualized treatment rules using outcome weighted learning. *Journal of the American Statistical Association 107*(499), 1106–1118.

Zhou, X. and M. R. Kosorok (2017). Augmented outcome-weighted learning for optimal treatment regimes. *arXiv preprint arXiv:1711.10654*.