Minimax-Angle Learning for Optimal Treatment Decision with Heterogeneous Data

Chengchun Shi

Department of Statistics
North Carolina State University

Joint work with Wenbin Lu and Rui Song

August 3, 2016

A few words on causal inference

Data

- A: Treatment (0 or 1)
- X: Covariates
- Y: Observed outcome (usually the larger the better)

A few words on causal inference

Data

- A: Treatment (0 or 1)
- X: Covariates
- Y: Observed outcome (usually the larger the better)
- $Y^*(a)$: Potential outcome $a = 0, 1$

A few words on causal inference

Data

- A: Treatment (0 or 1)
- X: Covariates
- Y: Observed outcome (usually the larger the better)
- $Y^*(a)$: Potential outcome $a = 0, 1$

Objective

Identify the optimal regime d^{opt} to reach the best clinical outcome

A few words on causal inference

Data

- A: Treatment (0 or 1)
- X: Covariates
- Y: Observed outcome (usually the larger the better)
- $Y^*(a)$: Potential outcome $a = 0, 1$

Objective

Identify the optimal regime d^{opt} to reach the best clinical outcome

- Maximize $\mathrm{E}Y^*(d) = \mathrm{E}[d(X)Y^*(1) + \{1 - d(X)\}Y^*(0)]$

$$d : X \to \{0, 1\}.$$

Q, Contrast and Value function

- $Q(x, a) = \mathrm{E}[Y|X = x, A = a]$,

Q, Contrast and Value function

- $Q(x, a) = \mathrm{E}[Y|X = x, A = a]$,
- $C(x) = Q(x, 1) - Q(x, 0)$,

Q, Contrast and Value function

- $Q(x, a) = \mathrm{E}[Y|X = x, A = a]$,
- $C(x) = Q(x, 1) - Q(x, 0)$,
- $V(d) = \mathrm{E}Y^*(d) = \mathrm{E}[d(X)Y^*(1) + \{1 - d(X)\}Y^*(0)]$.

Q, Contrast and Value function

- $Q(x,a) = \mathrm{E}[Y|X=x, A=a]$,
- $C(x) = Q(x,1) - Q(x,0)$,
- $V(d) = \mathrm{E}Y^*(d) = \mathrm{E}[d(X)Y^*(1) + \{1-d(X)\}Y^*(0)]$.

Optimal treatment regime

- SUTVA, no unmeasured confounders, positivity assumption
- optimal treatment regime

$$d^{opt}(x) = I(C(x) > 0).$$

Q, Contrast and Value function

- $Q(x, a) = \mathrm{E}[Y|X = x, A = a]$,
- $C(x) = Q(x, 1) - Q(x, 0)$,
- $V(d) = \mathrm{E}Y^*(d) = \mathrm{E}[d(X)Y^*(1) + \{1 - d(X)\}Y^*(0)]$.

Optimal treatment regime

- SUTVA, no unmeasured confounders, positivity assumption
- optimal treatment regime

$$d^{opt}(x) = I(C(x) > 0).$$

Heterogeneity

- Optimal treatment regime (OTR): captures patient's heterogeneous response

Heterogeneity

- Optimal treatment regime (OTR): captures patient's heterogeneous response
- However, OTR may vary across patients.

Heterogeneity

- Optimal treatment regime (OTR): captures patient's heterogeneous response
- However, OTR may vary across patients.
- Data integration (Meta analysis)
  - Results combined from different studies to identify similar patterns.
  - Heterogeneity due to different populations of the data

Heterogeneity

- Optimal treatment regime (OTR): captures patient's heterogeneous response
- However, OTR may vary across patients.
- Data integration (Meta analysis)
  - Results combined from different studies to identify similar patterns.
  - Heterogeneity due to different populations of the data

Examples

- Schizophrenia study: OTR varies across patients locations

Heterogeneity

- Optimal treatment regime (OTR): captures patient's heterogeneous response
- However, OTR may vary across patients.
- Data integration (Meta analysis)
  - Results combined from different studies to identify similar patterns.
  - Heterogeneity due to different populations of the data

Examples

- Schizophrenia study: OTR varies across patients locations
- Health assessment questionnaire (HAQ) progression data: OTR varies across patients enrollment time

Schizophrenia study

- A multi-center, randomized trial with an 18 months follow-up

Schizophrenia study

- A multi-center, randomized trial with an 18 months follow-up
- Over 400 patients from three geographical locations (Manchester/Salford, Liverpool and North Nottinghamshire)

Schizophrenia study

- A multi-center, randomized trial with an 18 months follow-up
- Over 400 patients from three geographical locations (Manchester/Salford, Liverpool and North Nottinghamshire)

HAQ data

- An observational study which enrolled 847 patients enrolled from 1990 to 2000.

Schizophrenia study

- A multi-center, randomized trial with an 18 months follow-up
- Over 400 patients from three geographical locations (Manchester/Salford, Liverpool and North Nottinghamshire)

HAQ data

- An observational study which enrolled 847 patients enrolled from 1990 to 2000.
- Patients enrolled at different times showing heterogeneity; we considered three groups: 1990 - 1992 ($G = 1$); 1993 - 1996 ($G = 2$); 1997 - 2000 ($G = 3$).

How to recommend treatment rule for future patients?

How to recommend treatment rule for future patients?

- Strategy 1: recommend OTR according to patients groups

How to recommend treatment rule for future patients?

- Strategy 1: recommend OTR according to patients groups (What if the future patients don't belong to any of current groups)

How to recommend treatment rule for future patients?

- Strategy 1: recommend OTR according to patients groups (What if the future patients don't belong to any of current groups)
- Strategy 2: combine the data together and obtain OTR based on the pooled data

How to recommend treatment rule for future patients?

- Strategy 1: recommend OTR according to patients groups (What if the future patients don't belong to any of current groups)
- Strategy 2: combine the data together and obtain OTR based on the pooled data (Doesn't take population heterogeneity into account)

How to recommend treatment rule for future patients?

- Strategy 1: recommend OTR according to patients groups (What if the future patients don't belong to any of current groups)
- Strategy 2: combine the data together and obtain OTR based on the pooled data (Doesn't take population heterogeneity into account)
- Our Strategy: focus on a single treatment regime

How to recommend treatment rule for future patients?

- Strategy 1: recommend OTR according to patients groups (What if the future patients don't belong to any of current groups)
- Strategy 2: combine the data together and obtain OTR based on the pooled data (Doesn't take population heterogeneity into account)
- Our Strategy: focus on a single treatment regime that accounts for population heterogeneities.

Models

- G different population groups:

$$Y_{gj} = h_g(X_{gj}) + A_{gj}\psi_g(X_{gj}^T\beta_g) + \varepsilon_{gj}$$

- $||\beta_g||_2 = 1, g = 1, \ldots, G, \ j = 1, \ldots, m$
- h_g arbitrary baseline function
- ψ_g arbitrary monotone function
- X_{gj} mean 0, covariance matrix I.

Models

- G different population groups:

$$Y_{gj} = h_g(X_{gj}) + A_{gj}\psi_g(X_{gj}^T\beta_g) + \varepsilon_{gj}$$

- $||\beta_g||_2 = 1, g = 1, \ldots, G, j = 1, \ldots, m$
- h_g arbitrary baseline function
- ψ_g arbitrary monotone function
- X_{gj} mean 0, covariance matrix I.

Objective

- Group-wise optimal regime: $I(X_0^T \beta_g > \psi_g^{-1}(0))$

Objective

- Group-wise optimal regime: $I(X_0^T\beta_g > \psi_g^{-1}(0))$
- Overall decision: $I(X_0^T\beta_0 > c_0)$ subject to $||\beta_0||_2 = 1$

Objective

- Group-wise optimal regime: $I(X_0^T\beta_g > \psi_g^{-1}(0))$
- Overall decision: $I(X_0^T\beta_0 > c_0)$ subject to $||\beta_0||_2 = 1$
- Two steps strategy:
  - Step 1: Fix c_0, search for some $\beta_0(c_0)$ achieves some "optimality"

Objective

- Group-wise optimal regime: $I(X_0^T\beta_g > \psi_g^{-1}(0))$
- Overall decision: $I(X_0^T\beta_0 > c_0)$ subject to $||\beta_0||_2 = 1$
- Two steps strategy:
  - Step 1: Fix c_0, search for some $\beta_0(c_0)$ achieves some "optimality"
  - Step 2: Optimize over c_0

Objective

- Group-wise optimal regime: $I(X_0^T\beta_g > \psi_g^{-1}(0))$
- Overall decision: $I(X_0^T\beta_0 > c_0)$ subject to $||\beta_0||_2 = 1$
- Two steps strategy:
  - Step 1: Fix c_0, search for some $\beta_0(c_0)$ achieves some "optimality"
  - Step 2: Optimize over c_0

How to define "optimality"

- For each β, define some loss function $L_g(\beta)$ given the decision

$$I(X_0^T\beta > c_0).$$

Objective

- Group-wise optimal regime: $I(X_0^T\beta_g > \psi_g^{-1}(0))$
- Overall decision: $I(X_0^T\beta_0 > c_0)$ subject to $||\beta_0||_2 = 1$
- Two steps strategy:
  - Step 1: Fix c_0, search for some $\beta_0(c_0)$ achieves some "optimality"
  - Step 2: Optimize over c_0

How to define "optimality"

- For each β, define some loss function $L_g(\beta)$ given the decision

$$I(X_0^T\beta > c_0).$$

- Minimax effects $\beta_0 = \arg\min_\beta \max_g L_g(\beta)$

Objective

- Group-wise optimal regime: $I(X_0^T \beta_g > \psi_g^{-1}(0))$
- Overall decision: $I(X_0^T \beta_0 > c_0)$ subject to $||\beta_0||_2 = 1$
- Two steps strategy:
  - Step 1: Fix c_0, search for some $\beta_0(c_0)$ achieves some "optimality"
  - Step 2: Optimize over c_0

How to define "optimality"

- For each β, define some loss function $L_g(\beta)$ given the decision

$$I(X_0^T \beta > c_0).$$

- Minimax effects $\beta_0 = \arg\min_\beta \max_g L_g(\beta)$
  - Maximize the minimum reward
  - Minimize the risk of the worst-case scenario (minimax strategy in game theory)

How to choose loss function

How to choose loss function

Example (Error rate)

Using error rate (average percentage of making the wrong decision),

$$L_g^{(1)}(\beta) = E|I(X_g^T \beta_g > \psi_g^{-1}(0)) - I(X_g^T \beta > c_0)|,$$

The minimax effects $\beta_0^{(1)} = \arg\min_{\beta:||\beta||_2=1} \max_g L_g^{(1)}(\beta)$.

How to choose loss function

Example (Error rate)

Using error rate (average percentage of making the wrong decision),

$$L_g^{(1)}(\beta) = E|I(X_g^T\beta_g > \psi_g^{-1}(0)) - I(X_g^T\beta > c_0)|,$$

The minimax effects $\beta_0^{(1)} = \arg\min_{\beta:||\beta||_2=1} \max_g L_g^{(1)}(\beta)$.

Example (Value difference function)

Using value difference (the difference of the value under overall decision and that under optimal groupwise decision)

$$L_g^{(2)}(\beta) = EY_g^{\star}(d_g^{opt}) - EY_g^{\star}(d(X_g,\beta)),$$

where $d(X_g,\beta) = I(X_g^T\beta > c_0)$.

The minimax effects $\beta_0^{(2)} = \arg\min_{\beta:||\beta||_2=1} \max_g L_g^{(2)}(\beta)$

An intuitive definition for the minimax effects

- Assume $\psi_1^{-1}(0) = \psi_2^{-1}(0) = \cdots = \psi_G^{-1}(0) = \bar{c}$, for each subgroup g, the optimal regime becomes

$$I(X_0^T \beta_g > \bar{c}),$$

An intuitive definition for the minimax effects

- Assume $\psi_1^{-1}(0) = \psi_2^{-1}(0) = \cdots = \psi_G^{-1}(0) = \bar{c}$, for each subgroup g, the optimal regime becomes

$$I(X_0^T \beta_g > \bar{c}),$$

- Note that $||\beta_g||_2 = 1$, each β_g represents the "direction".

An intuitive definition for the minimax effects

- Assume $\psi_1^{-1}(0) = \psi_2^{-1}(0) = \cdots = \psi_G^{-1}(0) = \bar{c}$, for each subgroup g, the optimal regime becomes

$$I(X_0^T \beta_g > \bar{c}),$$

- Note that $||\beta_g||_2 = 1$, each β_g represents the "direction".
- Intuitively, we can define the minimax effects through "angles":

$$\beta_0^{(3)} = \arg \min_{||\beta||_2=1} \max_g \angle(\beta, \beta_g).$$

An intuitive definition for the minimax effects

- Assume $\psi_1^{-1}(0) = \psi_2^{-1}(0) = \cdots = \psi_G^{-1}(0) = \bar{c}$, for each subgroup g, the optimal regime becomes

$$I(X_0^T \beta_g > \bar{c}),$$

- Note that $||\beta_g||_2 = 1$, each β_g represents the "direction".
- Intuitively, we can define the minimax effects through "angles":

$$\beta_0^{(3)} = \arg \min_{||\beta||_2=1} \max_g \angle(\beta, \beta_g).$$

- More formally, let

$$F(\beta) = \min_g \beta^T \beta_g,$$

and $\beta_0^{(3)}$ is defined as $\arg\max_{||\beta||_2=1} F(\beta)$ (Maximin correlation approach Avi-Itzhak et al., 1995).

Theorem (Equivalence of $\beta_0^{(1)}$ and $\beta_0^{(3)}$)

Assume $\psi_1^{-1}(0) = \psi_2^{-1}(0) = \cdots = \psi_G^{-1}(0) = \bar{c}$, each X_{ij} i.i.d spherically distributed, then for any c_0,

$$\beta_0^{(3)} = \beta_0^{(1)}.$$

Theorem (Equivalence of $\beta_0^{(1)}$ and $\beta_0^{(3)}$)

Assume $\psi_1^{-1}(0) = \psi_2^{-1}(0) = \cdots = \psi_G^{-1}(0) = \bar{c}$*, each* X_{ij} *i.i.d spherically distributed, then for any* c_0*,*

$$\beta_0^{(3)} = \beta_0^{(1)}.$$

Theorem (Equivalence of $\beta_0^{(2)}$ and $\beta_0^{(3)}$)

Assume $\psi_1 = \psi_2 = \cdots = \psi_G = \psi$*, each* X_{ij} *i.i.d spherically distributed, then for any* c_0*,*

$$\beta_0^{(3)} = \beta_0^{(2)}.$$

Only need to focus on the third definition !!

Refinement

- $\beta_0^{(3)} = \arg\max_{||\beta||_2=1} F(\beta)$
- $\beta_0^{(3)}$ always exists: the minimax effects is well defined

Refinement

- $\beta_0^{(3)} = \arg\max_{||\beta||_2=1} F(\beta)$
- $\beta_0^{(3)}$ always exists: the minimax effects is well defined
- May not be unique when $F_0 = \max_{||\beta||_2=1} F(\beta) < 0$
- The optimization problem

$$\arg\max_{||\beta||_2=1} F(\beta),$$

is a quasi-concave problem (difficult to solve globally).

Refinement

- $\beta_0^{(3)} = \arg\max_{||\beta||_2=1} F(\beta)$
- $\beta_0^{(3)}$ always exists: the minimax effects is well defined
- May not be unique when $F_0 = \max_{||\beta||_2=1} F(\beta) < 0$
- The optimization problem

$$\arg\max_{||\beta||_2=1} F(\beta),$$

  is a quasi-concave problem (difficult to solve globally).
- Consider $\beta_0^{(4)} = \arg\max_{||\beta||_2\leq 1} F(\beta)$,

Refinement

- $\beta_0^{(3)} = \arg\max_{||\beta||_2=1} F(\beta)$
- $\beta_0^{(3)}$ always exists: the minimax effects is well defined
- May not be unique when $F_0 = \max_{||\beta||_2=1} F(\beta) < 0$
- The optimization problem

$$\arg \max_{||\beta||_2=1} F(\beta),$$

  is a quasi-concave problem (difficult to solve globally).
- Consider $\beta_0^{(4)} = \arg\max_{||\beta||_2\leq 1} F(\beta)$,
- Solving $\beta_0^{(4)}$ is a tractable concave programming (Seung-Jean et al., 2008).

Refinement

- $\beta_0^{(3)} = \arg\max_{||\beta||_2=1} F(\beta)$
- $\beta_0^{(3)}$ always exists: the minimax effects is well defined
- May not be unique when $F_0 = \max_{||\beta||_2=1} F(\beta) < 0$
- The optimization problem

$$\arg\max_{||\beta||_2=1} F(\beta),$$

  is a quasi-concave problem (difficult to solve globally).
- Consider $\beta_0^{(4)} = \arg\max_{||\beta||_2\leq 1} F(\beta)$,
- Solving $\beta_0^{(4)}$ is a tractable concave programming (Seung-Jean et al., 2008).
- $\beta_0^{(4)}$ always exists, and is unique when $F_0 \neq 0$.

Estimating procedure

- Assume estimators $\hat{\beta}_1, \ldots, \hat{\beta}_G$ are available with $||\hat{\beta}_g||_2 = 1$ for any g.
- Concave optimization problem

$$\hat{\beta}_0 = \arg \max_{\beta:||\beta||_2 \le 1} \min_{g=1,\ldots,G} \beta^T \hat{\beta}_g.$$

- Equivalent to QCLP:

$$\begin{aligned} \text{maximize} \quad & t \in \mathbb{R} \\ \text{subject to} \quad & \beta^T \hat{\beta}_g \ge t, g = 1, \ldots, G \\ & \beta^T \beta \le 1, \end{aligned}$$

- Obtain $\hat{c}_0$ by maximizing IPWE (AIPWE):

$$\hat{c}_0 = \arg \max_c \frac{1}{mG} \sum_i \sum_j \frac{Y_{ij} I(X_{ij}^T \hat{\beta}_0 > c)}{A_i \hat{\pi}_i + (1 - A_i)(1 - \hat{\pi}_i)}$$

Theorem (Consistency)

Under certain conditions, if $F_0 \neq 0$, then with probability goes to 1, the estimated minimax effects $\hat{\beta}_0 = 0$. If $F_0 > 0$, then

$$||\hat{\beta}_0 - \beta_0||_2 = \sup_{g \in T_0} O(||\hat{\beta}_g - \beta_g||_2)$$

Theorem (Consistency)

Under certain conditions, if $F_0 \neq 0$, then with probability goes to 1, the estimated minimax effects $\hat{\beta}_0 = 0$. If $F_0 > 0$, then

$$||\hat{\beta}_0 - \beta_0||_2 = \sup_{g \in T_0} O(||\hat{\beta}_g - \beta_g||_2)$$

Theorem (Asymptotic normality)

Under conditions in Theorem 3, if $F_0 > 0$,

$$\sqrt{m}(\hat{\beta}_g - \beta_g) = \frac{1}{\sqrt{m}} \sum_{i=1}^{m} \psi_{ig} + o_p(1),$$

with $\Sigma_g = E\psi_g \psi_g^T$, $\max_{j=1,\dots,s} |\psi_{ig}^j|^3 < \infty$, then $\sqrt{m}(\hat{\beta}_0 - \beta_0)$ is asymptotically normally distributed with mean 0, and some covariance matrix Σ_0.

Over all value function

For each threshold c, we define the overall value function under the regime $I(x^T\beta_0 > c)$ as

$$V(\beta_0, c) = \frac{1}{G}\sum_{g=1}^{G} \mathrm{E}(h_g(X_{g0}) + \psi(X_{g0}^T\beta_0)I(X_{g0}^T\beta_0 > c)),$$

and denote c_0 to be the arg max of $V(\beta_0, c)$ over c.

Theorem

Under certain regularity conditions, we have

$$\hat{c}_0 - c_0 = O_p(m^{-1/3}).$$

Moreover, $\sqrt{m}(\hat{V}_m(\hat{\beta}_0, \hat{c}_0) - V(\beta_0, c_0))$ *is asymptotically normal with mean* 0*, variance* v_0^2*.*

Simulation studies

- Generate models from 4 different groups and estimate OTR for each group using A-learning estimating equation.

Simulation studies

- Generate models from 4 different groups and estimate OTR for each group using A-learning estimating equation.
- Compare the pooled treatment regime $d^P(x) = I(x^T\hat{\beta}^P > \hat{c}^P)$ with minimax treatment regime $d^M(x) = I(x^T\hat{\beta}^M > \hat{c}^M)$.

Simulation studies

- Generate models from 4 different groups and estimate OTR for each group using A-learning estimating equation.
- Compare the pooled treatment regime $d^P(x) = I(x^T \hat{\beta}^P > \hat{c}^P)$ with minimax treatment regime $d^M(x) = I(x^T \hat{\beta}^M > \hat{c}^M)$.
- Leave one group out cross validation procedure: obtain $d^P(x)$ and $d^M(x)$ based on any 3 groups of patients and evaluate the error rate, value function on the remaining group.

Thank you!

Simulation setting

- Three groups of patients, each generated according as

$$Y_{gj} = h(X_{gj}) + 2A_{gj}X_{gj}^T\beta_g + \varepsilon_{gj},$$

$X_{gj} \overset{i.i.d}{\sim} N(0, I_4)$ and $\varepsilon_{gj} \overset{i.i.d}{\sim} N(0, 0.25)$.

- Two baseline functions for h: linear and nonlinear.
- Two propensity score models for π: constant and probit.
- Four settings: subgroup estimator obtained using A-learning based on a linear model for h and logistic model for π:

1. S1: π correct, h correct,
2. S2: π correct, h wrong,
3. S3: π wrong, h correct,
4. S4: π wrong, h wrong.

Simulation setting (Continued)

- Two scenarios for subgroup parameters (representing different degrees of heterogeneity):
  - (large heterogeneity) $\beta_1 = (1, 0)$, $\beta_2 = (\cos(10^\circ), \sin(10^\circ))$, $\beta_3 = (\cos(70^\circ), \sin(70^\circ))$, $\beta_4 = (0, 1)$;
  - (small heterogeneity) $\beta_1 = (\cos(30^\circ), \sin(30^\circ))$, $\beta_2 = (\cos(45^\circ), \sin(45^\circ))$, $\beta_3 = (\cos(54^\circ), \sin(54^\circ))$, $\beta_4 = (\cos(60^\circ), \sin(60^\circ))$.
- $\beta_0 = (\cos(45^\circ), \sin(45^\circ))$ for both scenarios.

Table: Bias, standard deviation (in parenthesis) of $\hat{\beta}_0$ and coverage probability for confidence intervals of β_0.

		$\hat{\beta}_0^{(1)}$	$\hat{\beta}_0^{(2)}$	CP for $\hat{\beta}_0^{(1)}$	CP for $\hat{\beta}_0^{(2)}$
Sce 1	S1	$-0.018(0.037)$	$-0.028(0.051)$	96.6%	97.6%
	S2	$-0.015(0.045)$	$-0.025(0.053)$	97.6%	96.8%
	S3	$-0.016(0.048)$	$-0.024(0.055)$	97.2%	95.2%
	S4	$-0.010(0.061)$	$-0.020(0.069)$	98.8%	98.0%
Sce 2	S1	$3.6 \times 10^{-4}(0.018)$	$-0.001(0.018)$	96.0%	95.0%
	S2	$-0.006(0.033)$	$0.003(0.031)$	96.8%	96.8%
	S3	$-0.008(0.045)$	$0.002(0.042)$	96.6%	97.4%
	S4	$-0.012(0.064)$	$0.004(0.063)$	96.6%	97.8%

Table: Bias, standard deviation of $\hat{V}_m(\hat{\beta}^M, \hat{c}^M)$ and coverage probability for confidence intervals of $V(\beta^M, c^M)$

Scenario 1	Bias	SD	CI	Scenario 2	Bias	SD	CI
Setting 1	0.017	0.083	95.6%	Setting 1	0.007	0.099	95.4%
Setting 2	0.018	0.074	95.6%	Setting 2	0.005	0.075	95.2%
Setting 3	0.018	0.134	93.6%	Setting 3	0.011	0.101	95.2%
Setting 4	0.027	0.137	93.0%	Setting 4	0.003	0.115	95.2%

Comparison with simple method

- Methods to compare:
  - minimax treatment regime: $d^M(x) = I(x^T \hat{\beta}^M > \hat{c}^M)$
  - pooled treatment regime: $d^P(x) = I(x^T \hat{\beta}^P > \hat{c}^P)$
- Evaluation
  - obtain the estimated regime based on three groups and apply it to the remaining group;
  - compute error rate (using the estimated group-specific regime as the truth) and estimated value function (using A-learning) of the estimated regime for each group

Table: Groupwise and overall error rate and value function (in parenthesis) for the first scenario under estimated minimax OTR and pooled OTR

Testing group		First group	Second group	Third group	Fourth group
Setting 1	pooled	32.2%(1.42)	25.1%(1.56)	21.9%(1.61)	35.7%(1.35)
	minimax	28.3%(1.50)	20.2%(1.64)	15.0%(1.71)	31.1%(1.45)
Setting 2	pooled	32.2%(1.42)	25.2%(1.56)	21.7%(1.62)	35.8%(1.34)
	minimax	28.0%(1.51)	20.2%(1.64)	15.5%(1.71)	31.2%(1.44)
Setting 3	pooled	32.0%(1.43)	25.1%(1.56)	21.9%(1.61)	36.1%(1.34)
	minimax	28.7%(1.49)	21.0%(1.62)	16.2%(1.69)	31.4%(1.44)
Setting 4	pooled	32.0%(1.42)	25.2%(1.55)	22.0%(1.61)	35.9%(1.34)
	minimax	28.7%(1.49)	21.0%(1.63)	16.3%(1.69)	31.8%(1.43)

Table: Groupwise and overall error rate and value function (in parenthesis) for the second scenario under estimated minimax OTR and pooled OTR

Testing group		First group	Second group	Third group	Overall
Setting 1	pooled	12.8%(1.73)	2.0%(1.80)	5.0%(1.79)	9.5%(1.76)
	minimax	13.0%(1.73)	3.3%(1.79)	6.0%(1.78)	10.6%(1.75)
Setting 2	pooled	12.8%(1.73)	2.3%(1.80)	5.2%(1.79)	9.6%(1.76)
	minimax	13.1%(1.73)	3.7%(1.79)	6.2%(1.78)	10.7%(1.75)
Setting 3	pooled	12.9%(1.73)	2.5%(1.79)	4.9%(1.79)	9.3%(1.76)
	minimax	13.3%(1.73)	4.4%(1.79)	6.6%(1.78)	10.6%(1.75)
Setting 4	pooled	13.0%(1.73)	3.4%(1.79)	5.6%(1.78)	9.5%(1.76)
	minimax	14.2%(1.71)	5.6%(1.78)	7.7%(1.77)	11.1%(1.75)

Reference I

Avi-Itzhak, H., Van Mieghem, J. A., Rub, L., et al. (1995). Multiple subclass pattern recognition: a maximin correlation approach. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 17(4):418–431.

Chakraborty, B., Murphy, S., and Strecher, V. (2010). Inference for non-regular parameters in optimal dynamic treatment regimes. *Stat. Methods Med. Res.*, 19(3):317–343.

Murphy, S. A. (2003). Optimal dynamic treatment regimes. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 65(2):331–366.

Robins, J. M. (2004). Optimal structural nested models for optimal sequential decisions. In *Proceedings of the Second Seattle Symposium in Biostatistics*, volume 179 of *Lecture Notes in Statist.*, pages 189–326. Springer, New York.

Reference II

Seung-Jean, K., Almir, M., and Stephen, B. (2008). Maximin correlation. *Technical report*.

Watkins, C. and Dayan, P. (1992). Q-learning. *Mach. Learn.*, 8:279–292.

Zhang, B., Tsiatis, A. A., Laber, E. B., and Davidian, M. (2012). A robust method for estimating optimal treatment regimes. *Biometrics*, 68(4):1010–1018.

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