Residual Weighted Learning For Quantile Optimal Treatment Regimes

by

Xi Hu

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

Statistics

Department of Mathematical and Statistical Sciences
University of Alberta

Abstract

Optimal treatment regime, also called individualized treatment rule, is to seek a rule that assigns a treatment to the subject based on its covariates. It can be used in many areas such as: clinical studies, policy making and economics. In recent years, estimating optimal treatment regime has received considerable attention. However, most of the works focus on estimating the mean-optimal treatment regimes, while not many works have been done on estimating the quantile-optimal treatment regimes. In this thesis, we will focus on estimating quantile-optimal treatment regimes using the residuals.

The quantile-optimal treatment regime is very important in many cases. For instance, if the interest is to find the optimal treatment regime to increase the benefit on the lower tails, or if the outcome distribution is heavily skewed, then estimating the quantile treatment regime is more desirable than the mean. In the former case, the optimal treatment regime should be the one which maximizes some quantile of the potential outcomes, and in the latter case, the treatment regime maximizing the median is more desirable than the one maximizing the mean. Wang et al. firstly worked on the single-index rule case in estimating then quantile-optimal treatment regime. However, the parameters indexing the quantile treatment regime estimated by their framework have large variances [19]. This leads to the problem that their optimal treatment regime may be far away from the true optimal treatment regime, especially when the sample size is small.

To alleviate these problems, we proposed a new framework based on the

residuals, which can be derive by removing the estimated common effect from the outcomes. Motivated by the residual learning framework proposed by Zhou et al. [25], we remove the common effect from the observed outcomes, and define the rest as the residual. The quantile optimal treatment regime can be estimated from these residuals. However, it needs to note that quantile does not have the addition property as mean does. Therefore, this framework is limited to the case that removing the common effect does not change the order of the original outcomes.

From the four simulation examples, it can be seen that removing the common effect from the outcomes can significantly reduce the variances of the parameters indexing the treatment regime. This can stabilize the variance of value function. Further, even using smaller sample size data, the estimated quantile treatment regime is closer to the true treatment regime, comparing to the results from framework proposed by Wang et al.

We also analyzed Data ACTG175 using our proposed framework, and compare the results with Wang's framework on 9 quantile levels. It shows that in this data set, our new framework is comparable with the one from Wang et al. and on some of the quantile levels, our framework has better performance on providing higher value of the potential outcomes.

Acknowledgements

I would like to express my sincerest gratitude to my supervisor Dr. Linglong Kong. During my study in the University of Alberta, he gave tremendous help and patient guidance to my study and research. He is such a responsible supervisor that he went through almost every small detail in my thesis. This thesis can not be finished without his guidance and support.

I also want to thank Dr. Ivan Mizera, Dr. Adam Kashlak, and Dr. Matus Maciak, for being committee members in my thesis examination .

Thanks to the staff in the Department of Mathematical and Statistical Sciences, whose professional services are gratefully acknowledged. It has been a truly enjoyable experience to work with them

I also appreciate the encouragement, advices and help from all my friends in Edmonton during these two years life at University of Alberta.

Finally, special thanks to my family, for their long-lasting understanding and support.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

A treatment regime is a decision rule of whether providing the treatment to a subject based on its characteristics. The treatment regime which leads to the most favourable outcome is called the optimal treatment regime. Recently, studies on estimating the optimal treatment regime have received considerable attention.

Based on different targeting favourable outcomes, studies on estimating optimal treatment regimes mainly includes: estimating the mean-optimal treatment regimes and estimating the quantile-optimal treatment regimes. Currently, most of the studies focus on estimating the mean-optimal treatment regimes [7, 16, 24, 21, 22]. Assuming a larger outcome is preferred, the mean-optimal treatment regime followed by the whole population will yield the most favourable outcome on average [16, 24]. On the other hand, following the quantile-optimal treatment regime will maximize the most favourable outcomes on other criteria such as median or some quantiles. In some applications, maximizing the median outcome or some quantiles of the outcomes

can be more sensible than the average. For example, maximizing the median of the potential outcomes would be more desirable than maximizing the mean when the outcomes follow a heavily skewed distribution. Or in some cases, it is more important to find the best treatment regime which has more effects on the lower tail [19]. For instance, when finding a strategy to improve the earnings for those people whose earnings locate on the lower tail of the earning distribution, the optimal treatment regime should be considered to maximize some quantiles of the outcomes.

This thesis is based on some current studies to further improve the framework of estimating the quantile-optimal treatment regime. To better understand the optimal treatment regime and our proposed framework, we will review some researches in this chapter. In Section 1.1, we review some frameworks on estimating the mean-optimal treatment regime, especially, the residual weighted learning. In section 1.2, we review the frameworks and results of the quantile-optimal treatment regimes. The last section in this chapter is to introduce the contribution of this thesis.

1.1 Mean-Optimal Treatment Regime

In a two-arm randomized trial, we suppose the observed data are $\{(Y_i, A_i, \boldsymbol{X}_i), i = 1, 2, \ldots, n\}$ and assume these data are independent, identically distributed (iid) copies of $\{Y, A, \boldsymbol{X}\}$. Let $\boldsymbol{X}_i = (x_1, x_2, ..., x_l)^T$ denote the covariates of the i^{th} subject which are ascertained prior to the treatment and Y_i denotes the observed outcome. To keep consistent with other literatures in estimating optimal treatment regimes, we assume the larger value of Y_i is desirable

[7, 16, 24, 21, 22]. Let A_i denote the treatment assignment, where $A_i \in \{-1, 1\}$ (or $A_i \in \{0, 1\}$ depends on the methodology). $A_i = 1$ implies that the treatment is assigned to the subject, and $A_i = -1$ (or $A_i = 0$) means the subject does not receive the treatment.

A treatment regime is defined by a function d which maps the value of $\boldsymbol{X}$ to the treatment assignment A. For example, in a clinical study, we let X denote the age and the treatment assignment is denoted by $A \in \{0, 1\}$. A treatment regime $d(X) = I(X > 20)$ implies assigning the treatment to a patient older than 20 years old. The mean-optimal treatment regime is a rule d_0 maximizing the expected outcome when following the rule $d(X)$. There are two main approaches to estimate this mean-optimal treatment regime. We will review these two approaches respectively.

In the first approach, the expected outcome can be mathematically expressed as

$$V(d) = E\left[\frac{I(A = d(\boldsymbol{X}))}{p(A|\boldsymbol{X})}Y\right], \tag{1.1}$$

where $I(\cdot)$ is the indicator function. $V(d)$ is also called *Value Function* [16] associated with d. Then, for a given class of treatment regime D, finding $d_0(\boldsymbol{X})$ is equivalent to solve the following maximization problem:

$$d_0 \in \arg\max_{d \in D} V(d) = \arg\max_{d \in D} E\left[\frac{I(A = d(\boldsymbol{X}))}{p(A|\boldsymbol{X})}Y\right].$$

To estimate this d_0, Qian et al. proposed a two-stage procedure [16]. The first stage in the procedure is to estimate a conditional mean:

$$Q_0(\boldsymbol{X}, A) = E[Y|\boldsymbol{X}, A].$$

The second stage is to estimate the optimal treatment regime which maximizes this conditional mean.

However, the success of Qian's method highly depends on the estimation accuracy of the conditional mean [16, 24]. Moreover, if the approximation space does not provide the treatment effect term which is close enough to the treatment effect term in $Q_0(\boldsymbol{X}, A)$, then this two-stage procedure may not yield optimal treatment regime.

Zhao et al. provided a solution to solve this problem by proposing a framework called outcome weighted learning method. Instead of finding the correct model to approximate $Q_0(X, A)$, this framework can directly estimate the optimal treatment regime.

Zhao et al. claimed that the value function $V(d)$ can be approximated by

$$n^{-1}\sum_{i=1}^{n}\frac{Y_i}{p(A_i|\boldsymbol{X}_i)}I\left(A_i \neq \text{sign}(f(\boldsymbol{X}_i))\right), \tag{1.2}$$

where $f(x)$ is some decision function for the treatment regime. Consider (1.2) as a weighted classification problem, then by applying the hinge loss used in the SVM (Support Vector Machine) context (Corts and Vapnik [2]), $\hat{f}(X)$ can be estimated by minimizing

$$n^{-1}\sum_{i=1}^{n}\frac{Y_i}{p(A_i|\boldsymbol{X}_i)}(1 - A_i f(\boldsymbol{X}_i))^{+} + \lambda_n||f||^2, \tag{1.3}$$

where $x^{+} = \max(x, 0)$, $||f||$ is some norm for f and term $\lambda_n||f||^2$ is used to penalize the complexity of the decision function in order to avoid overfitting. Suppose the decision function $f^{*}(\boldsymbol{X})$ minimizes (1.3) and fix $\text{sign}(0) = -1$,

then the optimal treatment regime is $d^* = \text{sign}(f^*(\boldsymbol{X}))$.

Later, Zhou et al. found three shortages in Zhao's framework [25]. First, a simple shift of the outcomes can affect the estimated optimal treatment regime in the framework of Zhao et al.. This can be seen by adding an extremely large number c to the outcome Y, then the weights are almost identical and the weighed problem can be approximated by an unweighted one. According to this, the framework proposed by Zhao et al. does not hold the nice property that a simple shift on the outcomes will not change the estimated optimal treatment regime. Second, the estimated optimal treatment regime obtained in the framework tries to keep the actually received assignments of the treatments. Third, Zhao et al. did not include the variable selections in their framework.

To alleviate these problems, Zhou et al. proposed a new framework: residual weighted learning (RWL) to estimate the optimal treatment regime. In this framework, the optimal treatment regime is estimated by minimizing:

$$n^{-1}\sum_{i=1}^{n}\frac{y_i - \hat{g}^*(\boldsymbol{X}_i)}{p(A_i|\boldsymbol{X}_i)}I(a_i \neq \text{sign}(f(\boldsymbol{X}_i))), \tag{1.4}$$

where $\hat{g}^*$ is an estimate of $E\left[\frac{Y}{2p(A|\boldsymbol{X})}|\boldsymbol{X}\right]$ and reflects the common effects for both treatment arms. Moreover, the common effect can be estimated by any appropriate regression method. For example, Zhou et al. used $E\left[\frac{Y_i}{2p(A_i|\boldsymbol{X}_i)}|\boldsymbol{X}_i\right] = \hat{\beta}_0 + \boldsymbol{X}^T\hat{\beta}$ to estimate $\hat{g}^*(\boldsymbol{X})$.

Define $y_i - \hat{g}^*(\boldsymbol{X}_i)$ as the estimated residuals, then (1.4) can be viewed as a classification problem weighted by the residual $(y_i - \hat{g}^*(\boldsymbol{X}_i))/p(A_i|\boldsymbol{X}_i)$. To

solve (1.4), Zhou et al. applied the smoothed ramp loss defined by

$$T(u) = \begin{cases} 0, & \text{if } u \geq 1, \\ (1-u)^2, & \text{if } 0 \leq u < 1, \\ 2-(1+u)^2 & \text{if } -1 \leq u < 0, \\ 2 & \text{if } u < -1. \end{cases}$$

and concave-convex procedure to derive the optimal decision function f^*. Then optimal treatment regime is $d^* = \text{sign}(f^*(\boldsymbol{X}))$.

The second approach to estimating the mean-optimal treatment regime is based on the missing data analogy. The same with the first approach, Y denotes the continuous outcome and the larger value of Y is assumed desirable. $\boldsymbol{X}$ denotes the characteristics of the subject. A treatment regime is also defined as the same with the one in the first approach. However, the treatment assignments are taken from $A \in \{0, 1\}$ instead of $\{-1, 1\}$. Let $Y^*(A)$ denote the potential outcome if a subject is given the treatment A. Then it can be seen that $E[Y^*(A)]$ represents the mean of all the subjects when they receive the treatment A. For any treatment regime $d(\boldsymbol{X})$, we can further define the potential outcome as $Y^*(d) = Y^*(1)d(\boldsymbol{X}) + Y^*(0)(1 - d(\boldsymbol{X}))$ if following the regime. The optimal treatment regime d_0 in D (a class of treatment regimes of interest) is

$$d_0 = \arg\max_{d \in D} E\left[Y^*(d(\boldsymbol{X}))\right].$$

The expectation

$$V(d) = E\left[Y^*(d(\boldsymbol{X}))\right], \tag{1.5}$$

is called *Value function* associated with treatment regime d in this approach.

In order to estimate the value function (1.5) based on observed data, three assumptions must be satisfied:

1. The stable unit treatment value assumption.

2. Consistency assumption: $Y = Y^*(1)A + Y^*(0)(1-A)$

3. The no-unmeasured-confounders assumption: $A \perp \{Y^*(0), Y^*(1)|\boldsymbol{X}\}$

The stable unit treatment value assumption ensures that the potential outcome of one subject will not be affected by the assignment of the treatment to other subjects. The second assumption ensures that the unmeasured confounders do not exist.

Zhang et al. followed the missing data analogy as in Cao et al. [1] to estimate the optimal treatment regime d_0. Let $C(\beta) = Ad(\boldsymbol{X}, \beta) + (1-A)(1-d(\boldsymbol{X}, \beta))$, where β is a vector of the parameter indexing a treatment regime. In the missing data analogy, the 'full data' in this case is $\{Y^*(d), X\}$ and the observed data is $\{C, CY^*(d), \boldsymbol{X}\} = \{C, CY, \boldsymbol{X}\}$. Maximizing the value function (1.5) is then equivalent to maximize the simple inverse probability weighted estimator (IPWE), which is given in

$$IPWE(\beta) = n^{-1} \sum_{i=1}^{n} \frac{C(\beta)Y_i}{p(A_i|\boldsymbol{X}_i)}. \tag{1.6}$$

This estimator (1.6) is a consistent estimator of $V(d)$ if the propensity score $p(A_i|\boldsymbol{X}_i)$ is correctly specified [21].

In addition, following Robins et al. [17], Zhang et al. also provided an alternative doubly robust estimator: Augmented Inverse Probability Weighted

Estimator (AIPWE) which is given by

$$AIPWE(\beta) = n^{-1} \sum_{i=1}^{n} \left\{ \frac{C(\beta) Y_i}{p(A_i|\boldsymbol{X}_i)} - \frac{C(\beta) - p(A_i|\boldsymbol{X}_i)}{p(A_i|\boldsymbol{X}_i)} \hat{m}(\boldsymbol{X}_i) \right\}, \qquad (1.7)$$

where $\hat{m}(\mathbf{x}) = \hat{\mu_1}(\mathbf{x}) I(d(\mathbf{x}) = 1) + \hat{\mu_0}(\mathbf{x}) I(d(\mathbf{x}) = 0)$, $\hat{\mu_1}$ is the estimator of $E(Y|\boldsymbol{X} = \mathbf{x}, A = 1)$ and $\hat{\mu_0}$ is the estimator of $E(Y|\boldsymbol{X} = \mathbf{x}, A = 0)$. AIPWE is a consistent estimator of $V(d)$ if either the propensity score $p(A_i|\boldsymbol{X}_i)$ or $\hat{\mu_0}$ and $\hat{\mu_1}$ is misspecified.

It is worth noting that by taking $\hat{g}^*(\boldsymbol{X}) = \hat{m}(\boldsymbol{X})$, maximizing (1.7) is equivalent to minimizing (1.4). Thus, Zhou's methodology links AIPWE method to a classification method, and can be also viewed as an improved AIPWE methodology.

1.2 Quantile-Optimal Treatment Regime

If the interest is not of maximizing the average of the outcomes, but some τ^{th} quantile of the potential outcomes, then it is more desirable to consider the quantile-optimal treatment regime. For instance, we suppose the outcome Y can be modelled by $Y = 1 + 3X + A(2 - 5X) + (1 + AX)\epsilon$, where X follows $Uniform(0, 1)$ and denotes covariates of subjects, ϵ follows $Normal(0, 1)$ which is a random error, and $A \in \{0, 1\}$ denotes the treatment assignment. Consider 4 treatment regime: (1). $A_1 = I(X \leq 3/4)$, (2). $A_2 = I(X \leq 0.45)$, (3). $A_3 = I(X \leq 0.2)$, (4). $A_4 = I(X \leq 0.1)$. Based on a Monte Carlo experiment with 10^6 observations, we can derive the mean, median, the 0.25 quantile and 0.1 quantile of the potential outcome distribution corresponding

to these four regimes. The results are reported in Table 1.1. Since A_1 is the mean-optimal treatment regime, then when the interest is of maximizing the average, regime A_1 should be selected. If the interest is of maximizing the median, then it can be observed that the regime A_2 is better than A_1. Further, if the interest is to maximize the outcome on the lower tail, then regimes A_3 or A_4 are better choices. This example illustrates that in some cases, estimating quantile-optimal treatment regime is more desirable than the mean-optimal treatment regime.

Table 1.1: Summary results for the example

Regime	A_1	A_2	A_3	A_4
Mean	**2.901**	2.894	2.801	2.675
$Q_{0.10}$	-4.245	-4.376	-3.947	$\mathbf{-3.931}$
$Q_{0.25}$	-0.806	-0.865	$\mathbf{-0.744}$	-0.802
$Q_{0.50}$	2.928	**2.934**	2.801	2.670

Wang et al. proposed model-free framework to estimate the quantile-optimal treatment regime [19]. In the study of Wang et al., instead of maximizing the average of the potential outcome $E(Y^*(d))$ in the framework of Zhang et al. [21], they considered the maximization of some quantiles of the potential outcomes, which defined as $Q_\tau(Y^*(d))$, where τ is a quantile level of interest. This $Q_\tau(Y^*(d))$ is called value function associate with treatment regime d. Moreover, this quantile-optimal treatment regimes can be estimated by

$$\arg\max_{d\in\mathbb{D}} Q_\tau(Y^*(d)),$$

where $\mathbb{D}$ denotes a collection of all treatment regimes.

To estimate the quantile optimal treatment regime, Wang et al. used the

induced missing data framework motivated by Zhang et al. [21]. Defining

$$C(\beta) = Ad(\boldsymbol{X}, \beta) + (1 - A)(1 - d(\boldsymbol{X}, \beta),$$

then for any quantile level $\tau \in (0, 1)$, the τ^{th} quantile of the outcomes can be estimated by

$$\hat{Q}_\tau(\beta) = \arg\min_{a \in \mathbb{R}} n^{-1} \sum_{i=1}^{n} C_i(\beta)\rho_\tau(Y_i - a),$$

where $\rho(u) = u(\tau - I(u < 0))$ is the quantile loss function. Using genetic algorithm, the estimator of β_0 can be derived by solving

$$\hat{\beta}_n = \arg\max_{\beta \in \mathbb{B}} \hat{Q}_\tau(\beta),$$

where $\mathbb{B}$ is a compact subset of R^l and l is the number of the features (number of the columns of $\boldsymbol{X}$).

Another contribution of Wang et al. is that they investigated the asymptotic theory on estimating the quantile-optimal treatment regime. By introducing

$$\hat{\beta}_n = \arg\max_{\beta \in \mathbb{B}} P_n g(\cdot, \beta, \hat{m}_n),$$

as an alternative expression of $\hat{\beta}_n$, where P_n is the empirical expectation and

$$\begin{aligned} g(\cdot, \beta, \hat{m}_n) &= C(\beta) I\{Y - m > 0\} \\ \hat{m}_n &= \sup\{m : \sup P_n g(\cdot, \beta, m)\}, \end{aligned}$$

Wang et al. showed that the estimated parameter indexing the quantile-optimal treatment regime has a nonstandard convergence rate and a non-

normal limiting distribution. Their theoretical approach can be also applied to investigate the asymptotic distribution of the estimators in the framework of finding the mean-optimal treatment regime proposed by Zhang et al. [21].

To the best of our knowledge, Wang et al. are the first group who focused on estimating the static quantile optimal treatment regimes. Their rich work opens a gate of analyzing the quantile-optimal treatment regime.

However, even for some simple models which consists a single covariate, the estimated parameters indexing the quantile-treatment regime have relative large variances, some variances can be even larger than the absolute value of the parameters. Also, when the sample size is small, the estimated optimal treatment regime may not be stable and can be relatively far away from the true one.

1.3 Contributions of My Thesis

To alleviate the problems in the framework proposed by Wang et al., we propose a new framework which minimizes the value function of the residuals instead of the outcomes to estimate the quantile-optimal treatment regime. In this framework, the residuals are defined the same with the one in Zhou's residual weighted learning framework, which can be derived by removing the common effects from the outcomes [25].

One of the challenges needs to be noticed here is that, unlike calculating the expected value of the outcome, the τ^{th} quantile of the outcome may not equal to the sum of the τ^{th} quantile of the common effect and the τ^{th} quantile of the treatment effect. Therefore, to minimize the τ^{th} quantile of the residual

may not equivalent to minimize the τ^{th} quantile of the outcomes if not follows the treatment regime. It will be seen that by removing the common effect, the prediction accuracy is increasing with limited sample size, and the variances of the predicted coefficients become much smaller. Further, as the common effect can be estimated by regressions, then the feature selection techniques such as *least absolute shrinkage and selection operator* (LASSO) can be applied to select those most impactive features to the potential outcomes.

The rest of this thesis is organized as follows. In chapter 2, the framework of estimating the quantile-optimal treatment regime will be introduced. Simulation studies and a real data analysis will be reported in Chapter 3. Last, results and possible future works will be summarized in Chapter 4.

Chapter 2

Methodology

In this chapter, we propose a framework of estimating quantile optimal treatment regimes. Suppose the observed data are $(Y_i, A_i, \boldsymbol{X}_i)$ for $i = 1, 2, .., n$, and these data are independent, identically distributed (iid) copies of $\{Y, A, \boldsymbol{X}\}$ for all i. Let $A \in \{0, 1\}$ denote the treatment assignment, such that, if $A = 1$, the treatment will be assigned to the subject, and while if $A = 0$, the subject will not receive the treatment. $\boldsymbol{X} = (x_1, x_2, ..., x_l)^T$ denotes the covariates of the subjects which are ascertained prior to the treatment and Y denotes the observed outcome. We also assume that the larger value of Y is preferred to keep consistent with other literatures in estimating optimal treatment regimes [22, 23, 25]. Our target is to estimate the quantile-optimal treatment regime given a class of treatment regimes $\mathbb{D} = \{I(\boldsymbol{X}^T\beta > 0) : \beta \in \mathbb{B}\}$, where β is the parameter indexing a treatment regime and $\mathbb{B}$ is a compact subset of R^l. Due to its simplicity and easy interpretability, this class is popular in practice [19, 21, 22, 25].

Directly applying the framework provided by Wang et al. to estimate β,

the variances of the estimated $\hat{\beta}$ can be considerably large [19]. This may cause the problem that if limiting the sample size to a relatively small number, the estimated optimal treatment regime may not be stable and can be relatively far away from the true one.

The residual weighted learning framework proposed by Zhou et al. provides a potential solution, where the optimal treatment regime will be estimated by residuals [25]. Figure 2.1 is an example which illustrates the idea of the residual framework.

From Figure 2.1, it can be observed that a large value of Y may still be considered small comparing to other subjects with similar covariates. For example in the left figure in Figure 2.1, the subjects with $x > 0.5$ marked by dots have larger values of Y comparing to other subjects marked by dots. However, if compared to the subjects with $x > 0.5$ and marked by $+$, these subjects marked by dots have relatively smaller Y. Since the treatment effect is the one of interest, then it is more sensible to consider the residuals.

This motivates us to apply the similar methodology. Defining the residuals as removing the common effect from the outcomes, we use the residuals instead of the outcomes to estimate the quantile-optimal treatment regime.

However, it needs to note that, unlike the mean, quantile do not have the additive property. This leads to the fact that minimizing the τ^{th} quantile of the residuals may not be equivalent to minimizing the τ^{th} quantile of the outcomes if not follows the treatment regime. Therefore, this methodology is limited to the case that removing the common effect from the outcome would not change the original order of the outcome. See Figure 2.2 as an example.

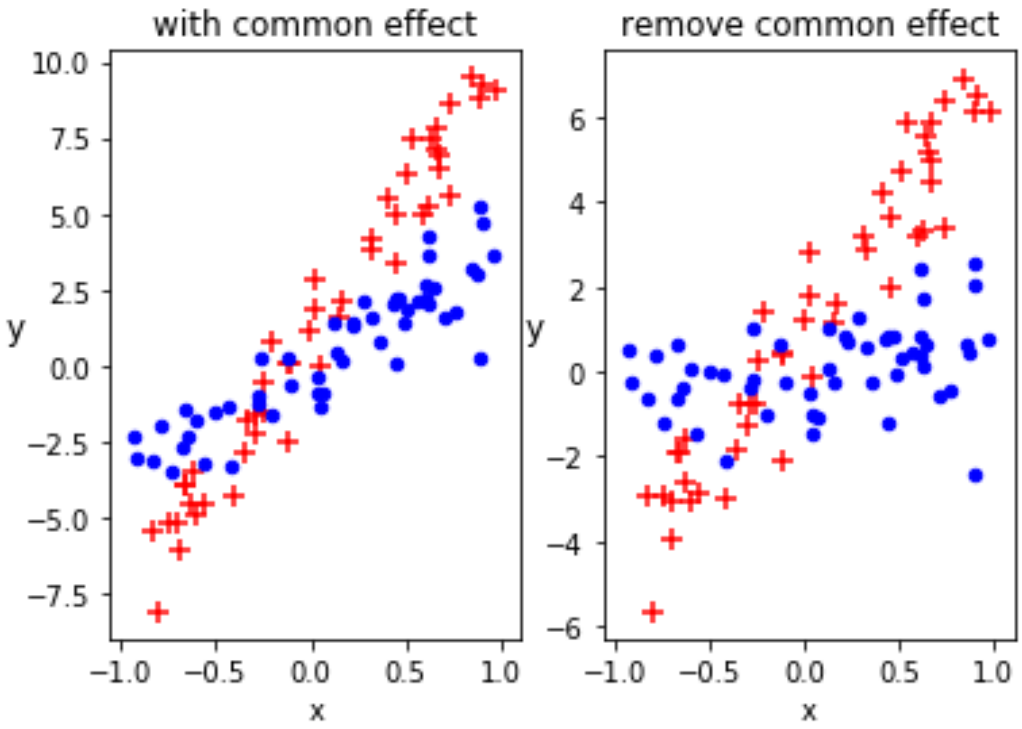

Figure 2.1: This is an example of estimating quantile-optimal treatment regime. The figure on the left is the raw data, which consists of a single covariate X, two treatment assignments $A = 0$ and $A = 1$ and continuous outcome Y. The '+' marker is used to denote $A = 1$ and the dot marker is used to denote $A = 0$. In this example, $E[Y|X, A] = 3\boldsymbol{X} + (6\boldsymbol{X} + 1)A$. The propensity score $P(A|\boldsymbol{X}) = 0.5$. In the figure on the right, the residuals $(6\boldsymbol{X} + 1)A$ was shown.

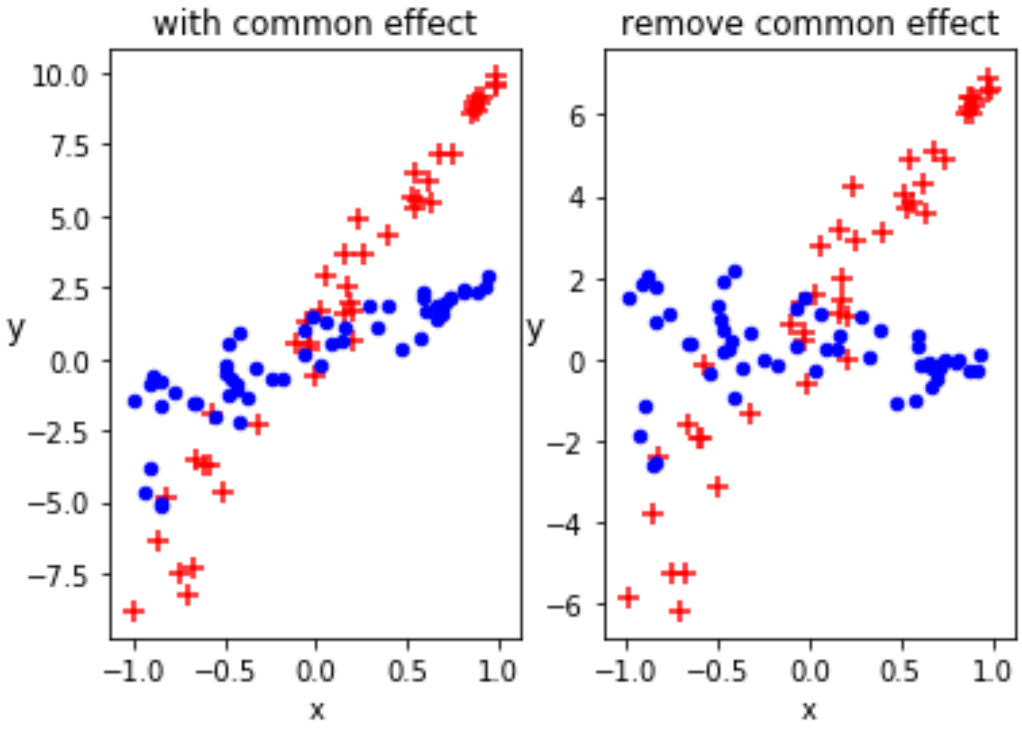

Figure 2.2: This is an example of estimating quantile-optimal treatment regime. The figure on the left is the raw data, which consists of a single covariate $\boldsymbol{X}$, two treatment assignments $A = 0$ and $A = 1$ and continuous outcome Y. The random error is $(1 - 0.8ax)\epsilon$, where ϵ follows standard normal distribution.The '+' marker is used to denote $A = 1$ and the dot marker is used to denote $A = 0$. In this example, $E[Y|X, A] = 3X + (6X + 1)A$. The propensity score $P(A|X) = 0.5$. In the figure on the right, the residuals $(6X + 1)A$ was shown.

2.1 Estimation of the common effect

Considering the model

$$Y = \mu(\boldsymbol{X}) + \delta(\boldsymbol{X})A + \omega(\boldsymbol{X}, A)\epsilon, \tag{2.1}$$

where $\mu(\boldsymbol{X})$ models the common effect, $\delta(\boldsymbol{X})$ models the treatment effect and $\omega(\boldsymbol{X}, A)\epsilon$ is the random error, we can see that for those subjects without taking the treatment, then $A = 0$. Equation (2.1) will then become to

$$Y = \mu(\boldsymbol{X}) + \omega(\boldsymbol{X}, 0)\epsilon. \tag{2.2}$$

If we assume that ϵ has mean 0, then

$$\mu(\boldsymbol{X}) = E[Y|\boldsymbol{X}, A = 0] = E[Y|\boldsymbol{X}].$$

This $\mu(\boldsymbol{X})$ can be posited by a parametric regression model $\mu(\boldsymbol{X}, \eta)$ such as linear regression with finite-dimensional parameter η.

Another benefit from this method is that it is easy to apply feature selection methods such as: *least absolute shrinkage and selection operator* (LASSO)[18] or *smoothly clipped absolute deviation* (SCAD) [8]. For example, LASSO is to solve

$$\hat{\boldsymbol{\eta}} = \arg\min_{\boldsymbol{\eta}} \sum_{i=1}^{m} \left(Y_i - \mathbf{x}_i^{\mathrm{T}}\boldsymbol{\eta}\right)^2 + P_\lambda\left(\boldsymbol{\eta}\right),$$

where m is the number of the subjects without assigning the treatment and

the penalization term $P_\lambda(\boldsymbol{\eta})$ is given by

$$\lambda \sum_{j=1}^{l} |\eta_j|$$

where $\lambda \geq 0$ is the regularization parameter which can be estimated by cross-validation. If $\lambda = 0$, then it can be seen that $\hat{\boldsymbol{\eta}}$ is estimated by regular linear regression. Further, as λ increases from 0 to sufficiently large, LASSO will continuously shrink some of the coefficients exactly to 0. Hence, LASSO can be efficient to select most important features and estimate the regression parameters simultaneously. In this thesis, we use linear regression with LASSO to select the most important features and estimate the common effect.

Denoting the estimated common effect as $\hat{\mu}(\boldsymbol{X})$, we can further define the estimated residuals as

$$\hat{R} = Y - \hat{\mu}(\boldsymbol{X}).$$

This $\hat{R}$ will be used to estimate the optimal treatment regime. Further, following the potential outcome framework, we let $\hat{R}^*(1)$ be the potential residual with assigning the treatment, and let $\hat{R}^*(0)$ be the potential residual without assigning the treatment. Then we can only observe either $\hat{R}^*(1)$ or $\hat{R}^*(0)$. Moreover, the potential residual outcome can be expressed as:

$$\hat{R}^* = \hat{R}^*(1)d(\boldsymbol{X}) + \hat{R}^*(0)(1 - d(\boldsymbol{X}))$$

2.2 Quantile-Optimal treatment regime

To estimate the τ^{th} quantile level optimal treatment regime where $\tau \in (0, 1)$, we will estimate it from $\hat{R}^*$ instead of Y^* in Wang's framework, which means, the optimal treatment regime will be defined by

$$\arg\max_{d\in\mathbb{D}} Q_\tau(\hat{R}^*(d)) \tag{2.3}$$

where $\mathbb{D}$ denotes a collection of treatment regimes, $Q_\tau(\hat{R}^*(d))$ is the τ^{th} quantile of $\hat{R}^*(d)$. More specifically,

$$Q_\tau(\hat{R}^*(d)) = \inf\{t : F(\hat{R}^*(d)) \geq \tau\}$$

with $F(t)$ is the distribution of $\hat{R}^*(d)$. By defining a treatment regime $d(\boldsymbol{X}, \beta) = I(\boldsymbol{X}^T\beta > 0)$ indexed by β, the goal is to estimate $\hat{\beta}$ which satisfy

$$\hat{\beta} = \arg\max_{\beta\in\mathbb{B}\subset R^l} Q_\tau(\hat{R}^*(d_\beta), \tag{2.4}$$

where $\hat{R}(d_\beta)$ denotes the residual which follows regime d_β indexed by β.

To estimate $\hat{\beta}$, we first consider a randomized trial where the propensity score $p(A|\boldsymbol{X})$ is ordinarily known and constant. We can then follow the induced missing data framework by defining

$$C(\beta) = Ad(\boldsymbol{X}, \beta) + (1 - A)(1 - d(\boldsymbol{X}, \beta).$$

We can consider the full but not completed observed data of interest are $\{\hat{R}^*(d_\beta), \boldsymbol{X}\}$ and the data which can be observed are $\{C(\beta), C(\beta)\hat{R}^*(d_\beta), \boldsymbol{X}\} =$

$\{C(\beta), C(\beta)\hat{R}, \boldsymbol{X}\}$. Note that $C(\beta) = 1$ if and only if $A = d(\boldsymbol{X}, \beta)$, while $C(\beta) = 0$ otherwise. This implies that the observed data are those data with the assignment of the treatment following the treatment regime.

Note that $Q_\tau(\hat{R}^*(d_\beta)$ can be estimated by quantile regression by defining:

$$\hat{Q}_\tau(\beta) = \arg\min_a \frac{1}{n} \sum_{i=1}^{n} C_i(\beta)\rho_\tau(\hat{R}_i^* - a),$$

as $\hat{Q}_\tau(\beta)$ is a consistent estimator of the τ^{th} quantile of $Q_\tau(\hat{R}^*(d_\beta)$ shown in the following theorem. Here $\rho_\tau(u) = u(\tau - I(u < 0))$ is the quantile loss function.

Theorem 1 *Suppose $R^*(0)$ and $R^*(1)$ follow continuous distributions and their density functions are bounded and second order differentiable, then for any $\beta \in \mathbb{B}$, we have*

$$\hat{Q}_\tau(\beta) \to Q_\tau(\hat{R}^*(d_\beta))$$

in probability.

Proof: The law of large numbers and convexity imply $1/n \sum_{i=1}^{n} C_i(\beta)\rho_\tau(\hat{R}_i - a)$ uniformly converge to $E[C_i(\beta)\rho_\tau(\hat{R}_i - a)]$ in a in probability. Further, when $C_i(\beta) = 1$, we have $\hat{R}_i = \hat{R}_i^*(d_\beta)$. Therefore, by considering a randomized trial case, we can have:

$$E[C_i(\beta)\rho_\tau(\hat{R}_i - a)] = E[C_i(\beta)\rho_\tau(\hat{R}^*(d_\beta) - a)]$$

By the Law of total expectation and the conditional independence of $C_i(\beta)$

and $\rho_\tau(\hat{R}^*(d_\beta) - a)$ given $\boldsymbol{X}_i$, we can further derive that

$$\begin{aligned} E[C_i(\beta)\rho_\tau(\hat{R}_i - a)] &= E[C_i(\beta)\rho_\tau(\hat{R}^*(d_\beta) - a)] \\ &= E\{E[C_i(\beta)\rho_\tau(\hat{R}^*(d_\beta) - a)|\boldsymbol{X}_i]\} \\ &= E\{E[C_i(\beta)|\boldsymbol{X}_i]E[\rho_\tau(\hat{R}^*(d_\beta) - a)|\boldsymbol{X}_i]\} \end{aligned}$$

Since in a randomized trial, $E[C_i(\beta)|\boldsymbol{X}] = 1/2$, then we have

$$E[C_i(\beta)\rho_\tau(\hat{R}_i - a)] = \frac{1}{2}E[\rho_\tau(\hat{R}^*(d_\beta) - a)]$$

Moreover, since $E[\rho_\tau(\hat{R}^*(d_\beta) - a)]$ is convex, continuous and can be minimized by $a = Q_\tau(R^*(d_\beta))$, then by the results of M-estimation, we can have

$$\hat{Q}_\tau(\beta) \to Q_\tau(\hat{R}^*(d_\beta)),$$

and the consistency of the estimation of $\hat{R}(d)$ can be derived.

By showing that $\hat{Q}_\tau(\beta)$ is a consistent estimator of the τ^{th} quantile of $Q_\tau(\hat{R}(d_\beta))$, the τ^{th} quantile treatment regime can be estimated by

$$\hat{\beta} = \arg\max \hat{Q}_\tau(\beta),$$

where motivated by Zhang et al. and Wang et al. This can be solved by Genetic Algorithm. The entire procedure can be summarized below

1. Select the subjects of those without assigning the treatment (those with $A = 0$), and apply regression method with LASSO or other feature selection technologies such as SCAD to select the most important features

and estimate the common effect. From this step, we can determine the number of the features l and estimate the common effect.

2. Derive the residuals R by removing the estimated common effect from the original responses.

3. Take β as the parameter indexing a treatment regime, and $C(\beta) = Ad(\boldsymbol{X}, \beta) + (1 - A)(1 - d(\boldsymbol{X}, \beta)$ as the weight, then using traditional quantile regression, we can derive a parametrictic function $f(\beta)$ for estimating $\hat{Q}_\tau(\beta)$.

4. Apply Genetic algorithm on this $f(\beta)$ to search for the value of β which indexes the quantile-optimal treatment regime.

So far we only considered the randomized trial with the propensity score $p(A|X)$ is constant and equals to 0.5. For the observational studies, the propensity score needs to be estimated. Following Zhang et al., let $p(A|X) = \pi(X, \gamma)$. This $\pi(X, \gamma)$ can be estimated by a logistics regression model $\pi(\boldsymbol{X}; \gamma) = \exp(\gamma^T \tilde{\boldsymbol{X}})/(1+\exp(\gamma^T \tilde{\boldsymbol{X}}))$, $\tilde{\boldsymbol{X}} = (1, \boldsymbol{X}^T)^T$. We can then estimate the propensity score for the observational study using

$$\hat{\pi} = \pi(\boldsymbol{X}, \hat{\gamma})d(\boldsymbol{X}, \beta) + (1 - \pi(\boldsymbol{X}, \hat{\gamma})(1 - d(\boldsymbol{X}, \beta)).$$

Following the missing data analogy, and IPWE estimator, in an observational study, the new $\hat{Q}_\tau(\beta)$ can be expressed as:

$$\hat{Q}_\tau(\beta) = \arg\min_a \frac{1}{n} \sum_{i=1}^{n} \frac{C_i(\beta)}{\hat{\pi}} \rho_\tau(\hat{R}_i^* - a),$$

By using the same procedure as before, we can obtain the quantile-optimal treatment regime in an observational study.

Chapter 3

Simulation

In this chapter, we will use numerical simulations to illustrate our proposed method and compare the estimated quantile optimal treatment regimes with the ones estimated by Wang's framework. In the simulation examples, we generate covariates $\boldsymbol{X} = (x_1, \ldots, x_4)$, which are independent uniform random variables following $U(-1, 1)$. The treatment assignment A, which is independent of $\boldsymbol{X}$, are taken from $\{0, 1\}$ with probability $P(A = 0) = P(A = 1) = 0.5$. Therefore, the propensity score $p(A|\boldsymbol{X}) = 0.5$ can be considered as constant. Further, the outcome $Y(\boldsymbol{X}, A)$ can be generated by the model:

$$Y(\boldsymbol{X}, A) = \mu(\boldsymbol{X}) + \delta(\boldsymbol{X})A + \omega(\boldsymbol{X}, A)\epsilon,$$

where $\mu(\boldsymbol{X})$ models the common effect, and $\delta(\boldsymbol{X})$ models the treatment effect. In the following examples, this random variable ϵ follows either standard normal distribution or *Gamma* distribution.

We will consider the class of treatment regimes having the form $I(\boldsymbol{X}^T\beta >$

0) due to its simplicity and interpretability. Unlikely to the mean optimal treatment regime, the true parameter value of quantile-optimal treatment regime is very complicated. Therefore, instead of calculating the true optimal treatment regime, we use the estimated quantile treatment regime obtained by Monte Carlo experiments with the sample size $n = 10^5$. This treatment regime will be considered as the ideal treatment regime and called *True values*.

In all of the simulation examples, the accuracy rate is defined by

$$\mathbb{P}_n^*[I(d(\boldsymbol{X}) = d^*(\boldsymbol{X}))] \tag{3.1}$$

where $\mathbb{P}_n^*$ is the empirical average, $I(\cdot)$ is the indicator function, $d(\boldsymbol{X})$ is the true optimal treatment regime estimated by Monte Carlo experiments, and $d^*(\boldsymbol{X})$ is the optimal treatment regime estimated either by our proposed framework or by the method proposed by Wang et al. This accuracy rate can be interpreted as: the higher the accuracy rate, the closer the estimated treatment regimes to the true ones.

3.1 Simulations

Example 1. We take a simple example with only one covariate x_1. In this example, $\mu(\boldsymbol{X}) = 3x_1$, $\delta(\boldsymbol{X}) = 6x_1 + 1$ and $\omega(\boldsymbol{X}, A) = 1 - 0.8ax_1$ and the quantiles are chosen as $\tau = 0.1$ and $\tau = 0.25$. The random error ϵ follows $Normal(0, 1)$. The class of treatment regime is $\mathbb{I}(\beta_0 + \beta_1 x_1 > 0)$. The treatment assignment which indicated by A follows a Bernoulli distribution, and will be considered in two cases based on different success probability. The first

case takes the success probability 0.5, while in the second case, the success probability satisfies $logit(P(A_i = 1|X_i) = logit(X_i - 1)$, where $logit(x)$ is defined by $logit(x) = \exp(x)/(1 + \exp(x))$. The sample size considered in this example is $n = 500$ and $n = 1000$, and simulation is repeated 500 times for both of the cases.

Note that *Example 1* is a simple example with only one feature, thus there is no need to apply LASSO to perform the feature selection. Therefore, the common effect can be easily estimated by the Linear Regression and then be removed from the original observed outcomes Y.

Case 1 in Example 1. Table (3.1) and Table (3.1) report the estimated β values, where *RCE* which short for *Remove Common Effect*, denotes our proposed method, and *WCE*, which is short for *With Common Effect*, denotes the method proposed by Wang et al.. '*true value*', as mentioned previously, is derived from Monte Carlo experiments with the sample size $n = 10^5$ and served as true the optimal treatment regime.

It can be seen that in either of two quantiles ($\tau = 0.1$ and $\tau = 0.25$), removing the common effect from the outcomes can significantly increase the accuracy rates. Further, the standard deviation also becomes smaller than the ones if not removing the common effect. Figure 3.1 and Figure 3.2 provide visualized comparisons for this example. These two figures are the scatter plots for a test data set with sample size $n = 500$. The vertical lines, defined as $\hat{x} = -\hat{\beta}_0/\hat{\beta}_1$, denote the true decision boundary and the estimated one respectively.

Case 2 in Example 1. This case of *Example 1* is used to simulate an observed data case. In this case, the treatment assignment A follows binomial

Table 3.1: Summary results for the first example when $\tau = 0.1$. RCE stands for *removing common effect* and WCE stands for *with common effect*. The numbers in the parenthesis are standard deviations. The last column is the accuracy rate defined by $\mathbb{P}_n^*$ using the true value in the last row.

		β_0	β_1	accuracy rate
$n = 500$	RCE	-0.385 (0.034)	0.922(0.014)	0.975
	WCE	-0.332(0.256)	0.370(0.774)	0.745
$n = 1000$	RCE	-0.387 (0.025)	0.921(0.01)	0.980
	WCE	-0.357(0.227)	0.371(0.764)	0.788
true value		-0.424	0.906	

Table 3.2: Summary results for the first example when $\tau = 0.25$. RCE stands for *removing common effect* and WCE stands for *with common effect*. The numbers in the parenthesis are standard deviations. The last column is the accuracy rate defined by $\mathbb{P}_n^*$ using the true value in the last row.

		β_0	β_1	accuracy rate
$n = 500$	RCE	-0.377 (0.037)	0.925(0.015)	0.99
	WCE	-0.279(0.247)	0.334(0.855)	0.802
$n = 1000$	RCE	-0.3774 (0.028)	0.926(0.011)	0.992
	WCE	-0.351(0.193)	0.604(0.686)	0.925
true value		-0.392	0.92	

distribution with the success probability satisfying $logit(P(A_i = 1|X_i) = X_i - 1$. The corresponding propensity score $\pi(x, a) = P(A = 1|X) = 0.272$ can be estimated by logistic regression.

Table (3.3) reports the simulation results of this observed data case. The same with case 1, removing the common effect can increase the accuracy rate comparing to the one with keeping the common effect in the outcomes. This implies that, with the same sample size, by removing the common effect, the estimated quantile optimal treatment regime can be much more closer to the ideal one than the regime estimated with common effect.

Further, from both of the cases in *Example 1*, it can be seen that increasing the sample size can also increase the accuracy rate, and the variances of the estimated coefficients of the optimal treatment regime are smaller when removing the common effect, which make the estimation of the optimal treatment regime more stable. This implies that removing the common effect from the outcomes can provide better estimation of the optimal treatment regime if the sample size is relatively small.

Table 3.3: Summary results for the first example when $\tau = 0.25$. RCE stands for *removing common effect* and WCE stands for *with common effect*. The numbers in the parenthesis are standard deviations. The last column is the accuracy rate defined by $\mathbb{P}_n^*$ using the true value in the last row.

		β_0	β_1	accuracy rate
$n = 500$	RCE	-0.372(0.043)	0.927(0.017)	0.981
	WCE	-0.245(0.245)	0.29(0.893)	0.794
$n = 1000$	RCE	-0.3782 (0.033)	0.925(0.013)	0.986
	WCE	-0.297(0.219)	0.457(0.81)	0.893
true value		-0.399	0.916	

Example 2. In the second example, we still use a single covariate and

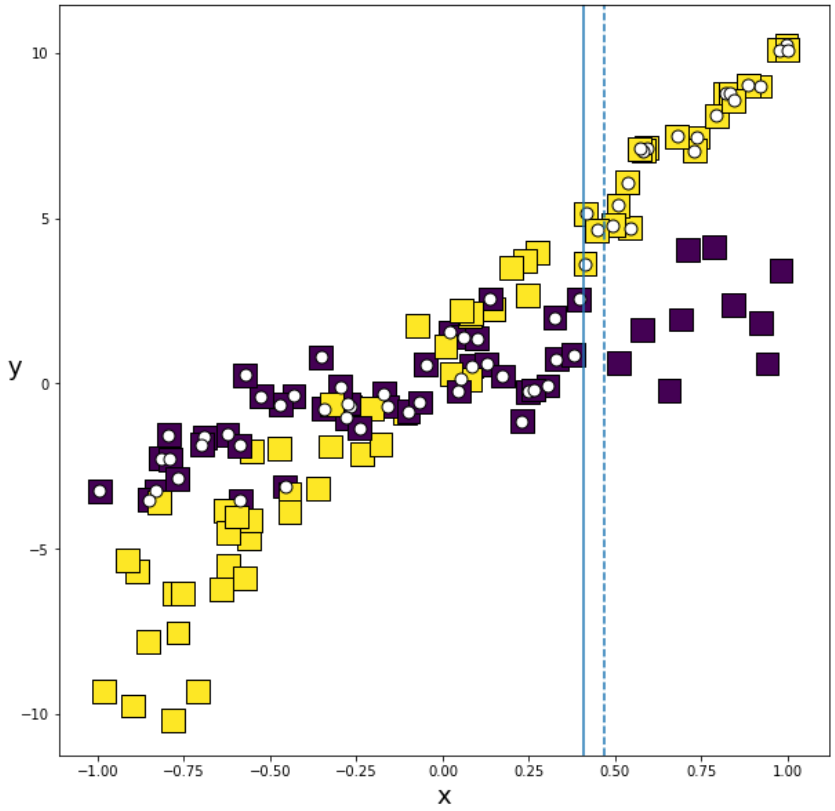

Figure 3.1: This is the simulation result for the first example when $\tau = 0.25$. In this figure, the light squares denote the subjects with treatment, while the dark squares are the ones without treatment. Both of the light and dark squares with smaller white circles inside them, denote the subjects follow the optimal treatment regime. The solid vertical line denote the estimated quantile treatment regime estimated by our framework, and the dashed vertical line denotes the true optimal treatment estimated by Monte Carlo experiment with sample size $n = 10^5$.

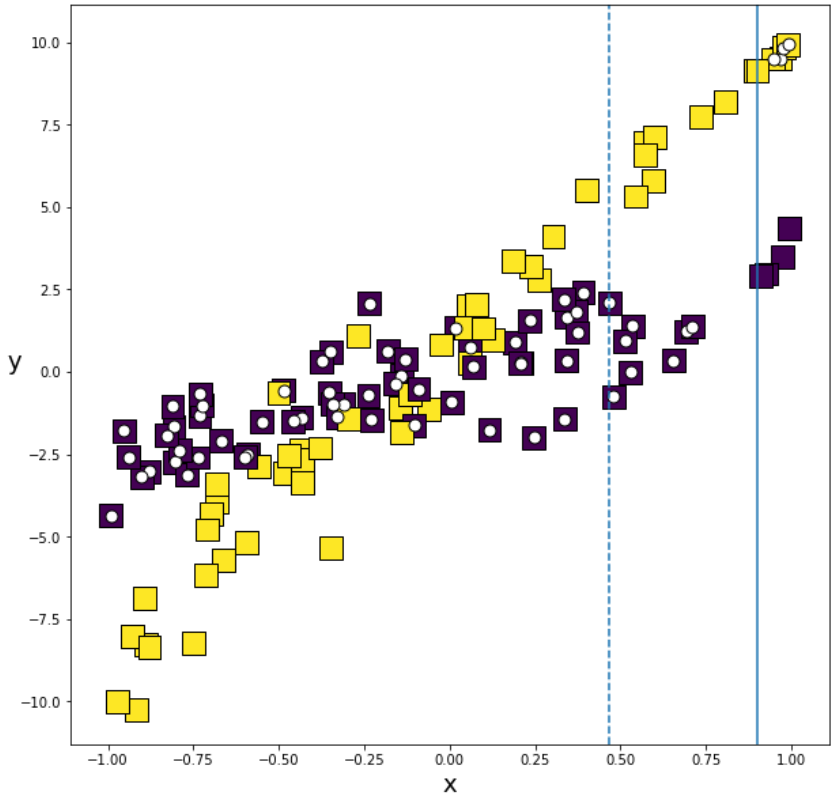

Figure 3.2: This is the simulation result for the first example when $\tau = 0.25$. In this figure, the light squares denote the subjects with treatment, while the dark squares are the ones without treatment. Both of the light and dark squares with smaller white circles inside them, denote the subjects follow the optimal treatment regime. The solid vertical line denote the estimated quantile treatment regime using the framework proposed by Wang et al., and the dashed vertical line denotes the true optimal treatment estimated by Monte Carlo experiment with sample size $n = 10^5$

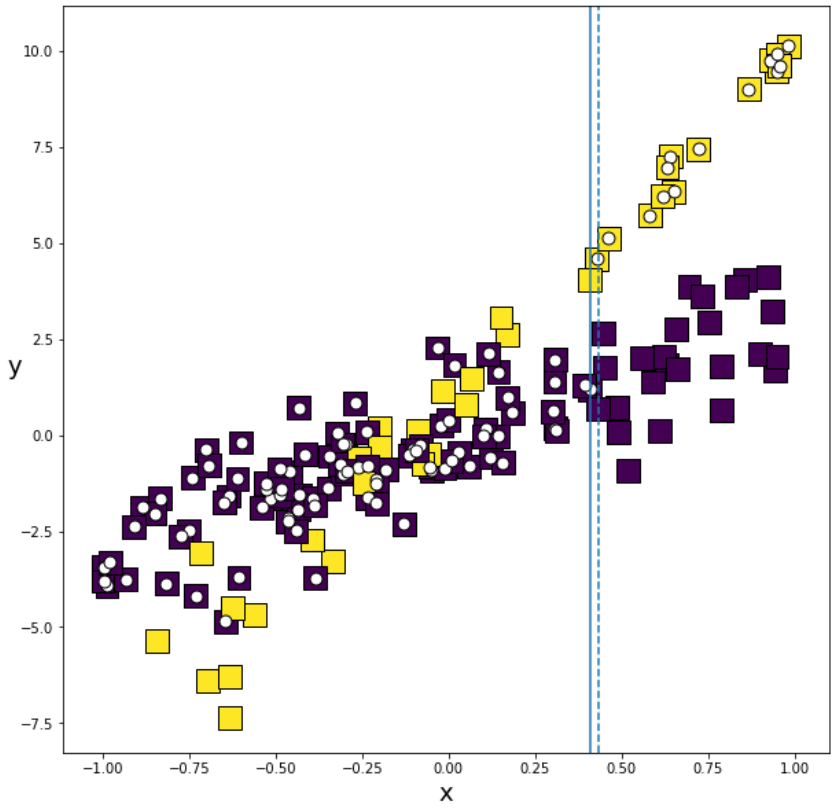

Figure 3.3: This is the simulation result for the first example considering as observed data when $\tau = 0.25$. In this figure, the light squares denote the subjects with treatment, while the dark squares are the ones without treatment. Both of the light and dark squares with smaller white circles inside them, denote the subjects follow the optimal treatment regime. The solid vertical line denote the estimated quantile treatment regime estimated by our framework, and the dashed vertical line denotes the true optimal treatment estimated by Monte Carlo experiment with sample size $n = 10^5$.

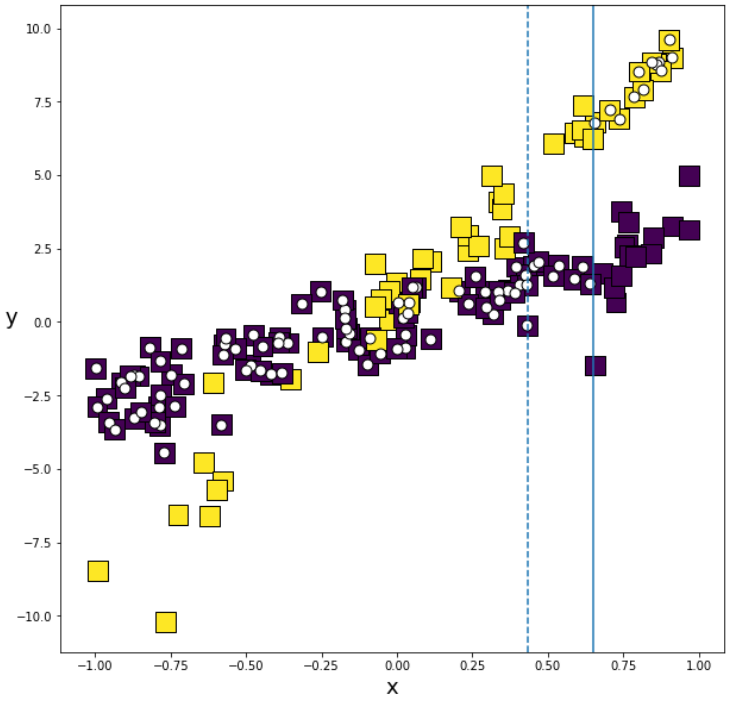

Figure 3.4: This is the simulation result for the first example when $\tau = 0.25$. In this figure, the light squares denote the subjects with treatment, while the dark squares are the ones without treatment. Both of the light and dark squares with smaller white circles inside them, denote the subjects follow the optimal treatment regime. The solid vertical line denote the estimated quantile treatment regime estimated by our framework, and the dashed vertical line denotes the true optimal treatment estimated by Monte Carlo experiment with sample size $n = 10^5$

consider $\mu(X) = 4 + 3x_1$, $\delta(X) = 5x_1 + 2$ and $\omega(X, A) = 1 - ax_1$, and the treatment assignment A follows a Bernoulli distribution with probability of 0.5. However, instead of using a normal distributed random error, in this example, ϵ follows $\Gamma(shape = x + 1, scale = 1)$.The quantiles considered in this example are still considered to take $\tau = 0.1$ and $\tau = 0.25$. The class of treatment regime is $\mathbb{I}(\beta_0 + \beta_1 x > 0)$. The sample size considered in this example is $n = 500$ and $n = 1000$. The same with *Example 1*, simulation is repeated 500 times for both of the methods.

In this example, the common effect can still be estimated by linear regression. However, note that the random error in this example follows Γ distribution with shape $x + 1$ and scale 1. This implies the estimation of the common effect is biased and simply using linear regression may not be appropriated. However, from the simulation where the results is reported in Table 3.5, it can be seen that even in this case, the accuracy of the prediction still increases, and meanwhile, the variances of the parameters are also reduced. Figure 3.5 and Figure 3.6 also provide visualized comparisons of the estimated optimal treatment regime without and with common effects in the outcomes.

Example 3. In the third example, we consider the scenario with two covariates (x_1, x_2). $\mu(X) = 1 + 3x_1 + 2x_2 + 5x_2^2$, $\delta(X) = 6x_1 + 5x_2 + 1$ and $\omega(X, A) = 1 - 0.8ax_1 - 0.8ax_2$. The random error ϵ follows $Normal(0, 1)$. The treatment assignment A follows a Bernoulli distribution with probability of 0.5. The quantile criterion is considered to take $\tau = 0.25$. The class of treatment regime is assumed as $\mathbb{I}(\beta_0 + \beta_1 x_1 + \beta_2 x_2 > 0)$.

It is worth to noting that the common effect in this example contains a quadratic term $5x_2^2$, thus a higher order of polynomial regression can estimate

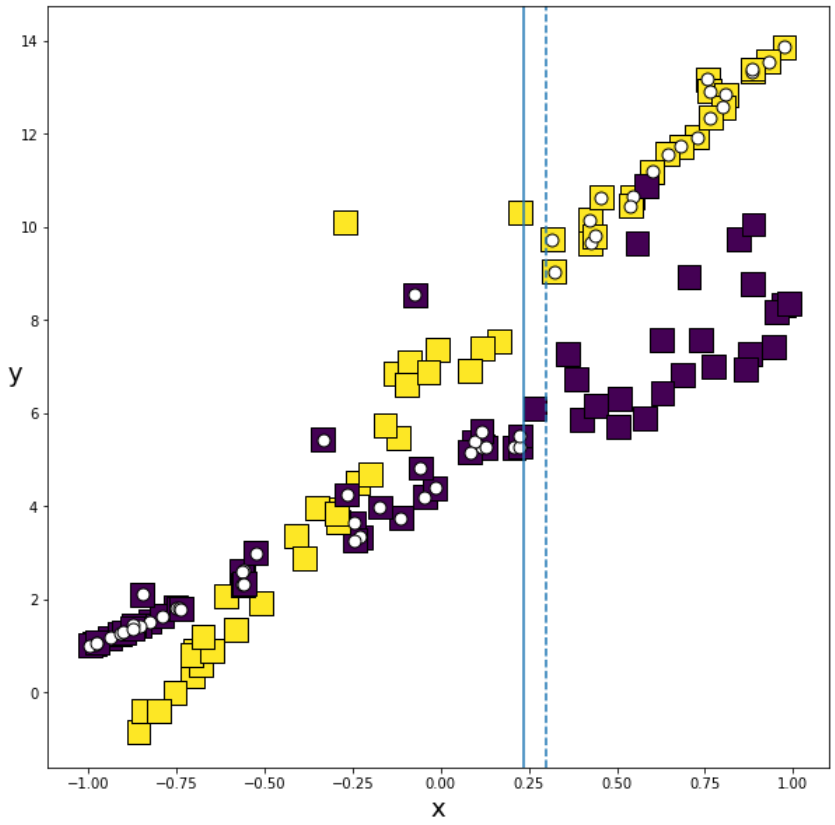

Figure 3.5: This is the simulation result for the second example when $\tau = 0.25$. In this figure, the light squares denote the subjects with treatment, while the dark squares are the ones without treatment. Both of the light and dark squares with smaller white circles inside them, denote the subjects follow the optimal treatment regime. The solid vertical line denote the estimated quantile treatment regime estimated by our framework, and the dashed vertical line denotes the true optimal treatment estimated by Monte Carlo experiment with sample size $n = 10^5$

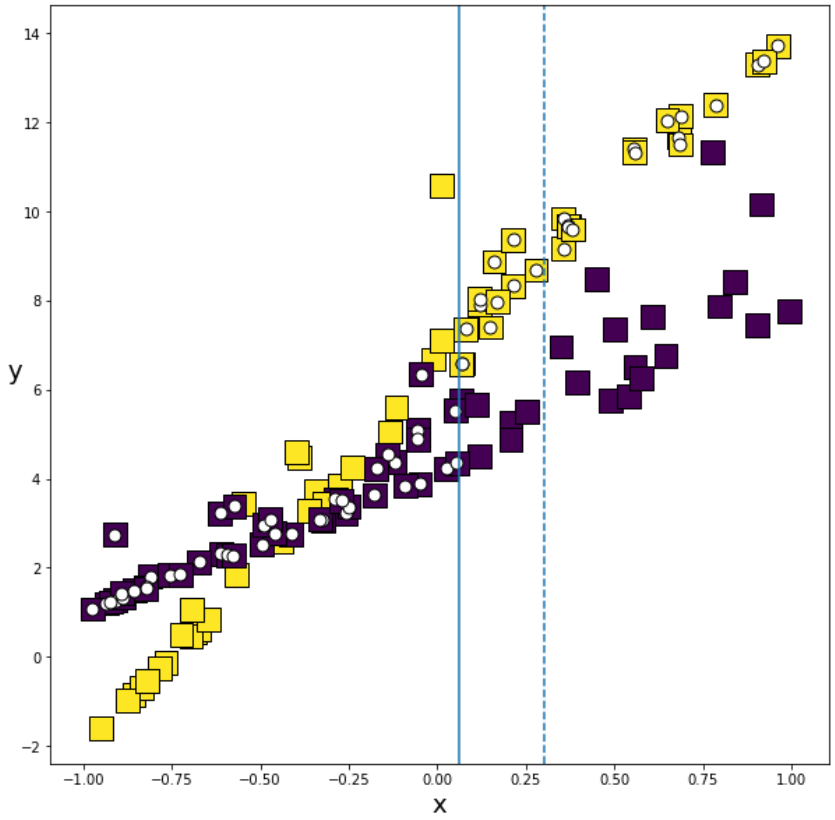

Figure 3.6: This is the simulation result for the second example when $\tau = 0.25$. In this figure, the light squares denote the subjects with treatment, while the dark squares are the ones without treatment. Both of the light and dark squares with smaller white circles inside them, denote the subjects follow the optimal treatment regime. The solid vertical line denote the estimated quantile treatment regime estimated by our framework, and the dashed vertical line denotes the true optimal treatment estimated by Monte Carlo experiment with sample size $n = 10^5$.

Table 3.4: Summary results for the second example when $\tau = 0.25$. RCE stands for *removing common effect* and WCE stands for *with common effect*. The numbers in the parenthesis are standard deviations. The last column is the accuracy rate defined by $\mathbb{P}_n^*$ using the true value in the last row.

		β_0	β_1	accuracy rate
$n = 500$	RCE	-0.229 (0.087)	0.968(0.061)	0.968
	WCE	-0.013(0.321)	0.216(0.921)	0.88
$n = 1000$	RCE	-0.248 (0.048)	0.967(0.011)	0.983
	WCE	-0.027(0.344)	0.121(0.929)	0.967
true value		-0.287	0.958	

the common effect more accurately than linear regression does. However, to illustrate the benefit of removing the common effect, we still use linear regression to estimate the common effect for simplicity. Table 3.5 reports the estimated β values and the accuracy rate of *Example 3*. It can be seen that even though the common effect is estimated by linear regression, the standard deviations decrease when we remove the common effect from the observed outcomes. Further, the accuracy rate increases. These two results can be easily seen in Figure 3.7. Further, Figure 3.8 shows the comparison of the two methodologies.

Example 4. In the forth example, we generate random data from $\mu(X) = 1+3x_1+2x_2^3$, $\delta(X) = 1+6x_1+5x_2+x_3+3x_4$ and $\omega(X, A) = 1-0.8ax_1-0.8ax_2$. The treatment assignment A follows a binomial distribution with probability of 0.5. The random error ϵ follows $Normal(0, 1)$. The quantile considered in this example is $\tau = 0.25$. The class of treatment regime is assumed as $\mathbb{I}(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 > 0)$.

In this example, the common effect is modelled by two of the four co-

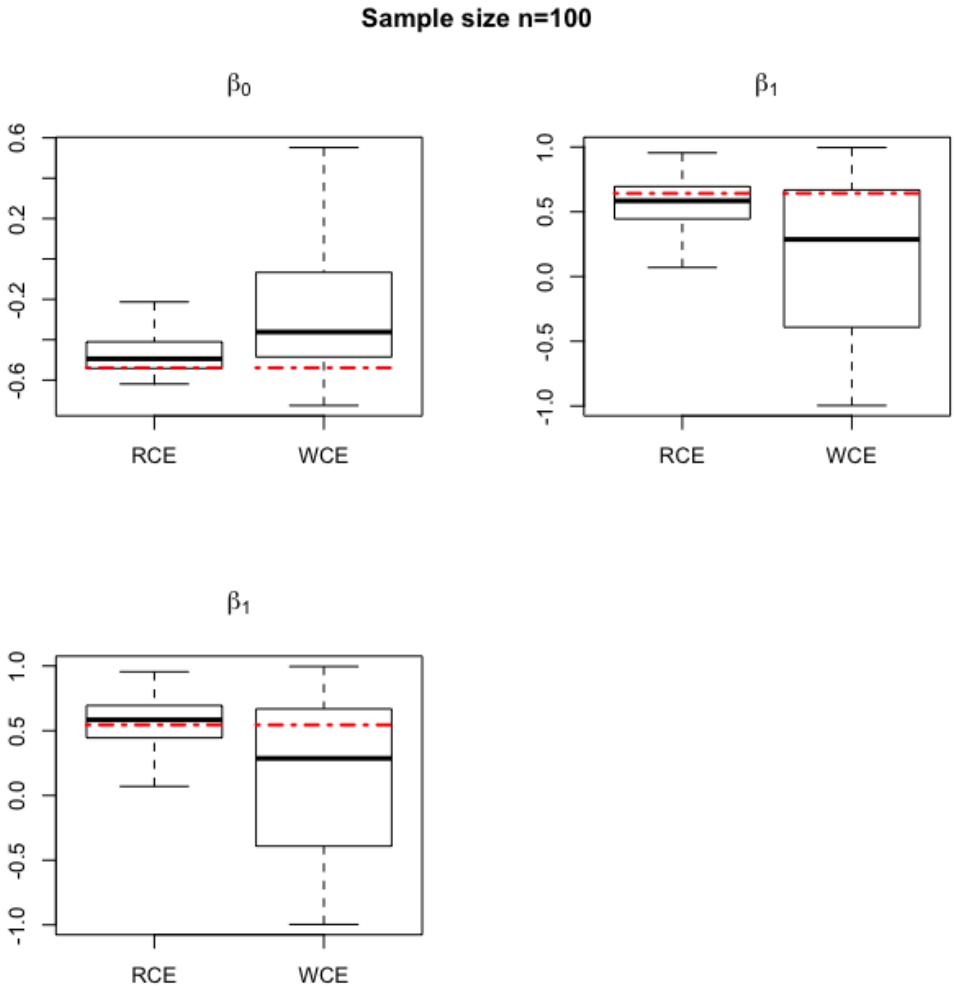

Figure 3.7: This is the box plot for the three parameters in the third example when $\tau = 0.25$. In this figure, dot-dash line denotes the true optimal treatment estimated by Monte Carlo experiment with sample size $n = 10^5$.

Table 3.5: Summary results for the third example when $\tau = 0.25$. RCE stands for *removing common effect* and WCE stands for *with common effect*. The numbers in the parenthesis are standard deviations. The last column is the accuracy rate defined by $\mathbb{P}_n^*$ using the true value in the last row.

		β_0	β_1	β_2	accuracy rate
$n = 100$	RCE	-0.463 (0.106)	0.558(0.226)	0.601 (0.226)	0.969
	WCE	-0.267(0.270)	0.119(0.581)	0.526 (0.476)	0.845
$n = 500$	RCE	-0.513 (0.056)	0.599(0.122)	0.586 (0.126)	0.986
	WCE	-0.422(0.182)	0.328(0.488)	0.607 (0.272)	0.943
$n = 1000$	RCE	-0.528 (0.03)	0.634(0.074)	0.552(0.082)	0.995
	WCE	-0.482(0.133)	0.471(0.375)	0.580 (0.224)	0.972
true value		-0.539	0.642	0.545	

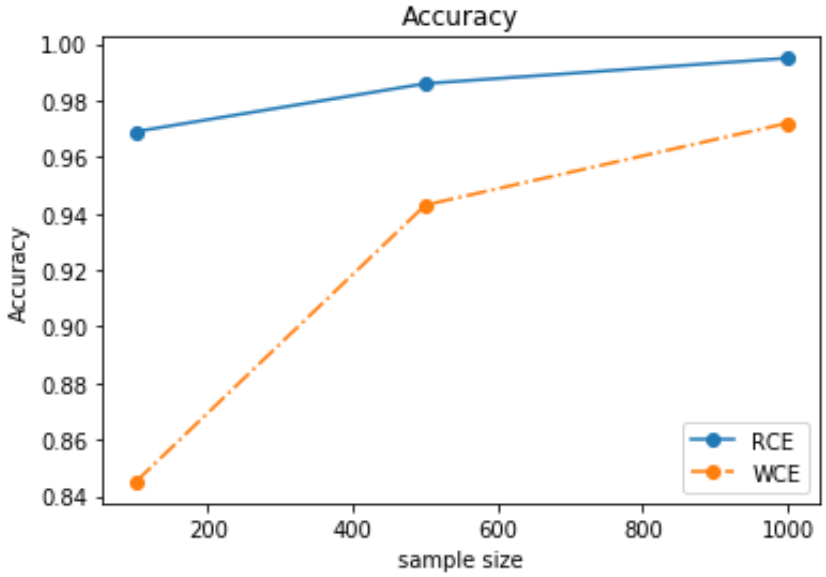

Figure 3.8: This figure shows the comparison of the three methods in the forth example with three sample size: 100, 500, 1000. The dot-dash line represents the method *with common effect*. The solid line represents the method in which the common effect is removed.

variates, (x_1, x_2). Therefore, we can apply LASSO method to perform the feature selection and then estimate the common effect more accurately. Table 3.6 reports the estimated result for this example. Here, WCE is short for *With Common Effect*, RL is short for *Remove common effect using LASSO* and RWL is for *Remove With Linear regression*. Note that there is a tuning parameter in the linear regression with LASSO (RL), this parameter was estimated by 10-fold cross-validation.

From Table 3.6 and Figure 3.9, it can be seen that the two methods (RL and RWL) have smaller variances and higher accuracy rates, comparing to the method WCE. This conclusion is the same with the one obtained in previous examples. Further, since applying LASSO can perform a better estimation on the common effect, then comparing the two accuracy rates between RL and RWL, we can see that better estimating the common effect can imply better estimating of quantile optimal treatment. Figure 3.10 shows the accuracy comparison of these three methods.

These four examples use four cases to compare the performances of the two methodologies. From these four cases, it can be seen that, removing the common effect can significantly reduce the variances of the parameters indexing the quantile-optimal treatment regime and increase the accuracy of the prediction.

3.2 Real Data Analysis

Similarly to Wang's work, we will apply our proposed new quantile OTR method on the data set ***ACTG175*** which can be found in the R package ***sp-***

Table 3.6: Summary results for the forth example when $\tau = 0.25$. RCE stands for *removing common effect* and WCE stands for *with common effect*. The numbers in the parenthesis are standard deviations. The last column is the accuracy rate defined by $\mathbb{P}_n^*$ using the true value in the last row.

	Method	β_0	β_1	β_2	β_3	β_4	accuracy rate
$n = 100$	RL	-0.522	0.529	0.431	0.057	0.245	0.965
		(0.163)	(0.152)	(0.176)	(0.248)	(0.238)	
	RWL	-0.529	0.523	0.437	0.071	0.245	0.961
		(0.162)	(0.152)	(0.179)	(0.242)	(0.226)	
	WCE	-0.242	-0.048	0.244	0.061	0.226	0.875
		(0.336)	(0.446)	(0.413)	(0.414)	(0.416)	
$n = 500$	RL	-0.617	0.548	0.465	0.07	0.273	0.985
		(0.03)	(0.053)	(0.061)	(0.089)	(0.082)	
	RWL	-0.618	0.554	0.458	0.072	0.272	0.984
		(0.031)	(0.052)	(0.064)	(0.09)	(0.075)	
	WCE	-0.479	-0.075	0.331	0.098	0.389	0.902
		(0.249)	(0.398)	(0.327)	(0.309)	(0.269)	
$n = 1000$	RL	-0.626	0.551	0.459	0.081	0.273	0.992
		(0.021)	(0.042)	(0.048)	(0.065)	(0.059)	
	RWL	-0.626	0.549	0.462	0.078	0.275	0.990
		(0.022)	(0.038)	(0.044)	(0.063)	(0.056)	
	WCE	-0.567	-0.062	0.399	0.100	0.409	0.930
		(0.194)	(0.384)	(0.258)	(0.214)	(0.205)	
true value		-0.644	0.517	0.485	0.064	0.279	

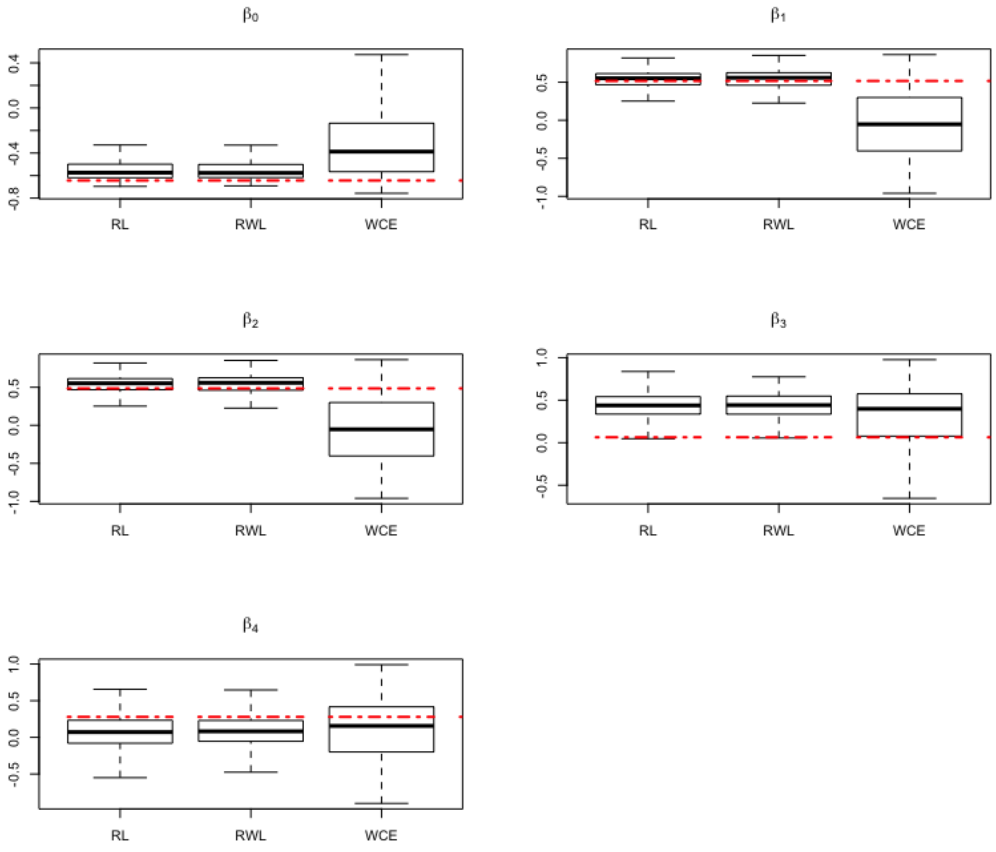

Figure 3.9: This is the box plot for the three parameters in the third example when $\tau = 0.25$. In this figure, dot-dash lines denote the true optimal treatment estimated by Monte Carlo experiment with sample size $n = 10^5$.

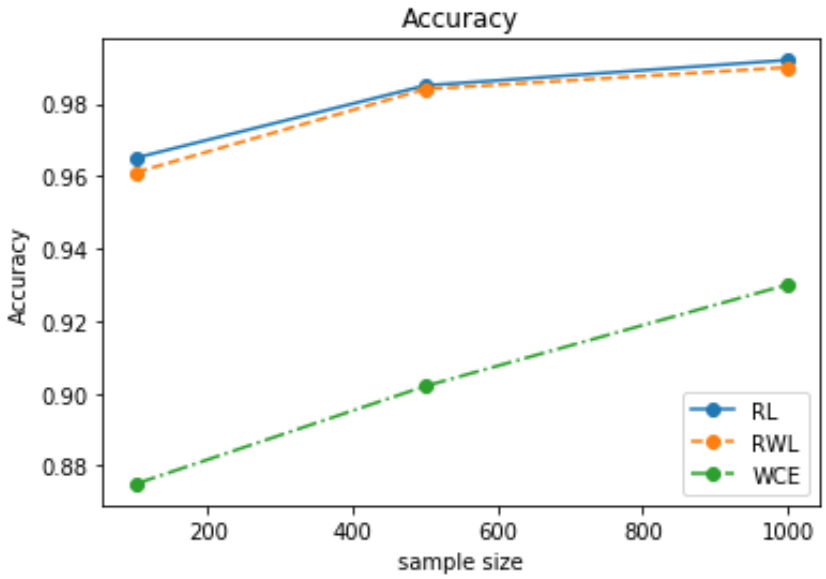

Figure 3.10: This figure shows the comparison of the three methods in the forth example with three sample size: 100, 500, 1000. The dot-dash line represents the method *with common effect*. The dash line represents the method which removes the common effect without using LASSO. The solid line represents the method in which the common effect is removed and also using LASSO for feature selection.

eff2trial. ***ACTG175*** data set records a randomized clinical trial to compare two single therapies: *zidovudine (ZI)*, *didanosine (DI)*, with two combination therapies: *zidovudine with didanosine (ZID)*, and *zidovudine with zalcitabine (ZIZ)*. The goal of this trial is to evaluate the single therapies and the combination therapy to check that which performs better to the patients with CD4 T cell between 200 and $500/mm^3$ [9], as the count of CD4 T cell is good indicator of the disease progression for the patients who with *HIV* infection.

This data set records 2139 *HIV-infected* patients with 27 variables such as *age*,*weight* and so on. The response variable can be considered either the CD4 count at week 20 (denoted as CD420) or the CD4 count at week 96 (denoted as CD496). Since there are missing data in CD496, we choose CD420 as our response variable. Figure 3.11 and Figure 3.12 indicate that the distributions of the response CD420 is asymmetric and right-skewed for both of the treatments. Moreover, all the patients participate have taken *ZI* before they entered this trial. Therefore, we compare two treatments: *DI* (with sample size 561) and *ZID* (with sample size 522) to find the treatment which performs better. Further, the treatment assignment is set to 0 for those patient only took *DI*, and is set to 1 if the patient took *ZID*.

The covariates initially considered here are: baseline age (X_1), and baseline count of CD4 cells (X_2) at the starting week of the trial. These two covariates were chosen based on the fact that the age [13] has impacts on the count of the CD4 cells and the count of CD4 at the week 20 is also related on the count of the cell at the beginning of the trial. The treatment regime considered in this case has the form of $I(\eta_0+\eta_1 X_1+\eta_2 X_2 > 0)$, where $\eta_i \in (-1, 1)$, $i = 0, 1, 2$ and $||\eta||_2 = 1$. If $\eta_0+\eta_1 X_1+\eta_2 X_2 \neq 0$, then the patient will take *DI*, otherwise, the

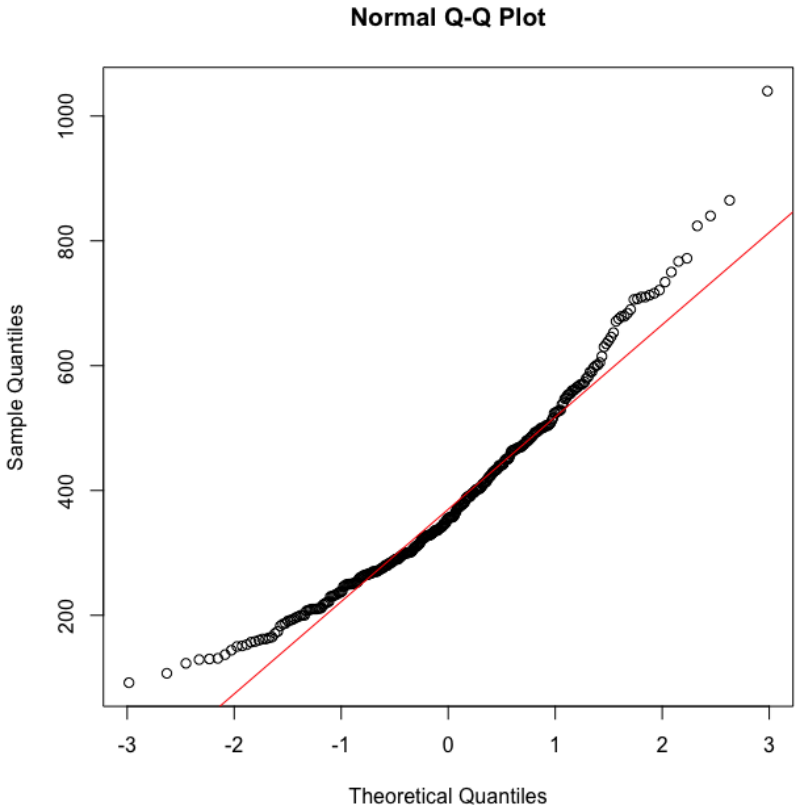

Figure 3.11: This is the Q-Q plot of the count of CD4 T cell with *DI* at week 20.

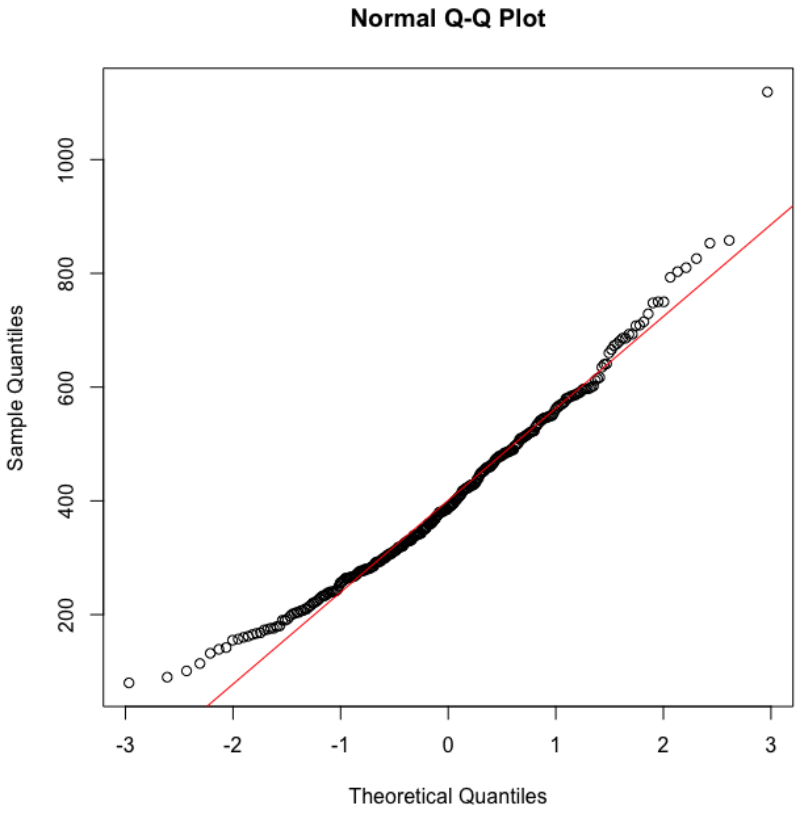

Figure 3.12: This is the Q-Q plot of the count of CD4 T cell with *ZID* at week 20.

patient will be assigned to *ZID*. Further, since the treatments are randomly assigned, it can be calculated that the propensity score in this case study is 0.482, which can also be considered as constant.

Note that the large count of CD4 cell at week 20 of a patient may either comes from an effective treatment, or comes from a large baseline count at week 0. This provides us an example of the reason to remove the common effect from the response. To estimate the common effect, we focus on the 561 samples with *DI*, and apply linear regression with L_1 penalty (*Lasso*). The common effect can be estimated by:

$$\widehat{\text{count}}_{week20} = 186.91 - 20.402\text{age} + 659.59\text{count}_{week0}.$$

From the above model of the common effect, it can be seen that the coefficient of the *Age* is negative. This implies that age has a negative impact on the count of CD4 cell at week 20. This result is consistent with the one shown by Means et al. [13], where they stated that the total immune reconstitution potential and the reconstitution rate are decreasing with age. This would lead to the fact that older patients' CD4 cell recovery is slower than the younger patients, and then further impact the total count of CD4 cells.

Removing the common effect, we can estimate the optimal treatment regime. Table 3.7 reports the comparison of method RCE (Remove Common Effect) and WCE (With Common Effect) by providing the estimated value function on 9 different levels of quantiles. It can be observed that in this data set, on most of values are the same. This may due to a relatively small sample size. However, it can also be seen that for some cases when $\tau = 0.3$ and on the me-

dian level ($\tau = 0.5$), a better treatment regime can be obtained by removing the common effect.

Table 3.7: Summary data analysis results for ACTG175

τ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
RCE $\hat{Q}_\tau$	221	277	**315**	349	**393**	431	484	540	608
WCE $\hat{Q}_\tau$	221	277	313	349	390	431	484	540	608

Table 3.8 and Table 3.9 report the comparisons among the two methodologies of estimating quantile optimal treatment regimes (with and without common effect), and the mean optimal treatment regime obtained by the method proposed by Zhang [21]. From these two tables, it can be seen that the mean-optimal treatment regime estimated by removing the common effects has the largest mean outcome 427.080. This means if the goal is the improve the overall outcomes of the target population, then choose the mean optimal treatment regime is the best choice. However, if $\tau = 0.3$ and $\tau = 0.5$ level quantiles of the outcomes are of interest, which means in this case that the goal of interest locates on the lower tail the the outcome distribution, then the quantile-optimal treatment regimes perform better than the regime relies on the mean. Further, the quantile-optimal treatment regime derived by removing the common effect provides a better regime than the regime estimated by framework proposed by Wang et al.

Table 3.8: Summary results on four different methods for ACTG175 when $\tau = 0.3$

	η_0	η_1	η_2	$\hat{Q}_{0.3}$	$\hat{Q}_{mean}$
RCE ($\tau = 0.3$)	-0.319 (0.116)	0.468 (0.144)	0.824 (0.433)	**315**	394.98
WCE ($\tau = 0.3$)	-0.320 (0.148)	0.454 (0.250)	0.831 (0.401)	313	394.87
Mean	-0.110 (0.411)	0.770 (0.664)	-0.629 (0.540)	313	414.37
Mean (WCE)	-0.372(0.348)	0.918(0.496)	0.141(0.545)	313	**427.080**

Table 3.9: Summary results on four different methods for ACTG175 when $\tau = 0.5$

	η_0	η_1	η_2	$\hat{Q}_{0.5}$	$\hat{Q}_{mean}$
RCE ($\tau = 0.5$)	0.305 (0.309)	0.192 (0.428)	-0.934 (0.555)	**393**	400.680
WCE ($\tau = 0.5$)	0.415 (0.206)	-0.686 (0.298)	0.597 (0.584)	390	402.230
Mean	-0.110 (0.411)	0.770 (0.664)	-0.629 (0.540)	390	414.37
Mean(WCE)	-0.372(0.348)	0.918(0.496)	0.141(0.545)	391	**427.080**

Note that in the previous data analysis, we used a different subset data of ACTG175 from the one used by Wang et al., and also, we considered a different model using age instead of using weighted. Therefore, to further compare with the result obtained by Wang et al., we last use the same data and features to do the data analysis.

To be consistent with the method used by Wang et al., we also rescale the weight and the count of CD4 cells at the starting week into $[-1, 1]$. Then using linear regression with LASSO to estimate the common effect, we can derive

$$\widehat{\text{count}}_{week96} = 129.322 + 72.057\text{weight} + 611.983\text{count}_{week0},$$

as the estimated common effect. Subtracting this from the observed counts of CD4 cells at week 96 and applying our proposed residual quantile-optimal treatment regime, we can have table 3.10 and table 3.11 which reports the comparison of the results from our proposed framework and the one from the framework proposed by Wang et al.. (The results from Wang et al. are used the same ones from their paper [19].) From the two tables, it can be observed that, when $\tau = 0.25$, the framework proposed by Wang et al. yields a higher value 261. When $\tau = 0.5$, our proposed framework yields 361, which is higher than the one derived by Wang et al. This implies that our framework

is comparable to the one proposed by Wang et al..

Table 3.10: Summary results on three different methods for ACTG175 when $\tau = 0.25$

	β_0	β_1	β_2	$\hat{Q}_{0.25}$
RCE ($\tau = 0.25$)	0.543 (0.241)	-0.346 (0.473)	-0.765 (0.382)	259
WCE ($\tau = 0.25$)	-0.231(0.303)	0.533(0.533)	0.814 (0.576)	**261**

Table 3.11: Summary results on three different methods for ACTG175 when $\tau = 0.5$

	β_0	β_1	β_2	$\hat{Q}_{0.5}$
RCE ($\tau = 0.5$)	0.552 (0.341)	-0.425(0.625)	-0.718 (0.420)	**361**
WCE ($\tau = 0.5$)	0.594 (0.399)	-0.564 (0.677)	-0.574 (0.574)	360

Chapter 4

Conclusion

A lot of researches have been done on finding the optimal treatment regimes which maximize the mean of the potential outcomes, but only a few have been done on finding the regimes maximizing certain quantiles of the outcomes. However, in many applications such as economics, it is of great interest to find the quantile optimal treatment regimes. Wang et al. are the ones who firstly engaged in this study and provided a good framework of finding the optimal treatment regimes based on quantile. Motivated by Zhou's work on the residual weighted learning on the mean-optimal treatment regimes, we further improve Wang's framework by removing the common effect from the response.

It has been shown that after removing the common effect, the estimated quantile optimal treatment regime is much closer to the true ones. Moreover, removing the common effect can greatly reduce the variance of the coefficients of the estimated treatment regime. It is worth noting that both Wang's framework and our proposed framework do not rely on a specified outcome regression

model. This leads to the difficulties of using feature selection techniques. However, we can apply feature selections when estimating the common effect. We can choose those features have more impacts on the response as the covariates and then calculate the quantile optimal treatment regimes.

Even though this framework made some progress in the study of finding optimal treatment regimes, there are still several aspects needed to be improve. First, this framework successfully reduce the variances of the estimated parameters indexing quantile-treatment regime, however, this framework only works under some conditions, where the treatment effect must be monotone and removing the common effect will not change the order of the original outcomes. This condition can not be easily checked in real data analysis. Therefore, it is important to analyze this deeply to find out a solution to alleviate this problem. Second, unlike in mean-optimal treatment regimes where the true values of coefficients of treatment regimes can be easily obtained, it is much more complicated to obtain the true values for the quantile-optimal treatment regimes to compare the estimated result. We, following the research of Wang et al. [19], used $n = 10^5$ Monte Carlo experiment to approximate the true values of coefficients in our study, but this method is not a perfect one due to the fact that the 'true' values still have some bias. This can be a problem especially when we want to compare the result and try to obtain a more accurate one. Therefore, further mathematical analysis needs to be done on this aspect to obtain the true values of the coefficients. Third, our proposed framework is also a two-step one. Thus, more theoretical work need to be developed to directly derive the quantile optimal treatment regimes. Forth, there has been some work on finding the mean-optimal treatment regimes on the data with

randomly censored response variables. To the best of our knowledge, no works have been done on finding the quantile-optimal treatment regimes on these censored data. Fifth, feature selection technologies are very important in machine learning, which can help to find the most important features and provide more accurate results. In the treatment regime studies, it should be noticed that the features used to estimate the treatment regimes may be different from the ones in common effects, and could be a subset of all the features. However, we have not considered this case in our study. Following the studies of using genetic algorithm to perform feature selections [20, 11, 15], we can further use genetic algorithm to find the most important features in the treatment effect and then further improve the performance of the quantile-optimal treatment regime. All these three are possible works for future studies.

Bibliography

[1] Weihua Cao, Anastasios A Tsiatis, and Marie Davidian. Improving efficiency and robustness of the doubly robust estimator for a population mean with incomplete data. *Biometrika*, 96(3):723–734, 2009.

[2] Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine learning*, 20(3):273–297, 1995.

[3] Griselda Deelstra, Jan Dhaene, and Michele Vanmaele. An overview of comonotonicity and its applications in finance and insurance. In *Advanced mathematical methods for finance*, pages 155–179. Springer, 2011.

[4] Jan Dhaene, Michel Denuit, Marc J Goovaerts, Rob Kaas, and David Vyncke. The concept of comonotonicity in actuarial science and finance: theory. *Insurance: Mathematics and Economics*, 31(1):3–33, 2002.

[5] Jan Dhaene, Steven Vanduffel, Marc J Goovaerts, Rob Kaas, Qihe Tang, and David Vyncke. Risk measures and comonotonicity: a review. *Stochastic models*, 22(4):573–606, 2006.

[6] Jan Dhaene, Steven Vanduffel, Marc J Goovaerts, Rob Kaas, and David Vyncke. Comonotonic approximations for optimal portfolio selection problems. *Journal of Risk and Insurance*, 72(2):253–300, 2005.

[7] Caiyun Fan, Wenbin Lu, Rui Song, and Yong Zhou. Concordance-assisted learning for estimating optimal individualized treatment regimes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 79(5):1565–1582, 2017.

[8] Jianqing Fan and Runze Li. Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American statistical Association*, 96(456):1348–1360, 2001.

[9] Scott M Hammer, David A Katzenstein, Michael D Hughes, Holly Gundacker, Robert T Schooley, Richard H Haubrich, W Keith Henry,

Michael M Lederman, John P Phair, Manette Niu, et al. A trial comparing nucleoside monotherapy with combination therapy in hiv-infected adults with cd4 cell counts from 200 to 500 per cubic millimeter. *New England Journal of Medicine*, 335(15):1081–1090, 1996.

[10] Roger Koenker and Kevin F Hallock. Quantile regression. *Journal of economic perspectives*, 15(4):143–156, 2001.

[11] Riccardo Leardi. Genetic algorithms in feature selection. In *Genetic algorithms in molecular modeling*, pages 67–86. Elsevier, 1996.

[12] Kristin A Linn, Eric B Laber, and Leonard A Stefanski. Interactive q-learning for quantiles. *Journal of the American Statistical Association*, 112(518):638–649, 2017.

[13] Arianna R Means, Kathryn A Risher, Eva L Ujeneza, Innocent Maposa, Joseph Nondi, and Steven E Bellan. Impact of age and sex on cd4+ cell count trajectories following treatment initiation: an analysis of the tanzanian hiv treatment database. *PloS one*, 11(10):e0164148, 2016.

[14] Erica EM Moodie, Nema Dean, and Yue Ru Sun. Q-learning: Flexible learning about useful utilities. *Statistics in Biosciences*, 6(2):223–243, 2014.

[15] Il-Seok Oh, Jin-Seon Lee, and Byung-Ro Moon. Hybrid genetic algorithms for feature selection. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (11):1424–1437, 2004.

[16] Min Qian and Susan A Murphy. Performance guarantees for individualized treatment rules. *Annals of statistics*, 39(2):1180, 2011.

[17] James M Robins, Andrea Rotnitzky, and Lue Ping Zhao. Estimation of regression coefficients when some regressors are not always observed. *Journal of the American statistical Association*, 89(427):846–866, 1994.

[18] Robert Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 267–288, 1996.

[19] Lan Wang, Yu Zhou, Rui Song, and Ben Sherwood. Quantile-optimal treatment regimes. *Journal of the American Statistical Association*, (just-accepted), 2017.

[20] Jihoon Yang and Vasant Honavar. Feature subset selection using a genetic algorithm. In *Feature extraction, construction and selection*, pages 117–136. Springer, 1998.

[21] Baqun Zhang, Anastasios A Tsiatis, Eric B Laber, and Marie Davidian. A robust method for estimating optimal treatment regimes. *Biometrics*, 68(4):1010–1018, 2012.

[22] Ying-Qi Zhao, Donglin Zeng, Eric B Laber, and Michael R Kosorok. New statistical learning methods for estimating optimal dynamic treatment regimes. *Journal of the American Statistical Association*, 110(510):583–598, 2015.

[23] Ying-Qi Zhao, Donglin Zeng, Eric B Laber, Rui Song, Ming Yuan, and Michael Rene Kosorok. Doubly robust learning for estimating individualized treatment with censored data. *Biometrika*, 102(1):151–168, 2014.

[24] Yingqi Zhao, Donglin Zeng, A John Rush, and Michael R Kosorok. Estimating individualized treatment rules using outcome weighted learning. *Journal of the American Statistical Association*, 107(499):1106–1118, 2012.

[25] Xin Zhou, Nicole Mayer-Hamblett, Umer Khan, and Michael R Kosorok. Residual weighted learning for estimating individualized treatment rules. *Journal of the American Statistical Association*, 112(517):169–187, 2017.