The Supervised Learning Approach To Estimating Heterogeneous Causal Regime Effects

Thai T. Pham

Stanford Graduate School of Business
thaipham@stanford.edu

May, 2016

Observations

- **Many sequential treatment settings**: patients make adjustments in medications in multiple periods; students decide whether to follow an educational honors program over multiple years; in labor market, the unemployed might participate in a set of programs (job search, subsidized job, training) sequentially.
- **Heterogeneity in treatment sequence reactions**: medication effects can be heterogeneous across patients and across time; the same for educational program and labor market.
- **Hard to set up sequential randomized experiments in reality.**

Contributions

Develop a nonparametric framework using supervised learning to estimate heterogeneous treatment regime effects from observational (or experimental) data.

Treatment Regime: a set of functions of characteristics and intermediate outcomes.

- Propose using supervised learning approach (deep learning), which gives good estimation accuracy and which is robust to model misspecification.
- Propose matching based testing method for the estimation of heterogeneous treatment regime effects.
- Propose matching based kernel estimator for variance of heterogeneous treatment regime effects (*time allows*).

Contributions (cont'd)

In this paper, we

- Focus on dynamic setting with multiple treatments applied sequentially (in contrast to a single treatment).
- Focus on the heterogeneous (in contrast to average) effect of a sequence of treatments, i.e. a treatment regime.
- Focus on observational data (in contrast to experimental data).

Setup - Motivational Dataset

- The North Carolina Honors Program Dataset.
- There are $24,112$ observations in total.
- $X_0 = [Y_0, d_1, d_2, d_3]$, where Y_0 is the Math test score at the end of 8th grade and d_1, d_2, d_3 are census-data dummy variables.
- $W_0, W_1 \in \{0, 1\}$ are treatment variables.
- Y_1: end of 9th grade Math test score.
- Y_2: end of 10th grade Math test score (object of interest).
- Y_0, Y_1, Y_2 are pre-scaled to have zero mean and unit variance.

Setup - Model

- **End of eighth grade:**
  - Students' initial information X_0, which includes Math test score Y_0 and other personal information (d_1, d_2, d_3), is observed.
  - Decide to follow honors ($W_0 = 1$) or standard ($W_0 = 0$) program.
- **End of ninth grade:**
  - X_0, W_0, and Math test score Y_1 are observed.
  - Decide to switch or stay in current program ($W_1 = 1$ or 0).
- **End of tenth grade:**
  - X_0, W_0, Y_1, W_1, and Math test score Y_2 are observed.
- Object of interest: Y_2 (It could be any functions of X_0, Y_1, Y_2).

Potential Outcome (PO) Framework

- Treatment regime $d = (d_0, d_1)$ has

$$d_0 : X_0 \to W_0 \in \{0,1\} \text{ and } d_1 : X_0 \times W_0 \times Y_1 \to W_1 \in \{0,1\}.$$

- Potential Outcome $Y_1(W_0) = Y_1$. Also, the observed outcome

$$Y_1 = W_0 \cdot Y_1(1) + (1 - W_0) \cdot Y_1(0).$$

- Similarly, $Y_2^d = Y_2$ if the subject follows regime d. We also write $Y_2 = Y_2^d = Y_2(W_0, W_1)$ when d_0 maps to W_0 and d_1 maps to W_1. We have

$$Y_2 = W_0 W_1 \cdot Y_2(1,1) + W_0(1 - W_1) \cdot Y_2(1,0) + \\ (1 - W_0) W_1 \cdot Y_2(0,1) + (1 - W_0)(1 - W_1) \cdot Y_2(0,0).$$

Types of Treatment Regime

- **Static Treatment Regime**: subjects specify (or are specified) the whole treatment plan based only on the initial covariates (X_0).

  So $d : X_0 \to (W_0, W_1) \in \{0, 1\}^2$.

- **Dynamic Treatment Regime**: subjects choose (or are assigned) the initial treatment based on the initial covariates (X_0); then subsequently choose (or are assigned) the next treatment based on the initial covariates (X_0), the first period treatment (W_0), and the intermediate outcome (Y_1); and so on.

  This is our original setup.

Potential Outcome (PO) Framework (Cont'd)

Objective: Estimate $\mathbb{E}\left[Y_2^d - Y_2^{d'}\right]$ for individuals (or average), and derive heterogeneous optimal regime

$$d^*(C) = \arg\max_d \mathbb{E}\left[Y_2^d \,\middle|\, C\right] \text{ for individual covariates } C.$$

Difficulties:

- Fundamental Problem of Causal Inference: for each subject, we never observe both Y_2^d and $Y_2^{d'}$.
- Selection Bias: students following d may fundamentally be different from those following d' (e.g., *students with good test scores choose the honors program in each period*).

Identification Result - Static Treatment Regime

Theorem (Identification Result - STR)

Let $d_0 = d_0(X_0)$ and $d_1 = d_1(X_0)$. Then (with Assumptions)

$$\mathbb{E}\left[\frac{Y_2 \cdot \mathbb{1}\{W_0 = d_0\} \cdot \mathbb{1}\{W_1 = d_1\}}{\mathbb{P}(W_0 = d_0 | X_0) \cdot \mathbb{P}(W_1 = d_1 | X_0)} \middle| X_0\right] = \mathbb{E}\left[Y_2^d \middle| X_0\right].$$

Corollary:

$$\mathbb{E}\Bigg[\overbrace{Y_2 \cdot \left[\frac{W_0 W_1}{e_0 e_1} - \frac{(1-W_0)(1-W_1)}{(1-e_0)(1-e_1)}\right]}^{\text{observed/estimable, Transformed Outcome}} \Bigg| X_0\Bigg] = \mathbb{E}\Big[\overbrace{Y_2(1,1) - Y_2(0,0)}^{\text{unobserved, PO}} \Big| X_0\Big].$$

Here, $e_0 = \mathbb{P}(W_0 = 1 | X_0)$ and $e_1 = \mathbb{P}(W_1 = 1 | X_0)$.

Matching Based Testing Method | STR Estimation

Identification Result - Dynamic Treatment Regime

Theorem (Identification Result - DTR)

Let $d_0 = d_0(X_0)$ and $d_1 = d_1(X_0, X_1, Y_1, W_0)$. Then (with Assumptions)

In period $T = 1$:

$$\mathbb{E}\Bigg[\underbrace{Y_2 \cdot \frac{\mathbf{1}\{W_1 = d_1\}}{\mathbb{P}(W_1 = d_1 | X_0, X_1, Y_1, W_0)}}_{\text{observed/estimable}} \Bigg| X_0, X_1, Y_1, W_0\Bigg]$$

$$= \mathbb{E}\big[\underbrace{Y_2^{d_1}}_{PO} \big| X_0, X_1, Y_1, W_0\big].$$

In period $T = 0$:

$$\mathbb{E}\left[\frac{Y_2 \cdot \mathbf{1}\{W_1 = d_1\} \cdot \mathbf{1}\{W_0 = d_0\}}{\mathbb{P}(W_1 = d_1 | X_0, X_1, Y_1, W_0) \cdot \mathbb{P}(W_0 = d_0 | X_0)} \middle| X_0\right] = \mathbb{E}\big[Y_2^d \big| X_0\big].$$

▸ DTR Estimation

Challenges In Traditional Approach

Goal: Specify a relation b/w transformed outcome T and covariates C.

Econometric approaches assume $T = h(C;\beta) + \epsilon$ for a fixed (linear) function $h(\cdot)$ and $\mathbb{E}[\epsilon|C] = 0$, and estimate β by minimizing

$$||\mathbf{T} - h(\mathbf{C};\beta)||^2.$$

Problem: Linear models need not give good estimates in general.

Machine Learning Approach

- Machine learning methods generally give much more accurate estimates than traditional econometric models.
- Empirical comparisons of different machine learning methods with linear regressions:
  - Caruana and Niculescu-Mizil (2006)[1]
  - Morton, Marzban, Giannoulis, Patel, Aparasu, and Kakadiaris (2014)[2]

---

[1] Caruana, R. and A. Niculescu-Mizil, (2006), "An Empirical Comparison of Supervised Learning Algorithms," *Proceedings of the 23rd International Conference on Machine Learning*, Pittsburgh, PA.

[2] Morton, A., E. Marzban, G. Giannoulis, A. Patel, R. Aparasu, and I. A. Kakadiaris, (2014), "A Comparison of Supervised Machine Learning Techniques for Predicting Short-Term In-Hospital Length of Stay Among Diabetic Patients," *13th International Conference on Machine Learning and Applications*.

Machine Learning Approach (Cont'd)

Goal: Specify a relation b/w transformed outcome T and covariates C.

- Machine learning (ML) methods allow $h(\cdot)$ to vary in terms of complexity, and estimate β by minimizing

  $$||\mathbf{T} - h(\mathbf{C}; \beta)||^2 + \lambda g(\beta)$$ where g penalizes complex models.

- Data set = (Training, Validation, Test). Use Training set (with CV) to choose the optimal $h(\cdot)$ in terms of complexity, validation set to choose the optimal 'hyperparameter' λ, and test set to evaluate the performance.

- RMSE is the comparison criterion.

Hence, ML approach is flexible and performance oriented.

Estimating Model

- **Propensity Score Estimation:** Use logistic regression or other ML techniques such as Random Forest, Gradient Boosting, etc.
- **Full Model Estimation:** (Though many ML techniques would work) We use a deep learning method in machine learning literature called "Multilayer Perceptron."
  - It possesses the *universal approximation property*: it can approximate any continuous function on any compact subset of $\mathbb{R}^n$.

Multilayer Perceptron (MLP)

- Assume we want to estimate $T = h(C;\beta) + \epsilon$. MLP (with one hidden layer) considers

$$h(C;\beta) = \sum_{j=1}^{K} \alpha_j \sigma(\gamma_j^T C + \theta_j) \text{ and } \beta = \left(K, (\alpha_j, \gamma_j, \theta_j)_{j=1}^{K}\right),$$

where σ is a sigmoid function such as $\sigma(x) = 1/(1 + exp(-x))$.

- Empirically, MLP (and deep learning in general) is shown to work very well. (Lecun et al.[3], Mnih et al.[4])

---

[3] LeCun, Y., Y. Bengio, and G. Hinton, (2015), "Deep Learning," *Nature* 521, 436-444 (28 May).

[4] Mnih, V., K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis, (2015), "Human-level control through deep reinforcement learning," *Nature* 518, 529-533 (26 February).

Matching Based Testing Method

The identification results relate unobserved difference of potential outcomes Z to observed (or estimable) transformed outcome T:

$$\mathbb{E}[T|C] = \mathbb{E}[Z|C].$$

For example in ▸ STR: $Z = Y_2^d - Y_2^{d'}$, $C = X_0$.

- Randomly draw M units with treatment regime d. Denote by $x_0^{d,m}$'s and $y_2^{d,m}$'s the covariates and corresponding outcomes.
- For each m, determine $x_0^{d',m} = \arg\min_{x_0^i \,|\, \text{regime} = d'} ||x_0^i - x_0^{d,m}||^2$.
- Let $\widetilde{\tau}_m = y_2^{d,m} - y_2^{d',m}$. Here, $\widetilde{\tau}$ is a proxy for the unobserved Z.
- Let $\widehat{\tau}$ be the estimator which fits x_0 to T.
- Define $\widehat{\tau}_m = \frac{1}{2}(\widehat{\tau}(x_0^{d,m}) + \widehat{\tau}(x_0^{d',m}))$.
- Define the **matching loss**$_M$: $\sqrt{\frac{1}{M}\sum_{m=1}^{M}(\widetilde{\tau}_m - \widehat{\tau}_m)^2}$.

Simulation Setup

Test the ability of our method in adapting to heterogeneity in the treatment regime effect.

- $50,000$ obs for training; $5,000$ obs for validation; $5,000$ for testing.
- $X_0 \sim U([0,1]^{10})$; $W_0 \in \{0,1\}$; $Y_1 \in \mathbb{R}$ with standard normal noise; $W_1 \in \{0,1\}$; $Y_2 \in \mathbb{R}$ with standard normal noise.
- $e_0(X_0) = e_1(X_0) = e_1(X_0, W_0, Y_1) = 0.5$.
- $\tau_1(X_0) = \mathbb{E}\left[Y_1^{W_0=1} - Y_1^{W_0=0} \,\middle|\, X_0 \right] = \xi(X_0[1])\xi(X_0[2])$; and

$$\tau_2(X_0, W_0, Y_1) = \mathbb{E}\left[Y_2^{W_1=1} - Y_2^{W_1=0} \,\middle|\, X_0, W_0, Y_1 \right] = \rho(Y_1)\rho(W_0)\xi(X_0[1])$$

where

$$\xi(x) = \frac{2}{1 + e^{-12(x-1/2)}}; \rho(x) = 1 + \frac{1}{1 + e^{-20(x-1/3)}}.$$

Simulation Results

Table: Performance In Terms of Root Mean Squared Error (RMSE)

Method	Linear Regression (LR)	Multilayer Perceptron (MLP)
STR	1.75	1.66
DTR: T = 0	0.74	0.13
DTR: T = 1	1.10	0.20

$*$ *sdv* (TO: T = 0) $= 2.41$; *sdv* (true effect: T = 0) $= 1.34$.
$*$ *sdv* (TO: T = 1) $= 3.21$; *sdv* (true effect: T = 1) $= 2.52$.

Comments:

- MLP returns really good results, and it outperforms LR.
- Static setting does not fit here as the RMSEs on STR are bad.

Simulation Results (cont'd)[5]

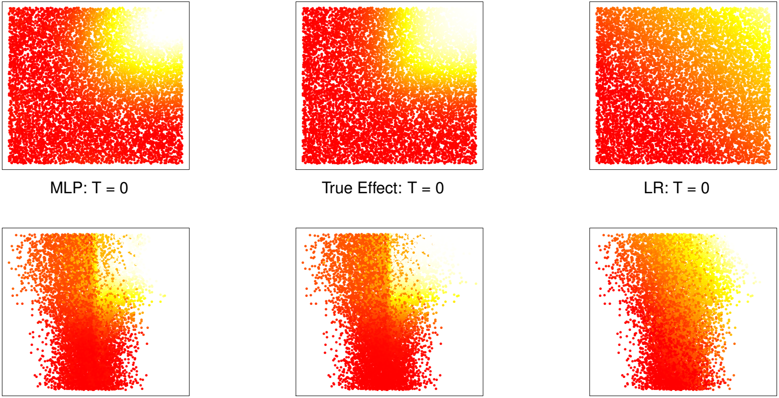


Figure: Heterogeneous Treatment Regime Effect Using Validation and Test Data. The first row corresponds to period $T = 0$ and the second row corresponds to period $T = 1$. In each period: the middle picture visualizes the true treatment effect; the left one is the estimated effect by using Multilayer Perceptron; and the right one is the estimated effect by using Linear Regression.

[5] We thank Wager and Athey (2015) for sharing their visualization code.

Propensity Score Estimation

Use the North Carolina Honors Program data (in illustrative model).

- Estimate

$$\mathbb{P}(W_0 = 1|X_0),\ \mathbb{P}(W_1 = 1|X_0)$$

and

$$\mathbb{P}(W_1 = 1|X_0, Y_1, W_0).$$

- Use Random Forest as a probabilistic classification problem.

Model Estimation - Static Treatment Regime

Use three methods: Linear Regression, Gradient Boosting, and Multilayer Perceptron.

Method	Validation Matching Loss	Test Matching Loss
Linear Regression	11.15	9.77
Gradient Boosting	5.03	4.89
Multilayer Perceptron	3.20	3.27

Comments:

- MLP outperforms other methods.
- All results are bad, which signals the dynamic nature of the data.

STR

Model Estimation - Dynamic Treatment Regime

Period T = 1 - Method	Validation Matching Loss	Test Matching Loss
Linear Regression	1.29	1.29
Gradient Boosting	0.94	1.01
Multilayer Perceptron	0.94	1.01

*sdv(TO: val) = 4.06; sdv(est. true effect: val) = 0.85.
*sdv(TO: test) = 4.03; sdv(est. true effect: test) = 0.92.

Model Estimation - Dynamic Treatment Regime (Cont'd)

Period T = 0 - Method	Validation Matching Loss	Test Matching Loss
Linear Regression	3.29	3.45
Gradient Boosting	1.51	1.63
Multilayer Perceptron	1.14	1.60

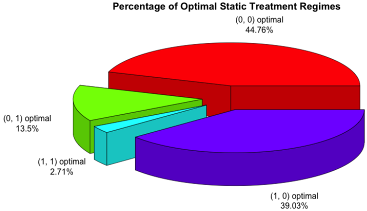
DTR - Use only students who follow the optimal treatment in $T = 1$.

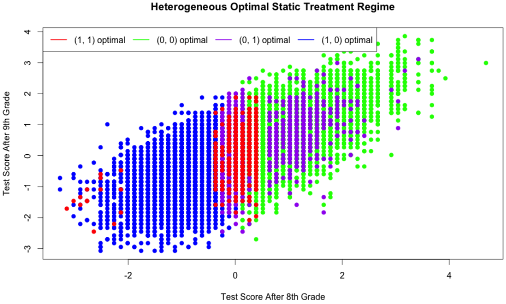
*sdv(TO: val) = 6.94; sdv(est. true effect: val) = 0.84.
*sdv(TO: test) = 7.45; sdv(est. true effect: test) = 0.98.

Remark: The results are worse than that in simulations due to unobserved heterogeneity.

Heterogeneous Optimal Regime

Static Treatment Regime





Heterogeneous Optimal Regime (cont'd)

STR: Estimated gain per student from heterogeneous optimal regime over homogeneous optimal regime $(0,0)$:

$$\frac{1}{\sum_{(0,0)\text{ not opt}} 1}\Bigg(\sum_{(1,1)\text{ opt}}\Big[\widehat{Y}_2(1,1)-\widehat{Y}_2(0,0)\Big]+$$

$$\sum_{(1,0)\text{ opt}}\Big[\widehat{Y}_2(1,0)-\widehat{Y}_2(0,0)\Big]+\sum_{(0,1)\text{ opt}}\Big[\widehat{Y}_2(0,1)-\widehat{Y}_2(0,0)\Big]\Bigg)=0.91.$$

DTR: Estimated gain per student in $T=0$ from heterogeneous optimal W_0 over homogeneous optimal treatment $W_0=0$:

$$\frac{\sum_{W_0=1\text{ optimal}}\Big[\widehat{Y}_2^{W_0=1}-\widehat{Y}_2^{W_0=0}\Big]}{\#\text{obs used in } T=0 \text{ s.t. } W_0=1\text{ opt}}=0.74.$$

$*$ $mean(Y_2)=0; \min(Y_2)=-4.06; \max(Y_2)=3.66$.

Conclusion

- We developed a nonparametric framework using supervised learning to estimate heterogeneous causal regime effects.
- Our model addresses the dynamic treatment setting, the population heterogeneity, and the difficulty in setting up sequential randomized experiments in reality.
- We introduced machine learning approach, in particular deep learning, which demonstrates its estimation power and which is robust to model misspecification.
- We also introduced matching based testing method for the estimation of heterogeneous treatment regime effects. A matching based kernel estimator for variance of these effects is introduced in Appendix.

Variance Estimation - Matching Kernel Approach

Matching

- M matching pairs $\left[(x_0^{d,m}, y_2^{d,m}); (x_0^{d',m}, y_2^{d',m})\right]$, $m = 1, ..., M$.
- Fix x_0^{new}. To estimate $\sigma^2(x_0^{new}) = Var(Y_2^d - Y_2^{d'} | x_0^{new})$, we define
$$\widehat{\epsilon}_m = \widetilde{\tau}_m - \widehat{\tau}_m.$$
- Let $x_0^{mean,m} = \frac{x_0^{d,m} + x_0^{d',m}}{2}$. An estimator for $\sigma^2(x_0^{new})$ is
$$\widehat{\sigma}^2(x_0^{new}) = \frac{\sum_{m=1}^{M} K(H^{-1}[x_0^{mean,m} - x_0^{new}])\widehat{\epsilon}_m^2}{\sum_{m=1}^{M} K(H^{-1}[x_0^{mean,m} - x_0^{new}])}.$$