Introduction to Dynamic Treatment Regimes
Homework 3, Spring 2019

1. Consider the HIV study data you worked with in Problems 2 and 3 of Homework 2, with $K = 4$. Recall the treatment regimes d_η with rules of the form, using the notation defined in Homework 2,

$$\begin{aligned} d_k(\overline{x}_k, \overline{a}_k; \eta) &= 1 \text{ if } c_k < \log(\eta) \text{ and } r_k = 0 \\ &= 0 \text{ otherwise,} \end{aligned} \tag{1}$$

   $k = 1, \ldots, 4$, for a fixed choice of the (common across decision points) threshold η, where we now make the dependence on η explicit. As noted in Problem 2 of Homework 2, at Decision k, the rule dictates administering therapy at Decision k if the current CD4 count is $< \eta$ as long as the individual's virus is not yet resistant; otherwise, do not administer therapy, so that a regime with rules (1) administers therapy only if CD4 count becomes sufficiently low (with no resistance), reflected by the choice of the threshold η.

   In Problems 2 and 3 of Homework 2, you estimated the value $\mathcal{V}(d_\eta)$ of fixed regimes d_η with rules of the form (1) for $\eta = 100, 200, 300$, and 400 cells/mm^3 using the g-computation algorithm and the IPW estimators $\widehat{\mathcal{V}}_{IPW}(d_\eta)$ and $\widehat{\mathcal{V}}_{IPW*}(d_\eta)$ and in each case obtained standard errors using a nonparametric bootstrap. In this problem, you will estimate $\mathcal{V}(d_\eta)$ for each of these fixed regimes d_η based on a marginal structural model (MSM) for the value as in (5.38) on Slide 312 of the form

$$\mu(\eta; \alpha) = \alpha_1 + \alpha_2 \log(\eta) + \alpha_3 \{\log(\eta)\}^2.$$

   (a) Estimate $\alpha = (\alpha_1, \alpha_2, \alpha_3)^T$ by solving an estimating equation like that on Slide 316.

   - Use the fitted propensity models you developed in Problem 3 of Homework 2
   - Choose the range of η values and the partition of it $\eta_{(j)}$ as you see fit
   - Take the weights $w(\eta_{(j)} \equiv 1$.

   (b) Using your fitted MSM, estimate the value $\mathcal{V}(d_\eta)$ for regimes with rules of the form (1) for $\eta = 100, 200, 300$, and 400. Obtain standard errors using a nonparametric bootstrap. (Note: you will have to refit the propensity models for each bootstrap data set just as you did for $\widehat{\mathcal{V}}_{IPW}(d_\eta)$ and $\widehat{\mathcal{V}}_{IPW*}(d_\eta)$ in Problem 3 of Homework 2).

   (c) Comment on how the results you obtain in (b) compare to those you obtained using the g-computation algorithm, $\widehat{\mathcal{V}}_{IPW}(d_\eta)$, and $\widehat{\mathcal{V}}_{IPW*}(d_\eta)$ in Homework 2.

2. In this problem, you will estimate an optimal treatment regime using the method of Q-learning. In the HIV study, at each decision point k, $\mathcal{A}_k = \{0, 1\}$. Recall that giving antiretroviral therapy to patients whose virus has become resistant is not feasible. Thus, there are two distinct subsets of $\mathcal{A}_k$ that are feasible sets, $\mathcal{A}_{k,1} = \{0, 1\} = \mathcal{A}_k$ and $\mathcal{A}_{k,2} = \{0\}$; and the feasible sets $\Psi_k(h_k)$ are $\mathcal{A}_{k,1}$ when $r_k = 0$ and $\mathcal{A}_{k,2}$ when $r_k = 1$. Define $s_k(h_k) = 1$ if $r_k = 0$ and $s_k(h_k) = 2$ if $r_k = 1$, so that $s_k(h_k) = r_k + 1$. Then, as on the bottom of Slide 323, the decision rule at Decision k is expressed as

$$d_k(h_k) = \sum_{l=1}^{2} \mathrm{I}\{s_k(h_k) = l\}\, d_{k,l}(h_k)$$

in terms of rules $d_{k,1}(h_k)$ and $d_{k,2}(h_k)$, where $d_{k,2}(h_k) \equiv 0$, so always selects the option of no therapy (for patients whose h_k includes $r_k = 1$, so whose virus is already resistant).

Thus, more precisely, you will estimate d^{opt} in the set $\mathcal{D}$ of all Ψ-specific treatment regimes using the method of Q-learning. Accordingly, the considerations on Slides 364-366 apply.

(a) Using the data from the HIV study, estimate an optimal treatment regime $d^{opt} \in \mathcal{D}$ using the method of Q-learning.

- The choice and development of the Q-function models is up to you
- You are encouraged to include in the kth model, $k = 4, \ldots, 1$, interaction terms between components of x_j and a_j, $j = k, \ldots, 1$, and to investigate the need for such terms.

You are free to program this yourself or to use `qlearn` in the `DynTxRegime` package (you may want to try both to compare).

(b) Estimate the value of d^{opt}, $\mathcal{V}(d^{opt})$.

(c) Express the rules comprising your estimated optimal regime in a way that could be understood and followed by physicians who are treating HIV infected patients.

3. In Problem 3 of Homework 2, you developed and programmed an expression for $\widehat{\mathcal{V}}_{IPW}(d_\eta)$ for a fixed η and thus for regimes in the restricted class of regimes $\mathcal{D}_\eta$ with rules as in (1).

   (a) Estimate an optimal regime $d_\eta^{opt} \in \mathcal{D}_\eta$ by maximizing $\widehat{\mathcal{V}}_{IPW}(d_\eta)$ in η. Note that this optimization is in one-dimension so can be carried out by a simple grid search.

   (b) Estimate the value of an optimal restricted regime, $\mathcal{V}(d_\eta^{opt})$, using $\widehat{\mathcal{V}}_{IPW}(d_\eta)$.

   (c) Compare the estimated value in (b) to that you obtained for an optimal unrestricted regime d^{opt} in (b) of Problem 2.

   (d) The current version of `DynTxRegime` does not allow the user to specify a class of regimes whose rules across decision points depend on the same parameter η. So instead consider the restricted class of regimes $\mathcal{D}_\eta$ with rules given by

$$\begin{aligned} d_k(\overline{x}_k, \overline{a}_k; \eta_k) &= 1 \quad \text{if } c_k < \log(\eta_k) \text{ and } r_k = 0 \\ &= 0 \quad \text{otherwise,} \end{aligned} \tag{2}$$

   $k = 1, \ldots, 4$, where now η_k is specific to Decision k, and $\eta = (\eta_1, \ldots, \eta_4)^T$.

   Use the `optimalSeq` function in `DynTxRegime` to estimate an optimal regime in $\mathcal{D}_\eta$ using the IPW estimator $\widehat{\mathcal{V}}_{IPW}(d_\eta)$, and obtain an estimate of its value. Compare the resulting estimated regime to the one you obtained in (a).

   Note: As in the previous problem, rules for regimes in the class $\mathcal{D}_\eta$ can be written as at the bottom of Slide 368; that is, at decision point k, $\mathcal{A}_k = \{0, 1\}$, and there are two distinct subsets of $\mathcal{A}_k$ that are feasible sets, $\mathcal{A}_{k,1} = \{0, 1\} = \mathcal{A}_k$ and $\mathcal{A}_{k,2} = \{0\}$. As above, $d_{k,1}(h_k; \eta) = 1$ if $c_k < \log(\eta)$ and = 0 otherwise, which takes as input h_k with $r_k = 0$, while $d_{k,2}(h_k; \eta) \equiv 0$, which takes as input h_k with $r_k = 1$; and $s_k(h_k) = r_k + 1$. In this simple situation, the rule $d_k(h_k; \eta)$ can be expressed directly as in (1). Similar considerations apply to (2).

4. Now let's repeat Problem 3 using the AIPW estimator $\widehat{\mathcal{V}}_{AIPW}(d_\eta)$ given in (6.39) on Slide 375. Program this estimator for a fixed η and thus for regimes in the restricted class of regimes

$\mathcal{D}_\eta$ with rules as in (1) using the same fitted propensity models you used in Problem 3 and taking

$$\mathcal{Q}_{d_\eta,k}(\overline{X}_{ki};\widehat{\beta}_k) = Q_k\{\overline{X}_{ki}, \overline{d}_{\eta,k}(\overline{X}_{ki});\widehat{\beta}_k\}, \quad k = 1,\dots,4,$$

where

$$Q_k(h_k, a_k;\widehat{\beta}_k) = Q_k(\overline{x}_k, \overline{a}_k;\widehat{\beta}_k), \quad k = 1,\dots,4,$$

are the Q-function models you developed and fitted in Problem 2.

(a) Estimate an optimal regime $d_\eta^{opt} \in \mathcal{D}_\eta$ by maximizing $\widehat{\mathcal{V}}_{AIPW}(d_\eta)$ in η. As in Problem 3, this optimization is in one-dimension so can be carried out by a simple grid search.

(b) Estimate the value of an optimal restricted regime $\mathcal{V}(d_\eta^{opt})$ using $\widehat{\mathcal{V}}_{AIPW}(d_\eta)$.

(c) Compare the estimated value in (b) to those you obtained for an optimal unrestricted regime d^{opt} in (b) of Problem 2 and in (b) of Problem 3 using $\widehat{\mathcal{V}}_{IPW}(d_\eta)$.

(d) Consider the restricted class of regimes $\mathcal{D}_\eta$ with rules given by (2). Use the `optimalSeq` function in `DynTxRegime` to estimate an optimal regime in $\mathcal{D}_\eta$ using the AIPW estimator $\widehat{\mathcal{V}}_{AIPW}(d_\eta)$, and obtain an estimate of its value. Compare the resulting estimated regime to the one you obtained in (a).

5. (a) Based on your fitted MSM in Problem 1, estimate an optimal regime in the restricted class $\mathcal{D}_\eta$ of regimes with rules of the form (1) as on Slide 434 and obtain an estimate of the value $\mathcal{V}(d_\eta^{opt})$ using $\widehat{\mathcal{V}}_{MSM}(d_\eta)$.

   (b) Compare the estimated value in (a) to those you obtained for an optimal unrestricted regime d^{opt} in (b) of Problem 2, in (b) of Problem 3 using $\widehat{\mathcal{V}}_{IPW}(d_\eta)$, and in (b) of Problem 4 using $\widehat{\mathcal{V}}_{AIPW}(d_\eta)$.

6. Recall that in the BOWL approach, reviewed on Slides 411-416, with $\mathcal{A}_k = \{-1, 1\}$, $k = 1,\dots,K$, treatment rules are represented in terms of a decision function $f_k(h_k;\eta_k)$ as

$$d_k(h_k;\eta_k) = \text{sign}\{f_k(h_k;\eta_k)\}.$$

   In Zhao et al. (2012, 2015), decision functions are represented in terms of a real-valued kernel function $\mathbb{K}(\cdot,\cdot)$; see the papers for details. `bowl` offers three choices for the kernel function, `linear`, `polynomial`, and `radial`; the first option boils down to representing the decision function as a linear combination of elements of h_k, while the other options correspond to more complex and flexible representations. See the papers for details.

   Use the `bowl` function in `DynTxRegime` to estimate an optimal restricted regime taking the decision function to be linear and the surrogate loss function to be the hinge loss (`hinge`). You can experiment with the value of the penalty tuning parameter (`lambdas`) or supply a vector of tuning parameters to be considered, in which case one of these will be chosen by `cvfolds`-fold cross-validation (as the parameter in the vector yielding the smallest cross-validation error). (A nice discussion of k-fold cross-validation can be found in `www.stat.cmu.edu/~ryantibs/datamining/lectures/18-val1.pdf`).

   Ideally, you will want to respect the definition of the feasible sets as discussed in Problems 2 and 3 using the `fSet` option. Alternatively, you could consider regimes whose rules allow a patient whose virus is resistant to receive therapy (so take the feasible sets to be $\mathcal{A}_k$ for all k). One would hope that an optimal regime so estimated would not select treatment option 1 for individuals with $r_k = 1$.