

Doubly-robust Q-estimation in observational studies with high-dimensional covariates

Hyobeen Lee^a, Yeji Kim^a, Hyungjun Cho^a, Sangbum Choi^{1,a}

^aDepartment of Statistics, Korea University

Abstract

Dynamic treatment regimes (DTRs) are decision-making rules designed to provide personalized treatment to individuals in multi-stage randomized trials. Unlike classical methods, in which all individuals are prescribed the same type of treatment, DTRs prescribe patient-tailored treatments which take into account individual characteristics that may change over time. The Q-learning method, one of regression-based algorithms to figure out optimal treatment rules, becomes more popular as it can be easily implemented. However, the performance of the Q-learning algorithm heavily relies on the correct specification of the Q-function for response, especially in observational studies. In this article, we examine a number of double-robust weighted least-squares estimating methods for Q-learning in high-dimensional settings, where treatment models for propensity score and penalization for sparse estimation are also investigated. We further consider flexible ensemble machine learning methods for the treatment model to achieve double-robustness, so that optimal decision rule can be correctly estimated as long as at least one of the outcome model or treatment model is correct. Extensive simulation studies show that the proposed methods work well with practical sample sizes. The practical utility of the proposed methods is proven with real data example.

Keywords: double robustness, precision medicine, propensity score, Q-learning, super learner, variable selection

1. 서론

유전적 배경, 인구 통계 및 임상적 특성에 따라 치료 체계를 개별화해야 하는 현대 시대에서 개인 맞춤형의료 (personalized medicine) 혹은 정밀의학(precision medicine)의 중요성이 날로 커지고 있다. 전통적으로 임상 의사결정은 의학 연구의 표준(gold-standard)이라 볼 수 있는 무작위 배정 임상시험(randomized controlled trial)에 기반하여 결정된다. 대부분의 임상 연구의 주요 목표는 전체 환자군에서 주어진 질병에 대한 최상의 치료법을 찾아내는 것이고, 이때 개별 환자의 이질성은 오히려 연구의 방해 요소로 간주되곤 하였다. 그러나 개인 맞춤형의료를 구현하기 위해서는 이러한 개별 환자의 특성과 치료 방법 간의 상관성을 확인하는 것이 중요하며, 이를 통해 증거 기반 의학의 맥락에서 치료를 개별화할 필요성이 제기되고 있다 (Sørensen, 1996; Longford와 Nelder, 1999; Kravitz 등, 2004). 이러한 배경에서 데이터 중심의 접근 방식을 통해 의사결정을

This research was supported by Korea University Research Fund (K2008341) and Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2019R1F1A1052239, 2019R1A4A1028134).

¹ Corresponding author: Department of Statistics, Korea University, 145 Anam-ro, Seungbuk-gu, Seoul 02841, Korea, E-mail: choisang@korea.ac.kr

개선하여 적시에 필요한 환자에게 올바른 치료를 제공하는 정밀의학의 개념이 등장하였다. 쉽게 말해 정밀의학은 유전정보, 생활습관, 예전 치료결과 등 개인 건강 빅데이터를 토대로 최적화된 진단 및 치료를 적용하는 새로운 헬스케어 패러다임이다.

정밀의학의 구현을 위한 대표적인 방법 가운데 순차적으로 치료를 결정하는 방법인 동적 치료 요법(dynamic treatment regimes; DTRs)이 있다. 암이나 AIDS, 우울증 등과 같은 만성 질병은 단 한번의 치료로 완치가 사실상 불가능하고 개별 환자의 질병 진행상태와 이전 치료의 결과에 따라 후속 치료방법을 적절히 선택함으로써 질병을 관리해 나가야 한다. 동적 치료 요법은 치료의 각 단계에서 개별 환자에 대한 개인 건강과 진단 및 치료 정보를 활용하여 최상의 치료 결과를 위한 결정 규칙을 수립하는데 그 목적이 있다. 동적 치료 요법에 대한 추정 방법에는 값찾기(value-search)방법 및 회귀 기반(regression-based)방법 등이 있다. 전자의 접근 방식은 직접 관심 결과를 최적화하여 연속적으로 최적의 치료 요법을 결정한다. 반면에 회귀 기반 방법은 결과 모형의 설정을 통하여 간접적으로 최적의 동적 치료 요법을 결정한다. 최적의 동적 치료 요법을 추정하기 위해 개발된 여러 분석 및 추론 방법들이 있다. 다양한 모수적 베이지안 접근 방법 (Bellman, 1957; Bertsekas와 Tsitsiklis, 1996; Thall 등 2000; Arjas와 Saarela, 2010) 이 활용되었지만, 근래에는 준모수적 방법이 가장 널리 사용되고 있다. 준모수적 방법으로는 반복적인 후회 점수 최소화(iterative minimization of regrets) (Murphy, 2003), G-추정(G-estimation) (Robins, 2004), 기계 학습 방법(machine-learning methods) (Pineau 등, 2007), 후회-회귀(regret-regression) (Henderson 등, 2010), 주변 구조 모형(marginal structural models) (Robins 등, 2000 ; Hernan 등, 2006 ; van der Laan과 Petersen, 2007; Robins 등, 2008) 등이 있다. 회귀 모형 기반 Q-학습 (Murphy, 2005; Chakraborty, 2011)은 모수적 방법이나 G-추정과 같은 준모수적 방법과의 연결이 가능하고 구현이 쉽기 때문에 대중적으로 널리 사용되고 있다 (Chakraborty 등, 2010).

동적 치료 요법 연구에서 Q-학습 방법은 무작위 시험 데이터에 대해 주로 사용되지만, Q-함수 모형이 정확하게 설정된다면 Q-학습 방법을 관찰 연구 데이터에 사용할 수 있다. 그러나 일반적인 관찰 연구에서 Q-함수 모형을 정확하게 아는 것은 어려우므로 모형의 오설정(misspecification)에도 강건성을 갖는 추정량을 찾기 위한 다양한 방법들이 연구되어 왔으며, 대표적인 방법으로 동적 가중 최소제곱법(dynamic weighted ordinary least squares)이 있다 (Wallace와 Moodie, 2015; Lu 등, 2011; Geng 등, 2015). 이 방법은 구현이 용이한 Q-학습의 특성과 재귀적 이중강건성(double-robustness)을 갖는 G-추정 (Robin, 2004)의 특성을 기반으로 한 방법이다. 이 방법을 사용하기 위해 결과 모형(outcome model)과 치료 모형(treatment model) 중 하나 이상이 정확하게 설정되어야 하지만, 현실에서는 둘 다 알려지지 않은 경우가 많다. 치료 모형을 추정하기 위하여 일반적으로 해석력이 좋은 로지스틱 회귀 분석을 사용하여 왔지만, 설명변수 간 비선형성이 존재하는 경우 편차가 발생하게 되며 이는 결과 모형의 추정에도 영향을 미치게 된다. 이를 보완하기 위하여 비선형적 데이터를 다루기 용이한 회귀나무, 랜덤 포레스트 등의 기계학습 방법들을 이용하여 성향점수 모형(propensity score model)을 학습하는 방법들이 연구되어 왔다 (Lee 등, 2010; Westreich 등, 2010). 기계학습 방법으로 성향점수를 추정할 경우 개별 치료에 대한 공변량 효과를 해석하기 어렵다는 단점이 있으나, 이를 가중치로 활용할 경우 이중강건성 성질에 의해 결과 모형을 보다 안정적으로 추정할 수 있는 이점을 갖게 된다. 본 논문에서는 다양한 Q-학습 모형을 설정하고 성향점수 모형에 사용되는 기계학습 방법론에 따라 결과 모형에 대한 추정치가 어떻게 변화하는지 살펴보고자 한다.

한편 고차원 공변량의 경우 다변량 회귀 분석을 이용하여 결과 모형을 찾아낼 때, 주효과와 교호작용 효과 각각에 대한 변수 선택이 고려되어야 한다. 동적 치료 요법은 교호작용까지 고려해야 하므로 치료가 진행될 수록 데이터의 차원 또한 급속하게 커지게 되어 이를 통제하기 위한 변수 선택법이 필수적으로 고려되어야 한다. 만약 치료와 공변량 간 교호작용 효과보다 결과에 더 큰 영향을 미치는 주효과가 존재하는 경우, 교호작용 효과를 모형화할 때 변수 선택의 어려움이 커질 수 있다. 이를 해결하기 위해 결과 모형에 주효과와 교호작용 효과 모두를 포함하는 것 대신 교호작용 효과만 고려하여 추정하는 방법이 제안되었다 (Tian 등, 2014). 기존의 변수 선택 방법들은 변동성을 줄이고 추정량의 정확도를 높이는 예측변수(predictive variables)

를 찾는 것이 주요 목표였으나 최적의 치료 요법의 추정에는 개별 특성에 맞는 치료 전략을 찾는 지표인 규정 변수(prescriptive variables)를 찾아야 한다. Tian 등 (2014)의 논문에서는 간단한 변수 조작을 통해 주효과를 모형에 포함하지 않고 교호작용만으로도 치료 방법에 대한 불편추정치를 찾을 수 있음을 제안하였고 이를 고차원 자료에도 활용할 수 있음을 보였다. 본 연구에서는 개인의 특성에 맞춰 최적의 치료를 추천하는 것을 목표로 회귀 기반의 Q-학습의 맥락에서 고차원 관찰 데이터에 적용할 수 있는 이중강건한 추정법을 연구한다. Tian 등 (2014)의 방법을 Q-학습 모형의 추정에 활용함으로써 교호작용 효과만을 모형화하여 일치성을 갖는 모형의 추정치를 찾아내고 이를 개인 맞춤형의료의 구현에 활용하는 것이 연구의 목적이다.

논문의 나머지 부분은 다음과 같이 구성된다. 2절에서는 선행연구를 소개하고 3절에서는 제안된 방법론을 기술하였다. 4절과 5절에서는 각각 모의실험 및 실제 데이터 분석 결과를 제시하였다. 6절에서는 이 연구의 시사점을 살펴보고 향후 연구 방향을 논의한다. 3절의 정리에 대한 증명은 부록에 첨부되어있다.

2. 선행연구

2.1. 동적 치료 요법 프레임워크: 단일 결정 설정

개인의 특성을 고려한 최적의 치료를 찾기 위해 종단적 데이터에 대하여 최적의 동적 치료 요법을 추정한다. 이 절에서는 한 시점의 임상적 상태만을 고려하는 단일 결정 설정(single-decision setting)에서 동적 치료 요법을 추정하기 위한 기본 개념들을 설명하고자 한다. n 명의 피험자를 표본 추출한 임상시험 또는 관찰 연구를 고려한다. $X \in \mathcal{X}$ 는 공변량으로 치료 전에 관측된 환자의 개인적인 특성을 포함한다. 대조군과 치료군, 두 가지 치료 옵션이 있을 때 $A \in \mathcal{A}$ 는 관측된 치료(treatment)로 편의상 이분형 변수라 가정한다. 즉 $\mathcal{A} = \{0, 1\}$ 이며 여기서 $A = 1$ 과 $A = 0$ 은 각각 실험군과 대조군을 의미한다. $Y \in \mathbb{R}$ 는 생존기간, 반응률과 같은 결괏값으로 Y 가 큰 값을 가지는 것에 관심이 있다고 하자. (Y, A, X) 가 알려지지 않은 분포를 따를 때, $\{(Y_i, A_i, X_i), i = 1, \dots, n\}$ 를 관측하며 관측된 데이터가 독립항등분포를 따른다고 가정한다. 우리의 목표는 이러한 데이터를 사용하여 다음과 같은 최적의 치료 요법을 수립하는 것이다.

함수 $d : \mathcal{X} \rightarrow \mathcal{A}$ 를 결정 법칙으로 정의한다. 즉, 관측된 공변량 $x \in \mathcal{X}$ 에 대해 치료 $a \in \mathcal{A}$ 를 사상하는 함수이다. $Y(a)$ 를 치료 $A = a$ 에서의 잠재적 결과(potential outcome)라 하면 결정 법칙 d 에서의 잠재적 결과는 $Y(d) = \sum_{a \in \mathcal{A}} Y(a) \mathbf{I}_{\{d(X)=a\}}$ 로 정의된다 (Robins 등, 2000; Robins, 2004). 이때, 최적의 치료 요법 $d^{\text{opt}}(x)$ 는 (i) 모든 $x \in \mathcal{X}$ 에 대해 $d^{\text{opt}}(x) \in \mathcal{A}$, (ii) 어떤 $x \in \mathcal{X}$ 에서, $d(x) \in \mathcal{A}$ 를 만족하는 모든 d 에 대해 $E[Y(d^{\text{opt}})] \geq E[Y(d)]$ 의 두 조건을 만족해야 한다. 즉, $d^{\text{opt}}(x)$ 는 $X = x$ 값을 가지는 환자 집단에 대한 최대 평균 결과를 산출하는 치료이다. 최적의 치료 요법을 추정하기 위해서는 블립 함수(blip function) 혹은 후회 함수(regret function)를 정의할 필요가 있다. 어떤 환자를 대조군에 배정하였을 때 잠재적 결과를 $Y(a^{\text{ref}})$ 라 할 때 블립 함수는

$$\gamma(x, a) = E \left[Y(a) - Y(a^{\text{ref}}) | X = x \right],$$

로 정의한다. 즉, 실제로 받은 치료에 대한 잠재적 결과와 대조군에 배정되었을 경우 잠재적 결과 간 기댓값의 차이이다. 이와 관련 있는 개념인 후회 함수는 최적의 치료를 받은 경우 잠재적 결과와 환자가 실제로 받은 치료에 대한 잠재적 결과 간 기댓값의 차이이다. 최적의 치료를 받은 경우 잠재적 결과를 $Y(d^{\text{opt}})$ 라고 할 때 후회 함수는 다음과 같이 정의한다.

$$\mu(x, a) = E \left[Y(d^{\text{opt}}) - Y(a) | X = x \right].$$

따라서 블립 함수와 후회 함수는 $\mu(x, a) = \gamma(x, d^{\text{opt}}) - \gamma(x, a)$ 와 같은 관계를 갖는다. 환자 집단에서 최대 평균 결과를 산출하는 치료 요법은 블립 함수를 최대화하거나 후회 함수를 최소화하는 치료 요법임을 알 수 있다.

관찰 연구에서 인과 추론을 가능하게 하기 위해서 다음의 가정들이 필요하다.

- 교환가능성(exchangeability) : $\{Y(a) \mid a \in \mathcal{A}\} \perp\!\!\!\perp A \mid X$,
- 양수성(positivity) : $0 < c < 1$ 가 존재하여, 모든 $a \in \mathcal{A}$ 와 $x \in \mathcal{X}$ 에 대해 $P(A = a \mid X = x) \geq c$,
- 일치성(consistency) : 모든 $a \in \mathcal{A}$ 대해, $Y = Y(a)$.

교환가능성은 측정되지 않은 교란변수(confounding variable)가 없다는 것을 의미한다. 양수성은 모든 환자가 치료변수의 모든 값에 대해 0과 1사이의 확률값을 갖는다는 가정이다. $P(A = a \mid X = x)$ 의 값이 0 또는 1이라는 것은 공변량 $X = x$ 인 환자들이 대조군 혹은 치료군 중 한 군에만 배정받았다는 것이므로, 양수성은 임의의 $X = x$ 가 공변량으로 관측된 환자들에 대하여 치료군과 대조군이 한쪽으로 크게 치우치지 않았음을 의미한다. 일치성은 치료가 오류 없이 기록되었음을 보장한다. 즉, 동일한 치료군에 속한 개체들이 실제로 동일한 치료를 받았음을 의미한다.

2.2. Q-학습 모형과 개인 의료 추천 방법

Wallace와 Moodie (2015)는 기존의 Q-학습 방법과 G-추정 방법을 기반으로 하는 새로운 접근 방식을 제시하였다. Q-학습 방법은 반응변수에 대한 결과 모형을 기반으로 최소제곱법을 순차적으로 적용하기에 G-추정과 비교하였을 때 상대적으로 구현하기 간단하다는 장점이 있다. 하지만 Q-학습 방법으로 구한 추정량의 일치성은 결과 모형의 정확한 설정(correct specification)에 의존하기 때문에 관찰 연구에서 적용이 어렵다. 한편 G-추정 방법으로 블립 함수 모수의 추정량을 구할 때, 결과 모형 또는 치료 모형 중 하나만 정확하게 설정되어도 추정량은 일치성을 갖는다. 이에 Q-학습 방법의 구현 용이성과 G-추정 방법의 이중강건성을 동시에 갖는 동적 가중 최소제곱법을 제안하였다.

이제 y 는 연속형 결과변수, a 는 0 또는 1 값을 갖는 이분형 치료일 때, $\{(y_i, x_i, a_i); i = 1, \dots, n\}$ 가 관측되었다고 하자. 결과값 y_i 는 p 개의 설명변수 $x_i = (1, x_{1i}, x_{2i}, \dots, x_{pi})^T$ 와 치료 $a_i \in \{0, 1\}$ 에 의존한다. 이때, 관측되지 않은 교란변수(confounding variable)는 없고 어떤 개체의 결과는 다른 개체의 치료 할당에 영향을 주지 않는다고 가정한다. Q-학습 모형 $E[Y(a)|X = x, A = a]$ 은 주효과 모형 $f(x^\beta; \beta)$ 와 블립 함수 $\gamma(x^\psi, a; \psi)$ 의 합으로 다음과 같이 표현할 수 있다.

$$E[Y(a)|X = x, A = a] = f(x^\beta; \beta) + \gamma(x^\psi, a; \psi). \quad (2.1)$$

x^β 와 x^ψ 는 모두 관측된 공변량 벡터 x 의 부분집합으로 본 논문에서는 편의상 $x^\beta = x^\psi = x$ 을 가정하도록 한다. β 를 모수로 갖는 주효과 모형은 치료에 영향을 받지 않는 공변량 x 의 효과를 모형화한다. 반면, ψ 를 모수로 갖는 블립 함수는 치료 a 와 공변량 x 사이의 교호작용 효과를 정량화한다. 예를 들어, 이분형 치료 a 와 단일 공변량 x 로 구성된 다음의 선형 결과 모형을 고려하자.

$$E[Y(a)|X = x, A = a] = \beta_0 + \beta_1 x + a(\psi_0 + \psi_1 x).$$

여기서 주효과 모형은 $f(x; \beta) = \beta_0 + \beta_1 x$ 이고 블립 함수는 $\gamma(x, a; \psi) = a(\psi_0 + \psi_1 x)$ 이다. 이 경우 a 에 대한 최적의 치료 요법은

$$d^{\text{opt}}(x) = \begin{cases} 1, & \text{if } \hat{\psi}_0 + \hat{\psi}_1 x > 0 \\ 0, & \text{otherwise} \end{cases} = I(\hat{\psi}_0 + \hat{\psi}_1 x > 0),$$

으로 계산될 수 있으며, 이때 모든 a 및 x 값에 대해 $E[Y(d^{\text{opt}})] \geq E[Y(a)]$ 가 성립함을 알 수 있다. 따라서 연구의 목적은 블립 함수 $\gamma(x, a; \psi)$ 를 최대한 정확히 추정하고 이를 통해 $d^{\text{opt}}(x)$ 를 적절히 찾아내는 데 있다.

그러나 주효과 모형 $f(x; \beta)$ 가 데이터에 알맞게 설정되지 않았을 경우 블립 함수의 추정에도 편향을 일으키게 된다. 이를 해결하기 위해 성향점수 $\pi(x) = P(A = 1|X = x)$ 에 기반한 가중치 $w(a; x)$ 를 설정함으로써

주효과 모형의 오설정으로 인한 편향 문제를 어느 정도 제거할 수 있는데 이를 이중강건성(double robustness)이라 한다. 즉, 주효과 모형 $f(x; \beta)$ 혹은 $\pi(x)$ 에 대한 모형 둘 중 하나만이라도 참 모형이 설정되면 블립 함수에 대한 편향을 제거할 수 있게 된다 (Robins, 2004; Wallace와 Moodie, 2015). 이러한 사실을 기초로 동적 가중 최소제곱법은 블립 함수 모수의 일치 추정량을 보장하는 값으로 가중치를 결정하는데, 이때 가중치 $w(a; x)$ 는

$$\pi(x)w(1; x) = (1 - \pi(x))w(0; x), \quad (2.2)$$

의 균형 조건 (Wallace와 Moodie, 2015)을 만족해야 한다. 가장 널리 쓰이는 가중치로는 역확률 가중치(inverse probability of treatment weighting; IPTW)가 있으며 보통

$$w(a; x) = \frac{I(a = 1)}{\pi(x)} + \frac{I(a = 0)}{1 - \pi(x)}, \quad (2.3)$$

와 같이 주어진다. 이 경우 만약 결과 모형에서 $f(x; \beta)$ 이 오설정되었다라도 가중치가 참 치료 모형 $\pi(x)$ 에 의해 선택되기만 한다면 ψ 에 대한 추정량은 일치성을 갖는다. 또한, 치료 모형이 오설정되어 가중치를 제대로 선택하지 못한 경우라도 결과 모형이 정확히 설정되는 한 불편추정량을 구할 수 있다. 하지만 두 경우 모두 블립 함수가 정확히 정의된 경우에만 ψ 의 추정량이 의미가 있기 때문에 최적의 치료 요법을 찾기 위해 블립 함수를 정확히 설정하는 것이 중요하다.

2.3. 변형된 공변량을 이용한 회귀분석

Tian 등 (2014)은 무작위 임상시험에서 결과 모형을 모를 때 공변량을 변형하는 것만으로 일치 추정량을 얻을 수 있는 방법을 제안하였다. 관측된 데이터에서 이분형 치료 A 를 $\{0, 1\}$ 로 입력하는 것 대신 $\{-1, 1\}$ 로 입력한다. 결과 Y 는 연속형이고 $Y(1)$ 와 $Y(-1)$ 은 각각 치료 $A = 1$ 및 -1 를 받은 환자의 잠재적 결과이다. 치료는 환자에게 무작위로 할당되어서 치료 A 와 공변량 X 가 독립이라고 가정한다. 평균 인과효과(average causal effect; ACE)를 간단하게 표현하기 위해 $P(A = 1) = P(A = -1) = 1/2$ 이라고 가정한다. 이제 치료와 공변량 간의 교호작용을 포함한 다음의 다변량 선형 회귀 모형을 고려한다.

$$y_i = f(x_i; \beta) + a_i (x_i^T \psi) + \epsilon_i, \quad (2.4)$$

여기서 ϵ 은 평균이 0인 랜덤 오차항이고 앞서와 마찬가지로 잠재적인 부분 균을 나눌 때 $x^T \psi$ 값의 부호를 기준으로 나눌 수 있다. 이때, 식 (2.4)에서 공변량 $X = x$ 를 가진 환자에 대해 평균 인과효과는 $\Delta(x) = E(Y(1) - Y(-1) | X = x) = 2(x^T \psi)$ 로 주어진다. 한편, 간단한 계산에 의해 $E(2YA | X = x) = \Delta(x)$ 임을 확인할 수 있으며 이를 통해 ψ 의 추정량을 아래와 같이 계산할 수 있다.

$$\hat{\psi} = \operatorname{argmin}_{\psi} \frac{1}{n} \sum_{i=1}^n (y_i a_i - x_i^T \psi)^2 = \operatorname{argmin}_{\psi} \frac{1}{n} \sum_{i=1}^n \{y_i - a_i (x_i^T \psi)\}^2. \quad (2.5)$$

식 (2.5)의 두 번째 항등식은 $a_i \in \{-1, 1\}$ 이기 때문에 성립하는데, 위와 같이 공변량을 (a_i, x_i) 로 변형한 후 일반 최소제곱법을 통하여 추정량을 구하는 방법을 변형된 공변량 방법(modified covariate method)라고 한다 (Tian 등, 2014). 이제 $z_i = a_i x_i$ 라 정의하면 식 (2.5)의 최우변은 다음의 모형에 대한 최소제곱법에 해당된다.

$$y_i = z_i^T \psi + \epsilon_i. \quad (2.6)$$

모형 (2.4)와 (2.6)을 비교해보면 식 (2.6)에서는 주효과 모형을 설정하지 않아도 ψ 를 추정할 수 있어 $f(x; \beta)$ 의 오설정에 대해 보다 강건하다. 다만 주효과 모형을 설정하지 않을 경우 ψ 추정치의 표본분산이 커질 수 있으며

4장의 모의실험을 통해 이를 비교하고자 한다. 마지막으로 식 (2.5)를 통해 $\hat{\psi}$ 를 얻으면 a_i 에 대한 최적의 치료 요법은 다음의 식을 통해 계산한다.

$$d^{\text{opt}}(x_i) = \begin{cases} 1, & \text{if } a_i(x_i^T \hat{\psi}) > 0 \\ -1, & \text{otherwise} \end{cases} = \text{sign}\{a_i(x_i^T \hat{\psi})\}. \quad (2.7)$$

3. 이중강건 동적 가중 최소제곱법

동적 가중 최소제곱법에서 일치하는 추정량을 얻기 위해서는 블립 함수를 정확히 설정해야 함을 확인하였다. 또한 최적의 요법을 결정하는 규정변수의 변수 중요도는 공변량과 치료 간 교호작용 효과의 크기에 의존한다고 알려진 바 있다 (Gunter 등, 2011). 이런 까닭에 최적의 동적 치료 요법을 찾는 것은 치료와 공변량 간 교호작용 효과인 블립 함수와 블립 함수의 영이 아닌 계수들의 변수 중요도에 의존한다. 이 절에서는 주효과 모형을 포함한 결과 모형에 대한 설정 없이 적용할 수 있는 새로운 추정 방법을 제안하고 이를 lasso 및 adaptive lasso를 이용하여 고차원에서 정규화하는 방법에 대해 논의한다.

Q-학습 방법을 이용하여 최적의 동적 치료 요법을 찾기 위해서는 블립 함수를 최대화시켜야 하며, 이는 결과 모형의 설정에 의존한다. 3장에서는 2장에서 논의한 Q-학습 모형에 동적 가중 최소제곱법과 변형된 공변량 방법을 기반으로 주효과 모형의 설정 없이 적용할 수 있는 추정 방법을 제안하고자 한다. 관측된 데이터 $\{(y_i, x_i, a_i); i = 1, \dots, n\}$ 에서 관찰 연구에서 인과 추론을 가능하게 하기 위한 가정을인 교환가능성, 양수성, 일치성이 성립한다 하자. 그러나 관찰 연구에서 X 가 교란변수(confounding variable)일 경우에는 2.3절의 변형된 공변량 기반 회귀를 사용할 수 없게 된다. Wallace와 Moodie (2015)는 균형 조건 (2.2)을 만족하는 w 로 가중치를 적용한 데이터 $((y_w)_i, (x_w)_i, (a_w)_i; i = 1, \dots, n)$ 를 생각할 때 $x_w \perp\!\!\!\perp a_w$ 가 성립함을 보였다. 이에 본 논문에서는

$$L(\psi) = \frac{1}{n} \sum_{i=1}^n w_i(a_i, x_i) \{y_i - a_i(x_i^T \psi)\}^2, \quad (3.1)$$

와 같이 손실함수를 정의하고 블립 함수의 규정변수를 추정하기 위해 아래와 같이 벌점화를 고려한다.

$$\hat{\psi} = \underset{\psi}{\text{argmin}} \left[L(\psi) + \lambda_n \sum_{j=0}^p J(|\psi_j|) \right]. \quad (3.2)$$

이때, λ_n 은 초모수이고 J 는 축소 벌점(shrinkage penalty)이다. $w_i(a_i, x_i)$ 에 대해서는 균형 조건 (2.2)를 만족하는 다양한 형태의 가중치를 고려할 수 있으나 본 논문에서는 식 (2.3)의 역확률 가중치를 사용한다. 제안된 추정량은 만약 블립 함수가 정확하게 설정되었다면 결과 모형의 설정 없이 블립 함수만으로 일치하는 추정량을 얻을 수 있음을 시사한다. 블립 함수의 설정을 평가하기 위한 방법이 연구되어 있다 (Wallace 등, 2017). 위 정리에서 가중 최소제곱 추정량을 구할 때는 상수항이 없는 모형을 사용한다. 이는 변형된 공변량 방법을 적용한 결과 상수항 $f(x; \beta)$ 와 z 가 직교하기 때문이다. 축소 벌점으로 볼록 함수(convex function)인 lasso 벌점 (Tibshirani, 1996)과 adaptive lasso 벌점 (Zou, 2006)을 사용한다. lasso 방법은 L_1 벌점을 사용한 방법이고 아래의 해를 통해 구할 수 있다

$$\hat{\psi}^{\text{Lasso}} = \underset{\psi}{\text{argmin}} \left[L(\psi) + \lambda_n \sum_{j=0}^p |\psi_j| \right]. \quad (3.3)$$

Adaptive lasso는 기존 방법을 개선하여 작은 추정값을 가지는 계수에 대해 더 큰 가중치를 주는 방법이며

구하는 식은 다음과 같다

$$\hat{\psi}^{\text{aLasso}} = \underset{\psi}{\operatorname{argmin}} \left[L(\psi) + \lambda_n \sum_{j=0}^p k_j |\psi_j| \right]. \quad (3.4)$$

이때, $k_j = 1/|\hat{\psi}_j|$, $j = 0, \dots, p$ 이며 $\hat{\psi}_j = (\hat{\psi}_0, \dots, \hat{\psi}_p)^T$ 는 손실함수 (3.1)을 최소화함으로써 추정한다.

독립항등분포 회귀 모형의 가정하에서 lasso 방법은 변수 선택의 일치성이 보장되지 않는 반면 adaptive lasso 방법은 신탁 성질(oracle property)을 보장한다 (Zou, 2006). Lasso 방법의 일치성을 보장하기 위해서는 대표할 수 없는 조건(irrepresentable condition)이 만족되어야 한다 (Zhao와 Yu, 2006). Lasso 방법과 adaptive lasso 방법 간의 성능 비교는 4절과 5절에서 다룬다. 제안된 추정 방법에 adaptive lasso 별점을 적용하였을 때, 추정량의 점근적 성질과 신탁 성질(oracle property)을 아래의 정리를 통해 보인다. $S = \{j : \psi_j \neq 0, j = 0, \dots, p\}$ 라 하자. 일반성을 잃지 않고 $\psi = (\psi_S^T, \psi_{S^c}^T)$ 라 쓸 수 있다. 그리고 $\hat{S} = \{j : \hat{\psi}_j \neq 0, j = 0, \dots, p\}$ 일 때 $\hat{\psi}_s$ 는 추정된 규정변수들로 이루어진 ψ 의 부분집합이라 하자. 이 때, 다음의 정리들이 성립하며 그 증명은 부록에 제시하였다.

Theorem 1. 만약 아래의 정규 조건 (RC1)–(RC3)이 성립하며 모형 (3.4)가 정확하게 설정되었고 치료 모형 $\pi(x)$ 가 알려져 있다면, $n \rightarrow \infty$ 일 때 $\sqrt{n}(\hat{\psi} - \psi) \rightarrow N(0, V)$ 이다. 이때 $U = 2E[w(a, x)a^2x(x)^T]$ 이고 $\Omega = E[q_{n,1}(\psi_0)q_{n,1}^T(\psi_0)]$ 일 때, $V = U^{-1}\Omega U^{-1}$ 이다.

(RC1) 각각의 관찰값들 x_i , $i = 0, \dots, n$ 은 독립이며 동일한 분포를 가진다.

(RC2) U 가 존재하며 비특이행렬이다.

(RC3) ψ^* 의 근방안에 점 ψ 에서 적분 가능한 함수 g 에 대해, 모든 $1 \leq i, j, k \leq p$ 와 모든 ψ 에서 $|\partial^3 L(x)/\partial \psi_i \psi_j \psi_k| < g(x)$ 이다. 이때, ψ^* 는 정리3.1의 증명 내에서 정의된다.

Theorem 2. $\sqrt{n}\lambda_n \rightarrow 0$ 이고 $n\lambda_n \rightarrow \infty$ 라 가정하자. 정리 3.1의 조건하에서 다음이 성립한다.

1. (선택 일치성) $\lim_{n \rightarrow \infty} P(\hat{S} = S) = 1$.

2. (점근적 정규성) $n \rightarrow \infty$ 일 때, $\sqrt{n}(\hat{\psi}_s - \psi_s) \xrightarrow{d} N(0, \Sigma)$.

만약 블립 함수가 정확하게 설정되어 있다면 정리 3.2에 따라 일치성을 보장할 수 있다. 또한 참 치료 모형 $\pi(x)$ 를 모르는 상황에서 모수 α 에 대한 모형 $\pi(x; \alpha)$ 를 추정하여 대체값을 찾을 수 있다. 만약 치료 모형 $\pi(x; \alpha)$ 를 정확하게 설정하였다면 모수 α 의 최대우도 추정량 $\hat{\alpha}$ 은 일치성을 가지므로 $\pi(x; \hat{\alpha})$ 값을 대체할 수 있다. 본 논문의 모의실험에서는 회귀나무, 랜덤 포레스트, Super learner 등 다양한 앙상블 기계 학습 기법을 활용하여 $\pi(x)$ 를 추정하였으며 이에 따른 ψ 추정치의 강건성을 평가하였다.

제안된 최적의 치료 요법은 아래의 절차를 따라 추정 가능하다.

Step 1. 참 치료 모형 $\pi(x)$ 이 알려진 경우 균형 조건 (2.2)를 만족하는 가중치 $w(a, x)$ 를 계산한다. 만약 참 치료 모형이 알려져 있지 않아도 $\pi(x; \hat{\alpha})$ 가 $\pi(x)$ 의 일치 추정량이라면 $\pi(x; \hat{\alpha})$ 를 사용한다.

Step 2. $a \in \{-1, 1\}$ 에 대해 *Step 1*에서 구한 가중치 $w(a, x)$ 를 적용하여 (3.2)로부터 해 $\hat{\psi}$ 를 구한다.

Step 3. 공변량 $X = x$ 인 개체의 최적의 치료 요법의 추정량 $\hat{d}^{\text{opt}}(x) = \operatorname{sign}\{a(x^T \hat{\psi})\}$ 로 계산한다.

마지막으로 제안된 방법론은 동적 치료 요법 구현을 위해 다중 결정 문제에 대해서도 확장 가능하나 본 논문에서는 단일 결정 모형에 집중한다. 다중 결정 모형으로의 확장에 관한 내용은 Wallace 등 (2017)을 참조하기 바란다.

Table 1: Five fitting cases for blip function estimation on simulation 1

Model	Regression y on	Weight (w)	Outcome model	Treatment model
(M1)	$(1, e^{x_1 x_2}, a, ax_1, ax_2)$	Not applied	Correct	-
(M2)	$(1, x_1, x_2, a, ax_1, ax_2)$	Not applied	Mis-specification	-
(M3)	$(1, x_1, x_2, a, ax_1, ax_2)$	Applied	Mis-specification	Logistic, Tree and SL
(M4)	(a, ax_1, ax_2)	Not applied	Omitted	-
(M5)	(a, ax_1, ax_2)	Applied	Omitted	Logistic, Tree and SL

4. 모의실험

4.1. 모의실험 1

첫 번째 모의실험에서는 총 5가지의 모형을 설정하고 각 모형이 참 데이터 생성 모형과 일치하는지 여부에 따라 ψ 에 대한 이중강건 추정치가 어떻게 변화하는지 살펴본다. 표본의 크기 $N = 100$ 과 $1,000$ 에 대해 두 공변량 $X = (X_1, X_2)$, $X_1 \sim \text{Uniform}(-1, 1)$ 및 $X_2 \in \{-1, 1\} \sim \text{Bernoulli}(0.5)$ 를 생성한다. 치료변수 $A \in \{-1, 1\}$ 는 $P(A = 1|X_1, X_2) = \text{expit}(X_1 + X_2 + m \cdot (X_1 X_2)) = 1 - P(A = -1|X_1, X_2)$ 의 확률을 따르며 여기서 $\text{expit}(x) = e^x / (1 + e^x)$ 이다. m 은 $\{0, 1, 2\}$ 중 하나의 값을 갖는 상수로 m 값이 커질수록 A 에 대한 교호작용 $X_1 X_2$ 의 효과가 증대된다. 즉, m 값이 커질수록 비선형향이 치료에 미치는 영향력이 커진다. 마지막으로 반응변수 Y 는 다음의 모형으로 추출하였다

$$Y = 5 + 0.5 \exp(X_1 X_2) + A(1 + X_1 + X_2) + \epsilon. \quad (4.1)$$

제안된 방법이 블립 함수의 모수 $\psi_0 = \psi_1 = \psi_2 = 1$ 을 잘 추정하는지 비교하기 위해 Table 1과 같이 총 5가지의 모형 적합을 시도하였다. 가령 (M1)에서는 Y 에 대해 다음 형태의 모형을 설정하였다.

$$(M1): E[Y|A = a, X = x] = \beta_0 + \beta_1 \exp(x_1 x_2) + a(\psi_0 + \psi_1 x_1 + \psi_2 x_2).$$

(M1)과 데이터 생성 모형 (4.1)이 동일함을 알 수 있다. 즉 결과 모형(outcome model)이 바르게 설정되었으며 이 경우 가중치 w 는 따로 적용하지 않았다. 반면 모형 (M2), (M3)의 결과 모형은

$$(M2), (M3): E[Y|A = a, X = x] = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + a(\psi_0 + \psi_1 x_1 + \psi_2 x_2),$$

으로 (4.1)과 서로 다르며 (M2)에서는 가중치 없이, (M3)에서는 가중치를 포함하여 추정하였다. 마지막으로 모형 (M4), (M5)에서는 주효과 모형을 제외하고 교호작용에 대한 블립 함수만으로 아래와 같이 모형을 설정하였으며 각각 가중치를 비포함, 포함하였다.

$$(M4), (M5): E[Y|A = a, X = x] = a(\psi_0 + \psi_1 x_1 + \psi_2 x_2).$$

치료 모형 $\pi(x) = P(A = 1|X = x)$ 에 대한 추정을 위해 로지스틱 회귀(logistic regression), CART (classification and regression tree) (Breiman 등, 1984), SL (super learner) (van der Laan 등, 2007)을 적용하였고, 이를 위해 각각 패키지 `rpart` (Therneau 등, 2014), `SuperLearner` (Polley 등, 2016)를 사용하였다. 로지스틱(logistic) 회귀에서는 교호작용을 제외한 주효과만을 포함하였다. SL은 서로 다른 형태의 분류 기법을 혼합하는 앙상블 기계학습 기법으로 본 모의실험에서는 `SL.glm`, `SL.rpart` 및 `SL.step.interacion`의 세 옵션을 선택하였다. 여기서 `SL.glm` 옵션은 일반화 회귀분석, `SL.rpart` 옵션은 분류 및 회귀나무, 그리고 `SL.step.interacion` 옵션은 대응변수 교호작용들(pairwisely variable interactions)에서 단계적 변수 선택을 위한 옵션이다. 만약 앙상블 기계학습 기법을 통해 치료 모형이 잘 추정되었다면 비록 (M3)에서 결과 모형이 오설정되었다라도 ψ 에 대한 불편추정량을 얻을 수 있으며

Table 2: Bias and MSE of parameter estimators in the blip function

	Model	PS	$\hat{\psi}_0$		$\hat{\psi}_1$		$\hat{\psi}_2$	
			Bias	MSE	Bias	MSE	Bias	MSE
$m = 0$	(M1)	-	0.0006	0.0034	-0.0011	0.0147	-0.0018	0.0046
	(M2)	-	-0.0015	0.0054	0.2888	0.0961	0.1040	0.0155
	(M3)	True	-0.0018	0.0056	0.0144	0.0204	0.0045	0.0053
		Logistic	-0.0016	0.0057	0.0118	0.0183	0.0031	0.0048
		Tree	-0.0025	0.0063	0.0461	0.0197	0.0305	0.0065
	(M4)	SL	-0.0017	0.0055	0.0162	0.0183	0.0048	0.0049
		-	0.0005	0.0037	0.2868	0.0948	0.1032	0.0153
		True	-0.0015	0.0054	0.0093	0.0198	0.0023	0.0052
		Logistic	-0.0014	0.0056	0.0074	0.0180	0.0012	0.0048
	(M5)	Tree	-0.0022	0.0059	0.0418	0.0195	0.0290	0.0064
SL		-0.0015	0.0053	0.0125	0.0180	0.0033	0.0049	
(M1)		-	-0.0012	0.0036	0.0010	0.0144	-0.0041	0.0045
(M2)		-	0.0915	0.0137	0.2822	0.0933	0.0905	0.0131
(M3)		True	0.0039	0.0132	0.0166	0.0197	0.0027	0.0053
	Logistic	0.0879	0.0060	0.0222	0.0229	0.0132	0.0058	
	Tree	0.0189	0.0065	0.0543	0.0221	0.0189	0.0062	
	SL	0.0113	0.0064	0.0242	0.0252	0.0051	0.0063	
	(M4)	-	0.0592	0.0073	0.2595	0.0805	0.0875	0.0123
(M5)	True	0.0015	0.0123	0.0096	0.0180	0.0001	0.0047	
	Logistic	0.0845	0.0058	-0.0060	0.0222	-0.0049	0.0057	
	Tree	0.0174	0.0062	0.0483	0.0209	0.0172	0.0060	
	SL	0.0100	0.0060	0.0196	0.0246	0.0035	0.0061	
	$m = 2$	(M1)	-	-0.0006	0.0041	0.0010	0.0127	-0.0025
(M2)		-	0.1670	0.0329	0.2688	0.0877	0.0681	0.0103
(M3)		True	0.0159	0.0090	0.0354	0.0369	0.0112	0.0088
		Logistic	0.1589	0.0303	0.0523	0.0223	0.0371	0.0072
		Tree	0.0442	0.0093	0.0840	0.0302	0.0206	0.0084
(M4)		SL	0.0206	0.0083	0.0476	0.0338	0.0142	0.0079
		-	0.1181	0.0179	0.1946	0.0505	0.0588	0.0080
		True	0.0090	0.0083	0.0218	0.0357	0.0054	0.0083
		Logistic	0.1492	0.0268	-0.0167	0.0158	-0.0058	0.0045
(M5)		Tree	0.0412	0.0089	0.0691	0.0253	0.0154	0.0074
	SL	0.0169	0.0078	0.0376	0.0317	0.0104	0.0074	

이는 이중강건성의 중요한 특성이다. (M5)의 추정량도 (M3)처럼 이중강건성을 보이는가 모의실험을 통해 살펴보고자 한다.

Table 2는 300개의 독립 표본으로부터 1,000번 모의실험을 반복한 결과를 나타낸다. 상수 m 의 크기 변화(즉, 비선형항의 영향력 변화), 적합 방법, 그리고 치료 모형 추정을 위해 사용한 기계학습 방법에 따라 블립 함수 모수 추정 결과가 제시되어 있다. 치료 모형 추정 시 사용한 기계학습 방법을 포함하여 참 치료 모형을 사용한 경우의 결과("True"로 표기)도 함께 제시하였다. 각 적합 방법의 성능을 평가하기 위해 추정량의 편차(bias)와 평균제곱오차(mean square error; MSE)를 사용하였고, 그 값이 작을수록 참 값에 근접하게 잘 추정한 것으로 평가한다. (M1)은 결과 모형을 정확하게 설정한 경우이고, (M2)는 결과 모형을 오설정된 경우로 두 결과를 비교하여 결과 모형의 비선형항이 성능에 얼마나 영향을 끼치는지 확인할 수 있다. 또한 (M2)와 (M3)

의 결과를 비교하였을 때 (M3)에 성향점수에 기반한 가중치를 사용함으로써 얼마나 추정량의 성능이 개선되었는지 확인할 수 있다. 이와 비슷하게 (M4)도 결과 모형과 치료 모형 모두 모형 오설정된 경우로 (M5)와 비교해서 가중치를 사용했을 경우 추정량의 성능을 평가할 수 있다. 또한 이를 동적 가중 최소제곱법을 이용한 (M3)의 성능과 비교해본다. 그리고 m 의 변화에 따른 결과를 살펴봄에 치료 모형에서 비선형항의 영향력이 커질 때 기계학습 방법 별로 성능의 차이가 있는지 확인해 보았다.

Table 2에서 $m = 0$ 인 경우, 즉 교호작용이 A 에 영향을 미치지 않는 경우 모든 모형이 블립 함수를 안정적으로 추정하고 있다. $m = 0$ 인 경우는 치료 모형의 추정 모형이 선형인 경우로 이 경우 로지스틱 회귀는 선형 결정 경계(decision boundary)를 만들어 내지만 분류 및 회귀나무는 비선형 결정 경계를 만들어 내기 때문에 앙상블 학습의 성능이 상대적으로 좋지 않은 것으로 보인다. 그러나 m 값이 커질수록 모형의 오설정으로 인한 편향이 발생하고 있으며 로지스틱 회귀를 이용하여 역확률 가중치를 고려한다 하더라도 편향이 줄어들지 않음을 확인할 수 있다. 비선형항의 영향이 커지는 경우에는 로지스틱 회귀의 경우 $\hat{\psi}_1$ 과 $\hat{\psi}_2$ 의 편차는 다른 방법과 비교하여 상대적으로 작은 반면 $\hat{\psi}_0$ 의 편차는 커짐을 확인할 수 있다. 특히 $m = 2$ 의 경우 (M4)와 로지스틱 회귀를 활용한 (M5)의 결과를 비교하면 오히려 가중치를 적용했을 때 오차가 커지는 현상도 발생하였다. 이 경우 분류 및 회귀나무(tree) 혹은 SL을 사용하여 상당 부분 편차를 줄일 수 있었으며, 특히 SL은 모든 모의실험 결과에서 대체로 안정적이고 우수한 성능을 보였다.

4.2. 모의실험 2

실제 고차원의 데이터를 분석할 때 관심있는 결과 Y 에 크게 영향을 끼치지 않는 공변량들이 포함되어 있는 경우가 있다. 모의실험 2에서는 모의실험 1과 마찬가지로 $X_1 \sim \text{Uniform}(-1, 1)$ 및 $X_2 \in \{-1, 1\} \sim \text{Bernoulli}(0.5)$ 를 생성하고 여기에 노이즈 변수 $X_3, X_4, \dots, X_{10} \sim N(0, 1)$ 을 추가하였다. 치료변수 A 와 결과값 Y 는 모의실험 1과 동일하게 추출하였다. $X = (1, X_1, X_2, \dots, X_{10})^T$ 라 하자. 모의실험 1의 Table 1에서와 같이 다음의 5가지 회귀모형을 적합하였다: (M1): $y \sim (e^{x_1 x_2}, ax)$; (M2), (M3): $y \sim (x, ax)$; (M4), (M5): $y \sim ax$. 여기에 (M3)와 (M5)에서는 성향점수 가중치를 추가하였다. 변수 선택을 위하여 (3.3)와 (3.4)의 lasso 및 adaptive lasso 방법을 사용한 벌점 회귀모형을 사용하였다. 모수의 추정은 극좌표 하강 알고리즘에 기반하였으며 10-분할 교차 검증(10-fold cross validation)으로 조율 모수 λ 를 추정하였다.

추정량의 변수 선택 성능을 평가하기 위해 추정된 계수들 중 결과에 영향을 끼치는 변수는 영이 아닌 값으로, 노이즈 변수들은 영으로 추정되었는지를 평가한다. 블립 함수의 모수는 $\psi_0 = (1, 1, 1, 0, 0, 0, 0, 0, 0, 0)^T$ 로 지정하였다. Table 3에서는 3개의 유의변수 (X_0, X_1, X_2)를 잘 찾아내는지, 그리고 8개의 노이즈 변수 (X_3, \dots, X_{10})를 잘 걸러내는지 여부를 Cor 및 Incor로 기록하였다. 즉 Cor은 실제로 영이 아닌 계수를 영이 아닌 것으로 추정했는지를, Incor은 실제로 영인 계수를 영으로 잘 추정했는지를 나타내며 각각 3과 0에 근접할수록 변수 선택이 잘된 것으로 평가한다. Cor과 Incor이 전체 계수를 두 범주로 나누어서 변수 선택의 성능을 평가하는 지표였다면 Exact는 개체 단위로 변수 선택의 성능을 평가하는 지표이다. 즉 Exact(%)는 전체 개체 중 모든 계수에 대해 영인 것은 영으로 영이 아닌 것은 영이 아닌 것으로 제대로 선택된 추정 결과 수의 비를 계산한 값이다.

Table 3의 결과를 보면 결과 모형의 오설정이 있더라도 역확률 가중치를 통해 보정을 해줄 경우 유의한 블립 함수 변수를 잘 찾아냄을 확인하였다. True의 결과와 비교하여 보면 Cor의 경우와 Incor의 경우 모두 Lasso 방법보다 adaptive lasso 방법이 더 변수 선택 성능이 좋은 것으로 평가된다. Exact의 결과 역시 lasso 방법보다 adaptive lasso 방법이 변수 선택 성능이 좋은 것으로 평가된다. 이는 adaptive lasso가 작은 추정량을 가지는 계수에 대해 더 많은 가중 벌점을 주어서 변수가 선택되지 못하게 하기 때문인 것으로 생각된다. (M3)와 (M5)에서 True의 결과를 비교하면 lasso와 adaptive lasso 방법 모두 큰 차이를 보이지 않는 것으로 판단된다.

마지막으로 치료 결정 규칙인 $I(\hat{z}_i^T \hat{\psi} > 0)$ 의 정확도를 평가하기 위해 $n = 300$ 개의 데이터 셋에서 올바른

Table 3: Simulation results of variable selection for parameter in the blip function

Penalty	Model	PS	$m = 0$			$m = 1$			$m = 2$			
			Cor	Incor	Exact(%)	Cor	Incor	Exact(%)	Cor	Incor	Exact(%)	
Lasso	(M1)	-	3	1.11	36.5	3	1.17	36.0	3	1.19	36.5	
	(M2)	-	3	1.01	39.0	3	1.17	38.0	3	1.27	34.5	
	(M3)	True	3	0.95	44.0	3	1.13	41.0	3	1.19	40.0	
		Logistic	3	0.97	43.5	3	1.24	38.5	3	1.48	29.0	
		Tree	3	1.05	40.5	3	1.11	41.5	3	1.22	35.0	
	SL	SL	3	0.99	41.0	3	1.09	43.0	3	1.13	37.5	
		(M4)	-	3	1.63	22.5	3	1.57	24.5	3	1.62	27.5
		(M5)	True	3	1.39	34.5	3	1.41	33.0	3	1.61	27.5
			Logistic	3	1.47	28.5	3	1.60	28.5	3	1.82	22.0
			Tree	3	1.52	26.0	3	1.59	23.0	3	1.65	23.0
	SL	3	1.38	29.0	3	1.53	30.0	3	1.53	28.5		
	Adaptive Lasso	(M1)	-	3	0.52	66.5	3	0.44	70.0	3	0.45	71.5
		(M2)	-	3	0.36	75.5	3	0.58	63.0	3	0.54	64.5
		(M3)	True	3	0.37	76.0	3	0.43	74.0	3	0.41	75.0
			Logistic	3	0.44	73.5	3	0.50	72.5	3	0.54	66.0
Tree			3	0.36	74.5	3	0.39	74.5	3	0.45	74.0	
SL		SL	3	0.43	71.5	3	0.44	73.0	3	0.43	72.5	
		(M4)	-	3	0.45	71.0	3	0.49	70.5	3	0.45	71.5
		(M5)	True	3	0.36	76.0	3	0.44	72.5	3	0.51	73.5
			Logistic	3	0.44	73.0	3	0.43	76.0	3	0.49	69.0
			Tree	3	0.41	74.5	3	0.40	77.5	3	0.53	69.0
SL		3	0.42	72.5	3	0.45	74.5	3	0.48	71.0		

Table 4: The percent of selecting the correct treatment rules

Model	PS	Lasso			Adaptive Lasso		
		$m = 0$	$m = 1$	$m = 2$	$m = 0$	$m = 1$	$m = 2$
(M1)	-	97.52	97.68	97.47	97.79	97.50	97.77
(M2)	-	97.43	97.49	97.95	98.03	97.33	97.60
(M3)	True	97.17	97.46	97.18	97.48	96.86	97.27
	Logistic	97.00	97.41	96.42	96.77	95.63	95.97
	Tree	96.67	97.21	96.93	97.36	96.56	96.96
	SL	97.08	97.45	97.13	97.42	97.07	97.41
	(M4)	-	97.27	97.43	97.85	98.07	97.35
(M5)	True	96.81	97.32	96.89	97.41	96.53	96.94
	Logistic	96.64	97.21	96.07	96.57	95.23	95.51
	Tree	96.31	96.95	96.63	97.22	96.24	96.75
	SL	96.78	97.30	96.91	97.32	96.84	97.28

결정을 한 경우의 비율(percent of making correct decision; PCD)을 아래와 같이 구한다.

$$1 - \sum_{i=1}^n |I(z_i^T \hat{\psi} > 0) - I(z_i^T \psi_0 > 0)| / n.f$$

200번의 반복실험에서 얻은 평균 백분율을 Table 4에 제시하였다. 상대적으로 비선형성이 약한 모형에서 자료

Table 5: List of variables in the lindner center data

변수명	변수 설명
cardbill	환자의 최초 PCI 후 6개월 이내에 발생한 심장 관련 비용(\$)
abcix	치료 선택 따른 지표. 0은 통상적인 PCI치료만 받은 것을 의미하고, 1은 통상적인 치료에 abciximab을 사용한 계획된 치료나 구제 치료를 포함한 것을 의미함
stent	관상 동맥 스텐트 전개 (1:예, 0:아니요)
height	키(cm)
female	여성 (1:예, 0:아니요)
diabetic	당뇨병 진단 여부 (1:예, 0:아니요)
acutem	지난 7일 이내의 급성 심근 경색 여부 (1:예, 0:아니요)
ejecfrac	좌심실 구혈률(%)
veslproc	환자의 최소 PCI 절차에 포함된 혈관 수

를 추출하였으므로 회귀 기반 모형인 (M2)의 성능이 나쁘지 않았다. (M3)과 (M5)의 True 결과를 비교해보면 (M5)의 방법이 동등한 성능을 보였으며, 특히 SL과 adaptive lasso를 사용했을 경우 모형의 오설정에도 불구하고 바른 치료 규칙을 전반적으로 잘 찾아냄을 확인하였다. 한편 Wenbin 등 (2011)과 Yuan 등 (2014) 연구의 모의실험에서 PCD는 주효과 모형 또는 교호작용 모형의 복잡성에 크게 영향을 받음을 확인할 수 있다. 본 모의실험의 경우 높은 결정 비율을 보이지만 주효과 모형이 더 복잡해 질수록 PCD는 낮아질 것으로 예상된다.

5. 예시

이 절에서는 제안된 방법을 실제 데이터에 적용해본다. Kereiakes 등 (2000)은 실제 임상 환경에서 제제의 임상 효능과 비용 효율성을 검증하기 위해 1997년 오하이오주 신시내티에 있는 The Christ Hospital에서 관찰 연구를 수행했다. 경피적관상동맥중재술(percutaneous coronary intervention; PCI)동안 혈전을 막기 위한 GP IIb/IIIa 수용체 차단제인 abciximab 제제에 대한 효능을 및 비용 효율성은 위약 대조군 무작위 실험을 통해 입증된 바 있었지만, 무작위 대조 시험에서 도출된 비용 효율성을 실제의 비용 효율성으로 적용하기엔 문제가 될 수 있기에 해당 연구를 진행한 것이다. 데이터 프라이버시 및 기밀성 문제로 인해 연구 데이터의 일부만 R 패키지 twang (Ridgeway 등, 2016)에 포함되어 있다. 사용된 Lindner Center 데이터 세트는 The Christ Hospital의 Ohio Heart Health Center에서 처음 PCI를 받은 996명의 환자에 대해 직후 6개월 동안 전향적으로 추적 관찰한 자료로 본 분석에서는 6개월 동안 생존한 환자 970명 만을 대상으로 하였다. 그중 697명은 abciximab 치료에 배정되었고 나머지 283명의 환자는 통상 치료만을 단독으로 받았다. Lindner Center 연구에서는 치료가 무작위 배정되지 않았기 때문에 중첩편향(confounding bias)이 존재한다.

Table 5에 결과변수와 치료변수 그리고 그 외 표본 집단의 특성을 나타내는 공변량들을 설명하였다. 분석에서는 기존 예측변수 7개와 각 변수들끼리의 이차 교호작용항 $7 \times 6/2 = 21$ 개를 추가하여 총 28개의 공변량을 사용하였다. cardbill은 환자가 수술을 받은 후 6개월 이내 카드 명세 비용으로 우리가 관심있는 결과변수로 분석에서는 로그 변환을 취하였다. abcix는 치료변수이고 치료 abciximab을 받은 그룹은 1로 일반 치료만을 받은 대조군 그룹은 -1로 입력하였다. 그 외의 변수는 수술받기 전 환자의 특성을 나타내는 공변량들이다. 분석의 목적은 환자의 특성에 근거하여 환자 집단의 평균 비용을 줄일 수 있는 최적의 치료를 찾는 것이고, 따라서 결과변수 cardbill은 작은 값일수록 선호된다. 블립 함수의 모수를 찾기에 앞서 로지스틱 회귀, 분류 및 회귀나무, SL 방법을 사용하여 치료 모형을 추정하며 세 방법의 설정은 4장과 동일하다. abciximab 치료를 받은 그룹 및 대조군을 받은 두 그룹에 대한 추정된 치료 모형의 분포가 Figure 1에 나타나 있다. 로지스틱 회귀와 SL의 분포는 상당히 유사하지만 분류 및 회귀나무는 비선형 결정 경계를 만들기 때문에 더 극단적인 분포의 형태를 띤다.

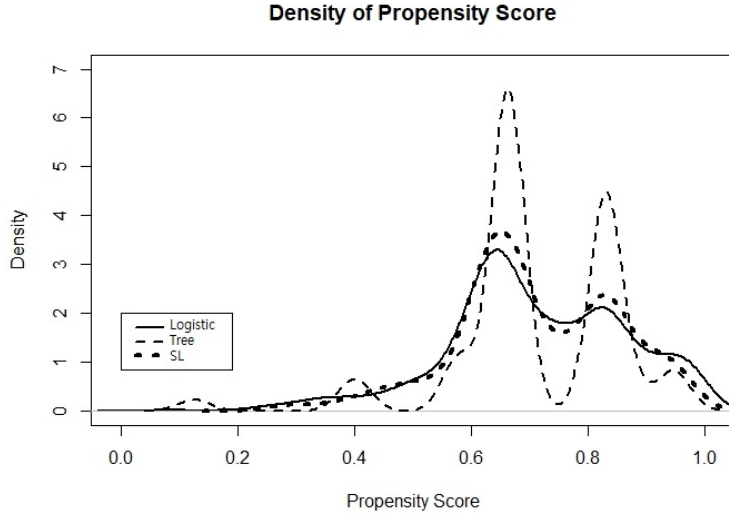


Figure 1: Comparison of kernel density function of propensity score model according to machine learning method.

분석을 위해 블립 함수에 총 28개(7개의 주요효과 및 21개의 교호작용)의 모든 공변량을 포함한 후 lasso 혹은 adaptive lasso 벌점화 방법을 사용하여 변수 선택을 진행하였다. 결과 모형을 알 수 없으므로 4장에서 제시한 (M2)–(M5) 모형을 사용하였다. 가령, 모형 (M3) 및 (M5)를 설정하고 super learner를 통한 가중치 및 adaptive lasso 기반 변수 선택법을 사용한 경우 아래와 같은 블립 함수의 형태를 얻을 수 있었다.

$$\begin{aligned} x^T \hat{\psi}_3 = & 9.43 + 0.08 \text{ stent} + 0.06 \text{ diabetic} - 0.10 \text{ stent} * \text{female} - 0.14 \text{ stent} * \text{diabetic} \\ & + 0.16 \text{ stent} * \text{acutemi} + 0.10 \text{ female} * \text{ves1proc} + 0.08 \text{ diabetic} * \text{acutemi} \\ & + 0.03 \text{ diabetic} * \text{ves1proc} \end{aligned}$$

$$\begin{aligned} x^T \hat{\psi}_5 = & 9.49 + 0.05 \text{ stent} + 0.15 \text{ diabetic} - 0.07 \text{ acutemi} - 0.10 \text{ stent} * \text{female} \\ & - 0.18 \text{ stent} * \text{diabetic} + 0.23 \text{ stent} * \text{acutemi} + 0.04 \text{ stent} * \text{ves1proc} \\ & - 0.01 \text{ height} * \text{acutemi} - 0.06 \text{ female} * \text{diabetic} - 0.04 \text{ female} * \text{acutemi} \\ & + 0.11 \text{ female} * \text{ves1proc} + 0.07 \text{ diabetic} * \text{acutemi} + 0.02 \text{ diabetic} * \text{ves1proc} \\ & + 0.04 \text{ acutemi} * \text{ves1proc}. \end{aligned}$$

(M3)에 비해 (M5)를 설정했을 때 보다 많은 변수가 치료 규칙을 찾는 데 사용되었다. 추정된 최적의 치료 요법 $\hat{d}^{\text{opt}}(x)$ 는 식 (2.7)과 같이 블립 함수값의 부호에 따라 결정한다. 만일 모든 환자가 최적의 치료법 $\hat{d}^{\text{opt}}(x)$ 을 받았다고 가정할 경우 $\log(\text{cardbill})$ 에 대한 예측 모형은 아래와 같을 것이다.

$$\log(\text{cardbill}) = f(x; \hat{\beta}) - |x^T \hat{\psi}|.$$

즉, 치료 방법을 개인별로 맞춤화함으로써 의료 비용을 줄이는 데 도움이 될 것이다. 치료 abciximab을 받은 그룹과 일반 치료를 받은 그룹의 최적화된 예측 비용에 대한 박스 플랏을 Figure 2에 제시하였다. 상단 및 하단의 그래프는 각각 lasso 및 adaptive lasso 방법을 적용한 결과이다. 그림에서 origin은 $E[\log(\text{cardbill})] = x^T \hat{\beta} + a(x^T \hat{\psi})$ 을 계산한 값이다. 로그 변환된 예상 비용 차이를 비교해보면 원래의 경우보다 최적의 치료를 받은 후 예상되는 비용이 작음을 확인할 수 있다. 이는 최적의 치료를 받는 것이 평균 인과효과를 작게 만들기 때문이다. (M3)와 (M5)를 비교하여 보면 원래의 경우와 비교하여서 (M3)의 결과는 분산의 크기가 비슷한

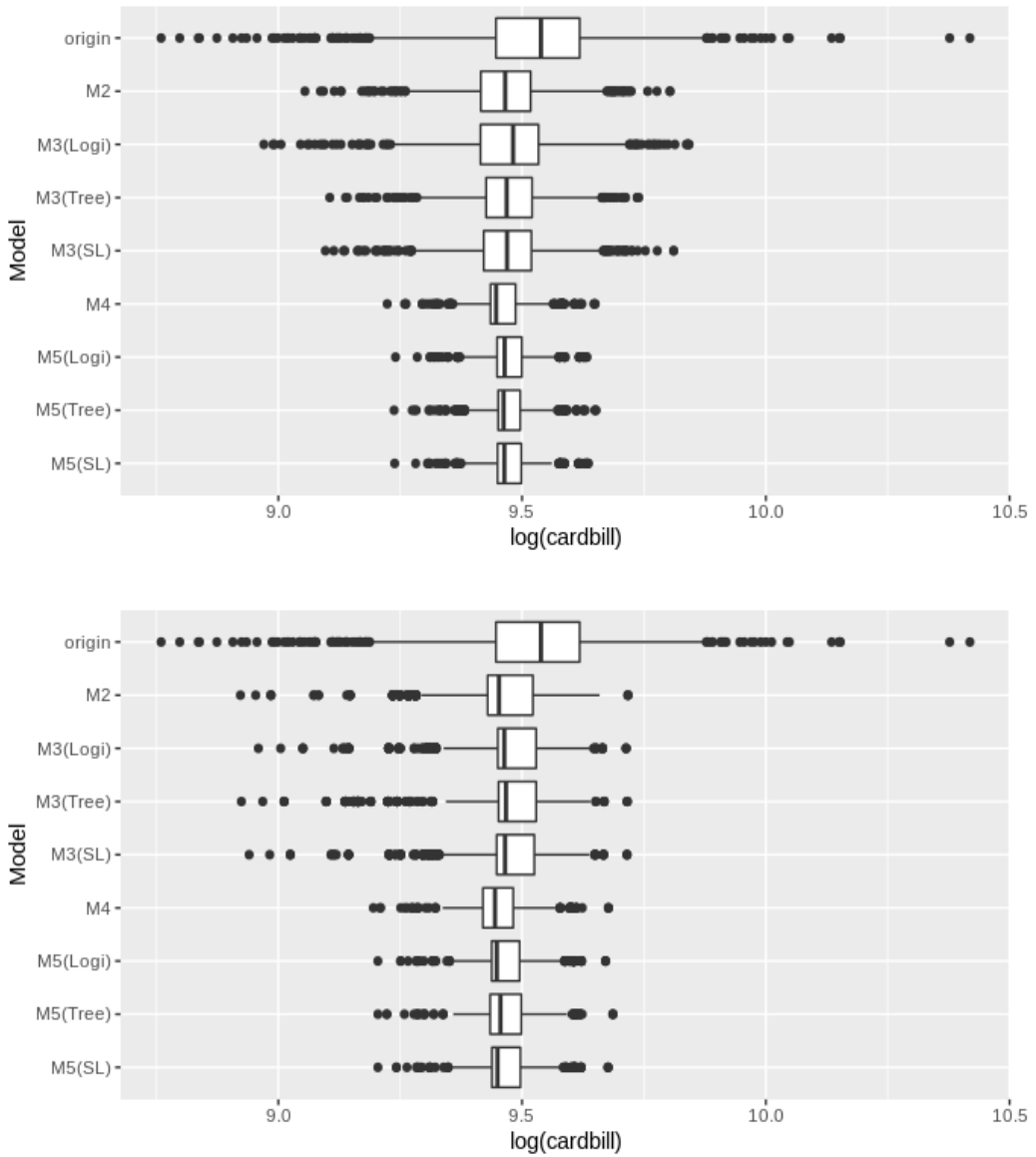


Figure 2: Comparison of predicted cost for log-transformed costs when the optimal treatment is prescribed using the lasso penalty (above) and the adaptive-lasso penalty (below).

반면 (M5)의 결과는 분산이 상대적으로 작아졌음을 확인할 수 있다. (M3)는 전체 모형을 적합하였기 때문에 교호작용항들만 적합한 경우보다 분산이 커지는 경향이 있는 것으로 보인다.

6. 논의

최적의 동적 치료 요법을 추정하기 위해 기존에 Q-학습, G-추정, 동적 가중 최소제곱법 등이 연구되었다. 동적 가중 최소제곱법은 기존의 Q-학습, G-추정의 장점을 지닌 추정량으로 이중강건성을 갖는다. 하지만 불립 함수의 정확한 설정에 의존하므로 공변량에 불필요한 변수가 많이 포함된 경우 적용이 어려울 수 있다. 특히나 고차원 데이터의 경우 이 방법을 적용하기 위해서는 변수 선택에 대한 연구가 필요하였다. 이에 본 논문에서는 치료 값의 입력 방법을 변형하여 불립 모형만으로 이중강건성을 갖는 추정량을 구할 수 있음을 밝혔다. 변수 선택을 위하여 lasso 벌점과 adaptive lasso 벌점을 고려하였다. 그리고 제안된 추정 방법에 adaptive lasso 벌점을 적용하였을 때의 추정량은 점근적 성질과 신탁 성질을 갖음을 확인하였다. 모의실험을 통하여 제안된 추정 방법과 기존의 동적 가중 최소제곱법을 비교한 결과 치료 모형의 비선형성이 커지더라도 동적 가중 최소제곱법보다 제안된 추정량의 성능이 좋았으며 특히 SL와 같은 앙상블 기계학습 모형을 이용하여 가중치를 추가함으로써 상당 부분의 편향을 제거할 수 있었다. 또한 adaptive lasso 방법을 적용하였을 때 두 방법의 변수 선택 성능과 치료 결정 규칙의 정확성이 거의 비슷함을 확인하였다. 제안된 방법은 주효과에 대한 고려 없이 공변량과 치료 사이의 교호작용 효과 크기만 모형화하므로 초고차원의 데이터 상황이나 주효과의 영향이 크지 않은 상황에서 더 유용할 것으로 생각된다. 본 연구에서는 단일 결정 문제의 설정만을 고려하였지만 불립 함수와 후회 함수의 모형화를 통해 다단계 결정 문제로도 확장가능하다. 이러한 문제는 후속 연구를 통해 추후에 규명하고자 한다.

Appendix A: Theorem 1에 대한 증명

Proof: $G(\psi) = E[w(a, x)\{Y - \psi^T ax\}]^2$ 라 정의하고, 참 결과 모형은 $Y = f_0(x; \beta) + \psi_0^T ax$ 라 하자. 정규 조건하에서

$$\frac{\partial G}{\partial \psi} = -2E[w(a, x)\{Y - \psi^T ax\}a(x)^T] = 2E[w(a, x)a^2 x(x)^T(\psi - \psi_0)]. \quad (\text{A.1})$$

이 때, $E(\epsilon) = E[E(\epsilon | A, X)] = 0$ 이고, $f_0(x; \beta) \perp a \cdot x$ 이다. $G(\psi)$ 가 최대화하는 값을 ψ^* 라 하자. 즉, $\psi^* = \psi_0$ 일 때 $G(\psi)$ 가 최대화된다.

우리는 식 (3.4)의 손실함수 $L_n(\psi)$ 를 최소화하는 값이 $G(\psi)$ 가 최대화하는 값으로 확률적으로 수렴함을 다음을 통해 알 수 있다. $L_n(\psi)$ 를 최소화하는 값을 $\psi = \tilde{\psi}$ 라 하자. 강대수의 법칙에 의해 다음이 성립한다.

$$L_n(\psi) \xrightarrow{P} G(\psi).$$

한편 아그맥스 연속 사상 정리(argmax continuous mapping theorem)에 의해서 다음이 성립한다.

$$\tilde{\psi} \xrightarrow{P} \psi_0.$$

즉, $\tilde{\psi}$ 는 ψ_0 의 일치하는 추정량이다.

$\psi = \psi_0$ 에서 L_n 의 테일러 전개를 사용하면 우리는 다음의 식을 얻을 수 있다. 여기서 $q_n(\psi) = \partial L_n / \partial \psi (\equiv 1/n \sum_{i=1}^n q_{ni}(\psi))$ 이고

$$Q_n(\psi) = \frac{\partial^2 L_n}{\partial \psi \partial \psi^T},$$

이다.

$$L_n(\psi) = L_n(\psi_0) + q_n(\psi_0)(\psi - \psi_0) + R_n,$$

양변을 ψ 로 미분하고 $\psi = \tilde{\psi}$ 를 대입하면 다음 식을 얻을 수 있다.

$$0 = q_n(\tilde{\psi}) = q_n(\psi_0) + Q_n(\psi_0)(\tilde{\psi} - \psi_0) + R_n.$$

즉,

$$\sqrt{n}(\tilde{\psi} - \psi_0) = -Q_n^{-1}(\psi_0) [\sqrt{n}q_n(\psi_0)] + \sqrt{n}R_n.$$

정규 조건(RC1)–(RC3)하에, 우리는 $\sqrt{n}R_n \rightarrow 0$ 임을 보일 수 있다. 대수의 법칙에 의하여 $-Q_n(\psi_0) \xrightarrow{p} U$ 이다. 식 (A.1)에 따라 $E[q_n(\psi_0)] = 0$ 이므로 중심 극한 정리에 의해서,

$$\sqrt{n}q_n(\psi_0) \xrightarrow{D} N(0, \Omega).$$

슬러츠키 정리에 의해서,

$$\sqrt{n}(\tilde{\psi} - \psi_0) \xrightarrow{D} N(0, V).$$

□

Appendix B: Theorem 2에 대한 증명

Proof: Adaptive lasso 벌점 모형을 사용한 경우 추정량의 신탁 성질은 Zou (2006) 연구에서 Theorem 2의 증명에 따라 보일 수 있다. □

References

- Arjas E and Saarela O (2010). Optimal dynamic regimes: presenting a case for predictive inference, *The International Journal of Biostatistics*, **6**.
- Bellman R (1957). *Dynamic Programming*, Princeton University Press, Princeton.
- Bertsekas D and Tsitsiklis J (1996). *Neuro-Dynamic Programming*, Athena Scientific, Massachusetts.
- Breiman L, Friedman JH, Olshen RA, and Stone CJ (1984). *Classification and Regression Trees*, Wadsworth, California.
- Chakraborty B, Murphy SA, and Strecher V (2010). Inference for non-regular parameters in optimal dynamic treatment regimes, *Statistical Methods in Medical Research*, **19**, 317–343.
- Chakraborty B (2011). Dynamic treatment regimes for managing chronic health conditions: a statistical perspective, *American Journal of Public Health*, **101**, 40–45.
- Geng Y, Zhang HH, and Lu W (2015). On optimal treatment regimes selection for mean survival time, *Statistics in Medicine*, **34**, 1169–1184.
- Gunter L, Zhu J, and Murphy SA (2011). Variable selection for qualitative interactions, *Statistical Methodology*, **8**, 42–55.
- Henderson R, Ansell P, and Alshibani D (2010). Regret-regression for optimal dynamic treatment regimes, *Biometrics*, **66**, 1192–1201.
- Hernán MA, Lanoy E, Costagliola D, and Robins JM (2006). Comparison of dynamic treatment regimes via inverse probability weighting, *Basic & Clinical Pharmacology & Toxicology*, **98**, 237–242.
- Kereiakes DJ, Obenchain RL, Barber BL, et al. (2000). Abciximab provides cost-effective survival advantage in high-volume interventional practice, *American Heart Journal*, **140**, 603–610.

- Kravitz RL, Duan N, and Braslow J (2004). Evidence-based medicine, heterogeneity of treatment effects and the trouble with averages, *Milbank Quarterly*, **82**, 661–687.
- Lee BK, Lessler J, and Stuart EA (2010). Improving propensity score weighting using machine learning, *Statistics in Medicine*, **29**, 337–346.
- Longford NT and Nelder JA (1999). Statistics versus statistical science in the regulatory process, *Statistics in Medicine*, **18**, 2311–2320.
- Lu W, Zhang HH, and Zeng D (2011). Variable selection for optimal treatment decision, *Statistical Methods in Medical Research*, **22**, 493–504.
- Murphy SA (2003). Optimal dynamic treatment regimes (with discussions), *Journal of the Royal Statistical Society: Series B*, **65**, 331–366.
- Murphy SA (2005). A generalization error for q-learning, *Journal of Machine Learning Research*, **6**, 1073–1097.
- Pineau J, Bellemare MG, Rush AJ, Ghizaru A, and Murphy SA (2007). Constructing evidence-based treatment strategies using methods from computer science, *Drug and Alcohol Dependence*, **88**, 52–60.
- Polley E, LeDell E, Kennedy C, Lendle S, and van der Laan M (2016). Package ‘superlearner’, Retrieved from: <https://cran.r-project.org/web/packages/SuperLearner/SuperLearner.pdf>
- Ridgeway G, McCaffrey D, Morral A, Griffin BA, and Burgette L (2016). Twang: toolkit for weighting and analysis of nonequivalent groups.
- Robins JM, Hernán MA, and Brumback B (2000). Marginal structural models and causal inference in epidemiology, *Epidemiology*, **11**, 550–560.
- Robins JM, Orellana L, and Rotnitzky A (2008). Estimation and extrapolation of optimal treatment and testing strategies, *Statistics in Medicine*, **27**, 4678–4721.
- Robins JM (2004). Optimal structural nested models for optimal sequential decisions. In *Proceedings of the Second Seattle Symposium on Biostatistics*, Springer, New York.
- Sørensen TI (1996). Which patients may be harmed by good treatments?, *Lancet*, **348**, 351–352.
- Thall PF, Millikan RE, and Sung HG (2000). Evaluating multiple treatment courses in clinical trials, *Statistics in Medicine*, **19**, 1011–1028.
- Therneau T, Atkinson B, and Ripley B (2014). Package ‘rpart’, Retrieved from: <http://cran.r-project.org/web/packages/rpart/rpart.pdf>
- Tian L, Alizaden AA, Gentles AJ, and Tibshirani R (2014). A simple method for estimating interactions between a treatment and a large number of covariates, *Journal of the American Statistical Association*, **109**, 1517–1532.
- Tibshirani R (1996). Regression shrinkage and selection via the lasso, *Journal of the Royal Statistical Society: Series B*, **58**, 267–288.
- Van der Laan MJ and Petersen ML (2007). Causal effect models for realistic individualized treatment and intention to treat rules, *The International Journal of Biostatistics*, **3**.
- Van der Laan MJ, Polley EC, and Hubbard AE (2007). Super learner, *Statistical Applications in Genetics and Molecular Biology*, **6**.
- Wallace MP and Moodie EEM (2015). Doubly-robust dynamic treatment regimen estimation via weighted least squares, *Biometrics*, **71**, 636–644.
- Wallace MP, Moodie EEM, and Stephens DA (2017). Model validation and selection for personalized medicine using dynamic-weighted ordinary least squares, *Statistical Methods in Medical Research*, **26**, 1641–1653.
- Westreich D, Lessler J, and Funk MJ (2010). Propensity score estimation: neural networks, support vector ma-

chines, decision trees (CART) and meta-classifiers as alternatives to logistic regression, *Journal of Clinical Epidemiology*, **63**, 826–833.

Zhao P and Yu B (2006). On model selection consistency of lasso, *Journal of Machine Learning Research*, **7**, 2541–2563.

Zou H (2006). The adaptive lasso and its oracle properties, *Journal of the American Statistical Association*, **101**, 1418–1429.

Received January 15, 2021; Revised February 15, 2021; Accepted February 18, 2021

고차원 관측자료에서의 Q-학습 모형에 대한 이중강건성 연구

이호빈^a, 김예지^a, 조형준^a, 최상범^{1,a}

^a고려대학교 통계학과

요약

동적 치료 요법(dynamic treatment regimes; DTRs)은 다단계 무작위 시험에서 개인에 맞는 치료를 제공하도록 설계된 의사결정 규칙이다. 모든 개인이 동일한 유형의 치료를 처방받는 고전적인 방법과 달리 DTR은 시간이 지남에 따라 변할 수 있는 개별 특성을 고려한 환자 맞춤형 치료를 제공한다. 최적의 치료 규칙을 파악하기 위한 회귀 기반 알고리즘 중 하나인 Q-학습 방법은 쉽게 구현될 수 있기 때문에 더욱 인기를 끌고 있다. 그러나 Q-학습 알고리즘의 성능은 Q-함수를 제대로 설정했는지의 여부에 크게 의존한다. 본 논문에서는 고차원 데이터가 수집되는 DTRs 문제에 대한 다양한 이중강건 Q-학습 알고리즘을 연구하고 가중 최소제곱 추정 방법을 제안한다. 이중강건성(double-robustness)은 반응변수에 대한 모형 혹은 처리변수에 대한 모형 둘 중 하나만 제대로 설정되어도 불편추정량을 얻을 수 있음을 의미한다. 다양한 모의실험 연구를 통해 제안된 방법이 여러 시나리오 하에서도 잘 작동함을 확인하였으며 실제 데이터 예제를 통해 방법론에 대한 예시를 제시하였다.

주요용어: 이중강건성, 정밀의학, 성향점수, Q-학습, SL, 변수 선택

이 연구는 2020학년도 고려대학교 교내학술연구비 지원(K2008341)과 정부(교육부)의 재원으로 한국연구재단의 지원(2019R1F1A1052239, 2019R1A4A1028134)을 받아 수행된 기초연구사업임.

¹교신저자: (02841) 서울특별시 성북구 안암로 145, 고려대학교 통계학과. E-mail: choisang@korea.ac.kr