UNC GILLINGS SCHOOL OF GLOBAL PUBLIC HEALTH

PHAIR

BIOS 740 Review/Preview
Intro to Precision Medicine

Yang Ruan, Anthony Wang

February 16th 2024

Content

Motivating Example

Sepsis and Mechanical Ventilation

- Sepsis
  - A life-threatening condition where the body reacts extremely to infection by damaging its own tissues and organs.
  - Every year, around 1.7 million people develop Sepsis. Every 1 in 3 patients who died in hospitals had Sepsis.
- Mechanical Ventilation (MV)
  - A treatment that generally brings improvement to Sepsis patients' outcomes (survival).[a]
  - Risk of death increases for Sepsis patients with lung-related complications given MV.
  - "Non-invasive" vs. "Invasive"

[a] Jeffrey E Gotts and Michael A Matthay. "Sepsis: pathophysiology and clinical management". In: *BMJ* (2016).

Content

Introduction

What & Why PM

- Definition of Precision Medicine
  - Precision Medicine is an innovative approach that takes into account individual differences in patients' genes, environments, and lifestyles.
  - Precision medicine refers to "the tailoring of medical treatment to the individual characteristics of each patient."
- Why Precision Medicine
  - Heterogeneity across patients:
    what works for one may not work for another.
  - Temporal variability within a patient:
    what works now may not work later.
  - ...

Introduction

Outline of Overall Pipeline[1]

- Observed data: $\{(X_i, A_i, Y_i)\}_{i=1}^n$
  - Tailoring Variable(s) $X \in \mathcal{X}$: baseline patient characteristics.
  - Assigned Treatment $A \in \mathcal{A}$
  - Outcomes (or utilities) $Y \in \mathbb{R}$: outcome coded so that higher values are better.
- Dynamic Treatment Regime (DTR): a map $d : \mathcal{X} \to \mathcal{A}$
  - Single decision: make a single recommendation for treatment in the middle of the trial (so still dynamic).
  - Multiple decision: make a series of interdependent recommendations depending on the intermediate outcomes for different individuals.

Goal:

- **Estimate the optimal treatment regime.**

[1] Michael R Kosorok and Eric B Laber. "Precision medicine". In: *Annual review of statistics and its application* 6 (2019), pp. 263–286.

Content

Potential Outcomes Framework Pt.1

Single Decision, Finite Horizon Setting

- X: Covariates - lung-disease indicator (Binary: 1 for positive; 0 for negative)
- A: Treatment - MV (Binary: 1 for invasive; -1 for non-invasive).
- Y: Observed Outcome - patient survival 90 days following treatment (1 for alive; 0 for dead)
- $Y^*(a)$: Potential outcome if patient is given treatment $A = a$.
- E.g. For patient i given treatment $A_i = 1$, $Y_i^*(1) = Y_i$ are observed, $Y_i^*(-1)$ can be estimated.

Potential Outcomes Framework Pt.2

Define Optimal DTR

- Recall goal of precision medicine: Estimate d^{opt}
- Potential Outcome under DTR d:

$$Y^*(d(X)) = \sum_{a \in \mathcal{A}} Y^*(a) I_{d(X)=a}$$

- For DTR to be optimal, it satisfies:

$$E[Y^*(d^{\text{opt}})] \geq E[Y^*(d)]$$

for all d such that $d(x) \in \psi(x)$ and $x \in \mathcal{X}$, where $\psi(x)$ is the set of allowable treatments the patient can receive given its X.

Potential Outcomes Framework Pt.3

CATE: Conditional Average Treatment Effect

- Subgroup Identification: d^{opt} can identify which population subgroup should be given each treatment to optimize patient outcome, i.e.
  - Patients with $\{x_i : d^{\text{opt}}(x_i) = 1\}$ will be given invasive MV;
  - Patients with $\{x_i : d^{\text{opt}}(x_i) = -1\}$ will be given non-invasive MV.
- CATE

$$\Delta(x) = E[Y^*(1) - Y^*(-1)|X = x]$$

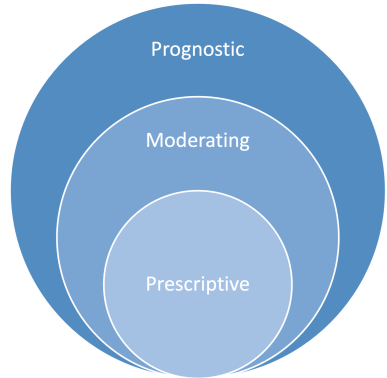
It can be deduced that $d^{\text{opt}}(x) = \text{sign}[\Delta(x)]$

Content

Biomarkers Pt.1

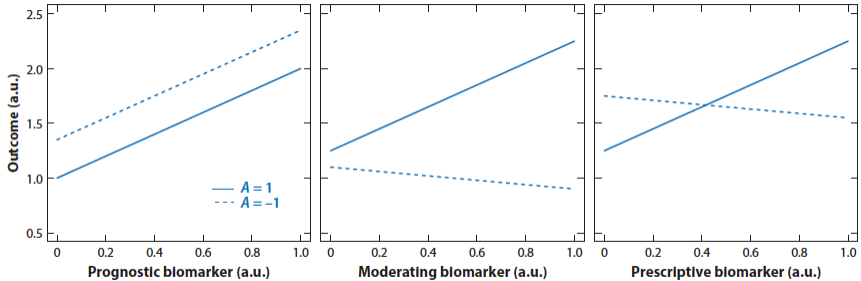
Why, What, and How (they relate and compare)

- Why: Precision medicine clinical goal - identifying biomarkers is important for finding optimal treatment.
- What: Scalar feature constructed from patient data.
- How: Three levels of specification for biomarkers.



Biomarkers Pt.2

Prognostic, Moderating, Prescriptive



- Prognostic biomarkers: predict the mean outcome of a patient
- Moderating biomarkers: predict the difference in mean outcomes across different treatments
- Prescriptive biomarkers: select the treatment that maximizes the mean outcome

Content

Estimating d^{opt} Pt.1

Causal Assumptions

1. Causal Consistency

$$Y = Y^*(A) \tag{1}$$

2. Positivity: Propensity score for each patient is greater than 0

$$P[A = a | X = x] > 0, \forall a \in \mathcal{A}, x \in \mathcal{X} \tag{2}$$

3. No Unmeasured Confounder/Strong Ignorability/(Full) Conditional Exchangeability

$$\{Y^*(a) : a \in \mathcal{A}\} \perp A | X \tag{3}$$

4. SUTVA: Stable Unit Treatment Value Assumption
   - No interference between patients.
   - Treatments are well defined. No multiple levels of a single treatment.

Estimating d^{opt} Pt.2

Regression-Based Estimation

- Define $Q(x, a) = E[Y|X = x, A = a]$
- Under the assumptions, we have $d^{\text{opt}} = \text{argmax}_{a \in \psi(x)} Q(x, a)$
- Regression-based estimation:
  - Construct estimator of $Q(x, a)$: $\widehat{Q}_n(x, a)$
  - Use plug-in estimator: $\widehat{d}_n = \text{argmax}_{a \in \psi(x)} \widehat{Q}_n(x, a)$
- Example: Linear Models
  - Define $Q(x, a; \beta) = \sum_{a' \in A} x_{a'} \beta_{a'} 1\{a = a'\}$
  - For the empirical measure $\mathbb{P}_n$, let $\widehat{\beta}_n = \text{argmin}_\beta \mathbb{P}_n\{Y - Q(X, A; \beta)\}^2$
  - Then $\widehat{d}_n = \text{argmax}_{a \in \psi(x)} Q(x, a; \widehat{\beta})$

Content

Q Learning

Multi-decision Setting[2]

- Recall that the observed data is of the form $\{(X_{t,i}, A_{t,i}, Y_{t,i})\}_{i=1}^{n}$. At each decision point $t = 1, \ldots, T$, assume that there is a finite set of all possible treatment options $\mathcal{A}_t$ with elements $A_t \in \mathcal{A}_t$.
- Let Y_T be the proximal outcome measured after the treatment at stage T.
- Denotes H_t as the set of available patient history at time t such that
  - $H_1 = X_1$
  - $H_t = (H_{t-1}, A_{t-1}, Y_{t-1}, X_t)$
- A dynamic Treatment Regime is a sequence of functions $\boldsymbol{d} = (d_1, \ldots, d_T)$ such that $d_t : \mathcal{H}_t \to \mathcal{A}_t$ for $t = 1, \ldots, T$.
- An optimal treatment regime maximizes the expectation of some (prespecified) cumulative outcome measure $Y = y(Y_1, \ldots, Y_T)$, e.g., $y(v_1, \ldots, v_T) = \sum_{t=1}^{T} v_t$, or $y(v_1, \ldots, v_T) = \max_t v_t$, or $y(v_1, \ldots, v_T) = v_T$.

[2] Phillip J Schulte et al. "Q-and A-learning methods for estimating optimal dynamic treatment regimes". In: *Statistical science: a review journal of the Institute of Mathematical Statistics* 29.4 (2014), p. 640.

Q Learning

Optimal Treatment Regime[4]

- Under these assumptions, we can express the optimal regimes in terms of the observed data. We now define the following:

$$Q_T(h_T, a_T) = E(Y_T \mid H_T = h_T, A_T = a_T)$$
$$V_T(h_T, a_T) = \max_{a_T} Q_T(h_T, a_T)$$

- and for $t = T-1, \ldots, 1$,

$$Q_t(h_t, a_t) = E(V_{t+1}(h_{t+1}, a_t) \mid H_t = h_t, A_t = a_t)$$
$$V_t(h_t, a_t) = \max_{a_t} Q_t(h_t, a_t)$$

- The optimal DTRs is:

$$d_t^{\text{opt}}(h_t) = \arg\max_{a_t} Q_t(h_t, a_t), \qquad \text{for} \quad t = 1, \ldots, T \tag{4}$$

- Q-learning is an approximate dynamic programming[3] algorithm based on (4). This immediately suggests a regression-based estimator $\widehat{Q}_{t,n}(h_t, a_t, \xi_t)$ of $Q_{t,n}(h_t, a_t)$ by regressing Y on H_t and A_t, where ξ_t is the parameters for estimating $\widehat{Q}_{t,n}$.

[3] Richard Bellman. "Dynamic programming". In: *Science* 153.3731 (1966), pp. 34–37.

[4] Schulte et al., "Q-and A-learning methods for estimating optimal dynamic treatment regimes".

Q Learning

Q-learning for Two Stages[5]

- Considering DTR with only two stages. We may fit linear models for

$$Q_1(h_1, a_1; \xi_1) = \mathcal{H}_1^T \beta_1 + a_1 \left(\mathcal{H}_1^T \psi_1\right)$$

$$Q_2(h_2, a_2; \xi_2) = \mathcal{H}_2^T \beta_2 + a_2 \left(\mathcal{H}_2^T \psi_2\right)$$

- where

$$\mathcal{H}_1 = (1, x_1^T)^T \quad \mathcal{H}_2 = (1, x_1^T, a_1, x_2^T)^T$$

$$\xi_t = (\beta_t^T, \psi_t^T)^T \quad t = 1, 2$$

- Here $Q_2(h_2, a_2; \xi_2)$ is a model for $E(Y \mid H_2 = h_2, A_2 = a_2)$, a standard regression problem involving observable data, whereas $Q_1(s_1, a_1; \xi_1)$ is a model for $E(V_2(h_2, a_1) \mid H_1 = h_1, A_1 = a_1)$

- The corresponding V-functions are

$$V_2(h_2, a_2; \xi_2) = \max_{a_2 \in \{-1,1\}} Q_2(h_2, a_2; \xi_2)$$

$$= \mathcal{H}_2^T \beta_2 + \left(\mathcal{H}_2^T \psi_2\right) \times \text{sign}\left(\mathcal{H}_2^T \psi_2\right), \text{ and}$$

$$V_1(s_1; \xi_1) = \max_{a \in \{-1,1\}} Q_1(s_1, a_1; \xi_1)$$

$$= \mathcal{H}_1^T \beta_1 + \left(\mathcal{H}_1^T \psi_1\right) \text{sign}\left(\mathcal{H}_1^T \psi_1\right)$$

- We can see that

$$d_1^{\text{opt}}(h_1; \xi_1) = \text{sign}\left(\mathcal{H}_1^T \psi_1\right)$$

$$d_2^{\text{opt}}(h_2, a_1; \xi_2) = \text{sign}\left(\mathcal{H}_2^T \psi_2\right)$$

- We can see that we only need to estimate the regression coefficients ψ_1 and ψ_2, which can be done via OLS and WLS, etc.

[5] Schulte et al., "Q-and A-learning methods for estimating optimal dynamic treatment regimes".

Content

Outcome Weighted Learning

Individualized Treatment Rules[6]

- Outcome weighted learning (OWL) directly estimates the decision rule that maximizes clinical response by formulating individualized treatment rule (ITR) $d(X) : \mathbb{R}^d \to \mathcal{A}$ estimation as a weighted classification problem
- The data consist of the triple (X, A, Y)
  - Prognostic variables, $X = (X_1, \ldots, X_d)^T \in \mathcal{X}$
  - Treatment $A \in \mathcal{A} = \{-1, 1\}$
  - Clinical outcome $Y > 0$ is bounded
- Before maximizing $d(X)$, we need to measure how well a given rule will perform
- Let P denote the distribution of (X, A, Y). Let P^d denote the distribution of (X, A, Y) given that $A = d(X)$
- We define the value function $\mathcal{V}(d)$ to be

$$\mathcal{V}(d) = E^d(Y) = \int R \frac{dP^d}{dP} dP = E\left[\frac{I(A = d(X))}{P(A \mid X)} Y\right]$$

[6] Yingqi Zhao et al. "Estimating individualized treatment rules using outcome weighted learning". In: *Journal of the American Statistical Association* 107.499 (2012), pp. 1106–1118.

Outcome Weighted Learning

Individualized Treatment Rules

- An optimal ITR is a rule that maximizes $\mathcal{V}(d)$

$$d^* = \arg\max_d E\left[\frac{I(A = d(X))}{P(A \mid X)} Y\right]$$

- Comparing the Regression-based method and OWL

$$\begin{cases} (X, A, Y) \xrightarrow[\text{Minimize Prediction Error}]{} \text{Predict} E(Y \mid A, X) \xrightarrow[\arg\max_A \widehat{E}(Y|A,X)]{} \text{Optimal ITR} \\ (X, A, Y) \xrightarrow[\text{Maximize} \mathcal{V}(\text{d})]{} \text{Optimal ITR} \end{cases}$$

- Notice that maximizing the value $\mathcal{V}(d)$, is equivalent to minimizing the risk of misclassification

Maximize the value

$$E\left[\frac{I(A = D(X))}{P(A|X)} Y\right]$$

Minimize the risk

$$E\left[\frac{I(A \neq D(X))}{P(A|X)} Y\right]$$

- We can derive the optimal ITR by solving a classification problem in machine learning approach.

Outcome Weighted Learning

Individualized Treatment Rules

- For any rule d, we may posit a model $d(X) = \text{sign}(f(X))$
- Empirical approximation to the risk function would be

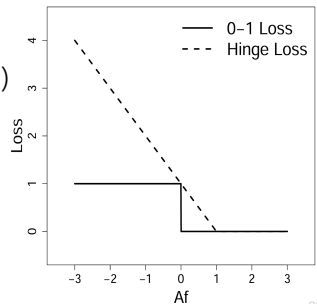
$$n^{-1} \sum_{i=1}^{n} \frac{Y_i}{P(A_i \mid X_i)} I(A_i \neq \text{sign}(f(X_i))$$

- Computation challenges: non-convexity and discontinuity of 0-1 loss.
  - If we replace the 0-1 loss with a hinge loss the problem becomes convex

$$n^{-1} \sum_{i=1}^{n} \frac{Y_i}{P(A_i \mid X_i)} \phi\left(A_i f(X_i)\right) + \lambda_n \|f\|^2 \qquad (5)$$

where $\phi(x) = \max(1 - x, 0)$, $\|f\|$ is some norm of f, and λ_n controls the severity of the penalty on the function.

  - The estimated optimal ITR is $\widehat{d}_n = \text{sign}(\widehat{f}_n(X))$ where $\widehat{f}_n$ minimizes (5)



Outcome Weighted Learning

Weighted Support Vector Machine[7]

- The modified classification error in (5) now looks like a weighted version of a support vector machine (SVM)
- If we assume that the decision function $f(x)$ is a linear function $f(x) = \langle \Psi, X \rangle + \Psi_0$, with $\|f\|$ as the Euclidean norm of Ψ, then (5) is equivalent to

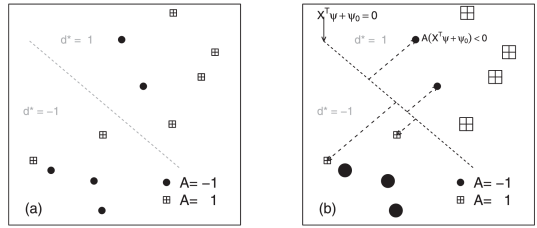


Figure 8.1. *Weighted classification idea in outcome weighted learning. (a) No weighting. (b) Outcome weighting. Symbol size is proportional to weight.*

$$\min \frac{1}{2}\|\Psi\|^2 + \kappa \sum_{i=1}^{n} \frac{Y_i}{P(A_i \mid X_i)} \xi_i \qquad (6)$$

- subject to $A_i\left(\langle \beta, X \rangle + \beta_0\right) \geq (1 - \xi_i)\,, \quad \xi_i > 0$, where ξ_i is the slack variable for subject i to allow a small portion of wrong classification.

[7] Michael R Kosorok and Erica EM Moodie. *Adaptive treatment strategies in practice: planning trials and analyzing data for personalized medicine.* SIAM, 2015.

Outcome Weighted Learning

Optimization & Kernel[8]

- As with standard SVM, we can find a dual problem of (6) as a convex optimization problem which can be solved by quadratic programming

$$\max_{\alpha} \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j A_i A_j \langle X_i, X_j \rangle$$

$$\text{subject} \quad \text{to} \quad 0 \leq \alpha_i \leq \kappa Y_i / \pi_i \quad \text{and} \quad \sum_{i=1}^{n} \alpha_i A_i = 0$$

- However, linear decision rules may be insufficient. By replacing $\langle X_i, X_j \rangle$ with a kernel $\mathcal{K}(X_i, X_j) : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$, we can fit non-linear decision boundaries.

- Zhao et al. were able to show that f which minimizes

Minimize the risk $E\left[\frac{I(A \neq \text{sign}(f(X)))}{P(A|X)} Y\right]$ is the same f which minimizes Minimize the ϕ -risk $E\left[\frac{\phi(Af(X))}{P(A|X)} Y\right]$

[8] Zhao et al., "Estimating individualized treatment rules using outcome weighted learning".

Thank you!

You can contact us at:
yangruan@unc.edu & andong@unc.edu