Appendix: Proofs of Theorems in "Outcome-Weighted Learning for Personalized Medicine with Multiple Treatment Options"

Xuan Zhou*, **Yuanjia Wang**†, **Donglin Zeng***

In this section, we prove Theorems 1 and 2 in "Outcome-Weighted Learning for Personalized Medicine with Multiple Treatment Options" by Zhou, Wang, and Zeng.

A.1 Proof of Theorem 1

We start from treatment category k following the order in SOM. First, we show $\mathcal{D}^*(x) = k$ if and only if $E(R|X = x, A = k) = \max_{l=1}^{k} E(R|X = x, A = l)$. For any x with $\mathcal{D}^*(x) = k$, by the definition of $\mathcal{D}^*$, there exists a permutation $(j_1, ..., j_{k-1})$ of $\{1, ..., k-1\}$ such that $\mathcal{D}_l^{*(k)}(x) = -1$ for $l = j_1, ..., j_{k-1}$. That is,

$$f^*_{j_1}(x) < 0, f^*_{j_2}(x) < 0, \ldots, f^*_{j_{k-1}}(x) < 0,$$

where $f^*_{j_l}$ is the counterpart of $\widehat{f}_{j_1}$ when $n = \infty$.

On the other hand, from the estimation of $\widehat{f}_{j_1}$, it is clear that $f^*_{j_1}$ is the minimizer of the expectation of a weighted hinge loss corresponding to $V_{n,j1}$, which is

$$\begin{aligned}
&E\left\{\frac{k-1}{k}\frac{R^+ I(A=j_1)}{\pi_{j_1}(x)}\{1-f(X)\}_+ \middle| X=x\right\} + E\left\{\frac{1}{k}\sum_{l=2}^{k}\frac{R^- I(A=j_l)}{\pi_{j_l}(x)}\{1-f(X)\}_+ \middle| X=x\right\} \\
&+ E\left\{\frac{1}{k}\sum_{l=2}^{k}\frac{R^+ I(A=j_l)}{\pi_{j_l}(x)}\{1+f(X)\}_+ \middle| X=x\right\} + E\left\{\frac{k-1}{k}\frac{R^- I(A=j_1)}{\pi_{j_1}(x)}\{1+f(X)\}_+ \middle| X=x\right\} \\
=&E\left(\frac{k-1}{k}R^+ \middle| X=x, A=j_1\right)\{1-f(x)\}_+ + \sum_{l=2}^{k} E\left(\frac{R^-}{k}\middle| X=x, A=j_l\right)\{1-f(x)\}_+ \\
&+ \sum_{l=2}^{k} E\left(\frac{R^+}{k}\middle| X=x, A=j_l\right)\{1+f(x)\}_+ + E\left(\frac{k-1}{k}R^- \middle| X=x, A=j_1\right)\{1+f(x)\}_+
\end{aligned}$$

where $R^+ = RI(R > 0)$, $R^- = -RI(R \leq 0)$, and $R = R^+ - R^-$.

*Department of Biostatistics, University of North Carolina at Chapel Hill
†Department of Biostatistics, Mailman School of Public Health, Columbia University

We first consider the case when $f(x) \in (-\infty, -1]$, the equation above can be reduced to

$$\left\{ E\left(\frac{k-1}{k} R^{+} \middle| X=x, A=j_1 \right) + \sum_{l=2}^{k} E\left(\frac{R^{-}}{k} \middle| X=x, A=j_l \right) \right\} \{-f(x)\} + constant \qquad \text{(A.1)}$$

It is clear that we cannot find a minimizer for (A.1). Similarly, the minimizer cannot be in the interval $f(x) \in [1, \infty)$. Therefore, we only consider $f(x) \in (-1, 1)$. Then the expectation of a weighted hinge loss corresponding to $V_{n,j1}$ above is:

$$\begin{aligned}
& E\left(\frac{k-1}{k} R^{+} \middle| X=x, A=j_1 \right) \{1-f(x)\}_{+} + \sum_{l=2}^{k} E\left(\frac{R^{-}}{k} \middle| X=x, A=j_l \right) \{1-f(x)\}_{+} \\
& + \sum_{l=2}^{k} E\left(\frac{R^{+}}{k} \middle| X=x, A=j_l \right) \{1+f(x)\}_{+} + E\left(\frac{k-1}{k} R^{-} \middle| X=x, A=j_1 \right) \{1+f(x)\}_{+} \\
= & \left\{ \sum_{l=2}^{k} E\left(\frac{R}{k} \middle| X=x, A=j_l \right) - E\left(\frac{k-1}{k} R \middle| X=x, A=j_1 \right) \right\} f(x) + constant
\end{aligned}$$

That is, $f^{*}_{j_1}(X) < 0$ is equivalent to

$$E\left(\frac{k-1}{k} R \middle| X=x, A=j_1 \right) < \sum_{l=2}^{k} E\left(\frac{R}{k} \middle| X=x, A=j_l \right),$$

which is equivalent to

$$E\left(R \middle| X=x, A=j_1 \right) < \frac{1}{k-1} \sum_{l=2}^{k} E\left(R \middle| X=x, A=j_l \right).$$

Next, when restricting data to those with $A \neq j_1$ and $f^*_{j_1}(X) < 0$, it is clear that $f^*_{j_2}$ minimizes

$$
\begin{aligned}
& E\left\{\frac{k-2}{k-1}\frac{R^+}{\pi_{j_2}(x)}I(A=j_2)\{1-f(X)\}\middle| X=x, A\neq j_1, f^*_{j_1}(X)<0\right\} \\
& + E\left\{\frac{1}{k-1}\sum_{l=3}^{k}\frac{R^-}{\pi_{j_l}(x)}I(A=j_l)\{1-f(X)\}\middle| X=x, A\neq j_1, f^*_{j_1}(X)<0\right\} \\
& + E\left\{\frac{1}{k-1}\sum_{l=3}^{k}\frac{R^+}{\pi_{j_l}(x)}I(A=j_l)\{1+f(X)\}\middle| X=x, A\neq j_1, f^*_{j_1}(X)<0\right\} \\
& + E\left\{\frac{k-2}{k-1}\frac{R^-}{\pi_{j_2}(x)}I(A=j_2)\{1+f(X)\}\middle| X=x, A\neq j_1, f^*_{j_1}(X)<0\right\} \\
=& E\left\{\frac{k-2}{k-1}R^+\middle| X=x, A=j_2, f^*_{j_1}(X)<0\right\}\{1-f(x)\} \\
& + \sum_{l=3}^{k}E\left\{\frac{R^-}{k-1}\middle| X=x, A=j_l, f^*_{j_1}(X)<0\right\}\{1-f(x)\} \\
& + \sum_{l=3}^{k}E\left\{\frac{R^+}{k-1}\middle| X=x, A=j_l, f^*_{j_1}(X)<0\right\}\{1+f(x)\} \\
& + E\left\{\frac{k-2}{k-1}R^-\middle| X=x, A=j_2, f^*_{j_1}(X)<0\right\}\{1+f(x)\} \\
=& \left[\sum_{l=3}^{k}E\left\{\frac{R}{k-1}\middle| X=x, A=j_l, f^*_{j_1}(X)<0\right\} - E\left\{\frac{k-2}{k-1}R\middle| X=x, A=j_2, f^*_{j_1}(X)<0\right\}\right]f(x) \\
& + constant
\end{aligned}
$$

Thus, we conclude that

$$
\mathrm{sign}\{f^*_{j_2}(x)\} = \mathrm{sign}[E\{(k-2)R|X=x, A=j_2\} - \sum_{l=3}^{k}E\{R|X=x, A=j_l\}]I\{f^*_{j_1}(x)<0\}.
$$

That is, $f^*_{j2}(x) < 0$ if and only if

$$
E(R|X=x, A=j_2) < \frac{1}{k-2}\sum_{l=3}^{k}E(R|X=x, A=j_l)
$$

Continue the same arguments so we establish the relationship between $f^*_{j_l}$ and $E(R|X=x, A=j_l)$ as

$$
\mathrm{sign}\{f^*_{j_l}(x)\} = \mathrm{sign}\left\{E(R|X=x, A=j_l) - \frac{1}{k-l}\sum_{h=l+1}^{k}E(R|X=x, A=j_h)\right\}
$$

In other words, we obtain that for this subject with $f^*_{j_1}(x) < 0, ..., f^*_{j_{k-1}}(x) < 0$, it holds

$$\begin{aligned}
E(R|X=x, A=j_1) < & \quad \frac{1}{k-1}\sum_{l=2}^{k} E(R|X=x, A=j_l), \\
E(R|X=x, A=j_2) < & \quad \frac{1}{k-2}\sum_{l=3}^{k} E(R|X=x, A=j_l), \\
& \vdots \\
E(R|X=x, A=j_{k-2}) < & \quad 1/2\{E(R|X=x, A=j_{k-1}) + E(R|X=x, A=k)\}, \\
E(R|X=x, A=j_{k-1}) < & \quad E(R|X=x, A=k).
\end{aligned}$$

Starting from the last inequality in the above, in turn, we have

$$\begin{aligned}
E(R|X=x, A=j_{k-1}) & < E(R|X=x, A=k) \\
E(R|X=x, A=j_{k-2}) & < 1/2\{E(R|X=x, A=j_{k-1}) + E(R|X=x, A=k)\} \\
& < E(R|X=x, A=k), \\
& \vdots \\
E(R|X=x, A=j_1) & < \frac{1}{k-1}\sum_{l=2}^{k} E(R|X=x, A=j_1) < E(R|X=x, A=k).
\end{aligned}$$

Therefore,

$$E(R|X=x, A=k) = \max_{l=1}^{k} E(R|X=x, A=l).$$

For the other direction, we suppose that

$$E(R|X=x, A=k) = \max_{l=1}^{k} E(R|X=x, A=l).$$

We order the expectations to obtain

$$E(R|X=x, A=j_1) \leq E(R|X=x, A=j_2) \leq ... \leq E(R|X=x, A=k)$$

Thus all the inequalities in (2)-(7) hold, from equivalence between $f^*_{j_l}$ and $E(R|X=x, A=j_l)$'s, it is straightforward to see that

$$f^*_{j_1}(x) < 0, ..., f^*_{j_{k-1}}(x) < 0.$$

In other words, $\mathcal{D}^*(x) = k$. Hence, we have proved that SOM learning correctly assigns subjects whose conditional mean outcomes are maximal in treatment k into the optimal treatment k.

To prove the consistency of the remaining classes, obtains the rule for class $(k-1)$ conditional on $A \neq k$ and $\mathcal{D}^*(x) \neq k$. Using the same proof as above, we conclude

$$\mathcal{D}^*(x) = (k-1) \text{ if and only if } (k-1) = \text{argmax}_{l=1}^{k-1} \widetilde{E}(R|X = x, A = l),$$

where $\widetilde{E}(R|X = x, A = j_l)$ is the conditional expectation of R given $X = x$, $A \neq k$ and $\mathcal{D}^*(x) \neq k$. Moreover, $\mathcal{D}^*(x) \neq k$ implies that $E(R|X = x, A = k)$ cannot be the maximum. Therefore,

$$(k-1) = \text{argmax}_{l=1}^{k-1} E(R|X = x, A = l) = \text{argmax}_{l=1}^{k} E(R|X = x, A = l).$$

That is,

$$\mathcal{D}^*(x) = (k-1) \text{ if and only if } (k-1) = \text{argmax}_{l=1}^{k} E(R|X = x, A = l).$$

We continue this proof for the remaining classes and finally obtain Fisher consistency.

A.2 Proof of Theorem 2

We first note

$$\begin{aligned}
& \mathcal{R}(\widehat{\mathcal{D}}) - \mathcal{R}(\mathcal{D}^*) \\
= & \sum_{l=1}^{k} \left[E\left\{ \frac{R}{\pi_l(X)} I(A = l, \widehat{\mathcal{D}}(X) \neq l) \right\} - E\left\{ \frac{R}{\pi_l(X)} I(A = l, \mathcal{D}^*(X) \neq l) \right\} \right] \\
= & \sum_{l=1}^{k} \left[E\left\{ \frac{R}{\pi_l(X)} I(A = l, \widehat{\mathcal{D}}(X) \neq l, \mathcal{D}^*(X) = l) \right\} - E\left\{ \frac{R}{\pi_l(X)} I(A = l, \mathcal{D}^*(X) \neq l, \widehat{\mathcal{D}}(X) = l) \right\} \right].
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\mathcal{R}(\widehat{\mathcal{D}}) - \mathcal{R}(\mathcal{D}^*) \\
= &\sum_{l=1}^{k}\left[E\left\{\frac{R}{\pi_l(X)}I(A=l,\widehat{\mathcal{D}}(X)\neq l,\mathcal{D}^*(X)=l)\right\} - E\left\{\frac{R}{\pi_A(X)}I(A\neq l,\widehat{\mathcal{D}}(X)\neq l,\mathcal{D}^*(X)=l)\right\}\right]. \\
\leq &\sum_{l=1}^{k}\left[E\left\{\frac{R^+}{\pi_A(X)}I(A=l,\widehat{\mathcal{D}}(X)\neq l,\mathcal{D}^*(X)=l)\right\} + E\left\{\frac{R^-}{\pi_A(X)}I(A\neq l,\widehat{\mathcal{D}}(X)\neq l,\mathcal{D}^*(X)=l)\right\}\right].
\end{aligned}$$

We let Δ_l to denote each term on the right-hand side of the above equation. That is,

$$\begin{aligned}
\Delta_l &= E\left\{\frac{R^+}{\pi_A(X)}I(A=l,\widehat{\mathcal{D}}(X)\neq l,\mathcal{D}^*(X)=l)\right\} + E\left\{\frac{R^-}{\pi_A(X)}I(A\neq l,\widehat{\mathcal{D}}(X)\neq l,\mathcal{D}^*(X)=l)\right\} \\
&= E\left\{\frac{|R|}{\pi_A(X)}I(Z_l\text{sign}(R)=1,\widehat{\mathcal{D}}(X)\neq l,\mathcal{D}^*(X)=l)\right\},
\end{aligned}$$

where we recall $Z_l = 2I(A=l)-1$.

We first examine Δ_k. For any x in the domain of X, we let $j_1, j_2, ..., j_{k-1}$ be the permutation of $\{1, ..., k-1\}$ such that

$$E(R|A=j_1, X=x) < ... < E(R|A=j_{k-1}, X=x).$$

Then according to SOM learning, $\mathcal{D}^*(x)=k$ implies that $f^*_{j_l(x)}(x)<0$ for any $l=1,..,k-1$, while $\widehat{\mathcal{D}}(X)\neq k$ implies that for this particular permutation, there exists some $l=1,...,k-1$ such that $\widehat{f}_{j_l}(x)>0$ so $\widehat{f}_{j_l}(x)f^*_{j_l}(x)<0$. Recall that $f^*_{j_l}(x)=\eta_{j_l,S}$ with $S=\{j_{l+1},...,k\}$ and it is the limit of $\widehat{f}_{j_l}$ from Theorem 3.1. Therefore, we obtain

$$\begin{aligned}
\Delta_k &\leq E\left[\frac{|R|}{\pi_A(X)}\left\{\sum_{(j_1,...,j_{k-1})} I(Z_k\text{sign}(R)=1,\text{there exists } l\leq k-1 \text{ s.t. } \widehat{f}_{j_l}(X)f^*_{j_l}(X)<0)\right\}\right] \\
&\leq \sum_{(j_1,...,j_{k-1})} E\left[\frac{|R|}{\pi_A(X)}I\left\{Z_{j_1}\text{sign}(R)=-1,...,Z_{j_{l-1}}\text{sign}(R)=-1,\widehat{f}_{j_l}(X)f^*_{j_l}(X)<0\right\}\right] \\
&\leq \sum_{(j_1,...,j_{k-1})} E\left[\frac{|R|}{\pi_A(X)}\{I(A=j_l)(k-l+1)+I(A\neq j_l)\}\right. \\
&\qquad \left.\times I\left\{Z_{j_1}\text{sign}(R)=-1,...,Z_{j_{l-1}}\text{sign}(R)=-1,\widehat{f}_{j_l}(X)f^*_{j_l}(X)<0\right\}\right].
\end{aligned}$$

Hence, it suffices to bound each term on the right-hand side of the above inequality.

When $l = 1$, under conditions 1-3, we use the same proof of Theorem 3.2 in Zhao et al. (2012), which extends the result in Stienwart and Christmann (2008) to a weighted support vector machine. Particularly, in their proof, we let the weight for subject i be

$$|R_i|/\pi_{A_i}(X_i)\{(k-1)I(A_i = j_1) + I(A_i \neq j_1)\}$$

and the class label be $Z_{j_1}\text{sign}(R_i)$. Furthermore, from the proof of Theorem 3.1, $f^*_{j_1}(x)$ has the same sign as $\eta_{j_1,\{j_2,\ldots,j_k\}}(x)$. Thus, from condition (C.1), we conclude that there exists at least probability $1 - 3e^{-\epsilon}$ and a constant C_1 such that it holds

$$E\left[\frac{|R|}{\pi_A(X)}\{(k-1)I(A = j_1) + I(A \neq j_1)\} I(Z_{j_1}\text{sign}(R)\widehat{f}_{j_1}(X) < 0)\right]$$
$$-E\left[\frac{|R|}{\pi_A(X)}\{(k-1)I(A = j_1) + I(A \neq j_1)\} I(Z_{j_1}\text{sign}(R)f^*_{j_1}(X) < 0)\right] \leq C_1 Q_n(\epsilon),$$

where

$$Q_n(\epsilon) = \left\{\lambda_n^{\frac{\tau}{2+\tau}}\sigma_n^{-\frac{d\tau}{d+\tau}} + \sigma_n^{\beta} + \epsilon\left(n\lambda_n^p\sigma_n^{\frac{1-p}{1+\epsilon_0}d}\right)^{-\frac{q+1}{q+2-p}}\right\}$$

with any constant $\epsilon_0 > 0$ and $d/(d+\tau) < p < 2$. Then according to the proof of Lemma 5 in Bartlett et al. (2006) and conditions 1 and 2, this gives

$$\Pr\{\widehat{f}_{j_1}(X)f^*_{j_1}(X) < 0\} \leq \{C_1' Q_n(\epsilon)\}^{\alpha},$$

where $\alpha = q/(1+q)$ and C_1' is a constant.

When $l = 2$, the step at j_2 in SOM is to minimize

$$n^{-1}\sum_{i=1}^{n} I\{Z_{ij_1} = -1, Z_{ij_1}\text{sign}(R_i)\widehat{f}_{j_1}(X_i) < 0\}w_i\{1 - Z_{ij_2}\text{sign}(R_i)f(X_i)\}_+ + \lambda_{n,j_2}\|f\|^2,$$

where $w_i = |R_i|/\pi_{A_i}(X_i)\{(k-2)I(A_i = j_2) + I(A_i \neq j_2)\}$. Thus, we can proceed the same proof of Theorem 3.2 in Zhao et al. (2012) except that only subjects in the random set

$$\left\{i : Z_{ij_1} = -1, Z_{ij_1}\text{sign}(R_i)\widehat{f}_{j_1}(X_i) < 0\right\}$$

are used in the derivation. We obtain that

$$\begin{aligned}
& E\left[\frac{|R|}{\pi_A(X)}\{(k-2)I(A=j_2)+I(A\neq j_2)\}I\{Z_{j_1}=-1, Z_{j_2}\text{sign}(R)\widehat{f}_{j_2}(X)<0\}\right] \\
& \quad -E\left[\frac{|R|}{\pi_A(X)}\{(k-2)I\{A=j_2)+I(A\neq j_2)\}I\{Z_{j_1}=-1, Z_{j_2}\text{sign}(R)f^*_{j_2}(X)<0\}\right] \\
\leq\ & C_2\left\{Q_n(\epsilon)+|\Pr(Z_{j_1}\text{sign}(R)\widehat{f}_{j_1}(X)>0)-\Pr(Z_{j_1}\text{sign}(R)f^*_{j_1}(X)>0)|\right\} \\
\leq\ & C_2\{Q_n(\epsilon)+Q_n(\epsilon)^\alpha\}
\end{aligned}$$

with a probability at least $1-3e^{-\epsilon}$ for a constant C_2. Note that the second term on the right-hand side is due to the estimated random set in this step. Again, the proof of Lemma 5 in Bartlett et al. (2006) gives

$$\Pr\{Z_{j_1}=-1, \widehat{f}_{j_2}(X)f^*_{j_2}(X)<0\}\leq\{C_2'Q_n(\epsilon)\}^\alpha.$$

We continue the same arguments for $l=3,...,k-1$ to obtain

$$\begin{aligned}
& E\left[\frac{|R|}{\pi_A(X)}\{(k-l+1)I(A=j_l)+I(A\neq j_l)\}I\left\{Z_{j_l}\text{sign}(R)\widehat{f}_{j_l}(X)<0, Z_{j_{l-1}}=-1,...,Z_{j_1}=-1\right\}\right] \\
- & E\left[\frac{|R|}{\pi_A(X)}\{(k-l+1)I(A=j_l)+I(A\neq j_l)\}I\left\{Z_{j_l}f^*_{j_l}(X)<0, Z_{j_{l-1}}=-1,...,Z_{j_1}=-1\right\}\right] \\
\leq\ & C_l\{Q_n(\epsilon)+Q_n(\epsilon)^\alpha\}
\end{aligned}$$

with a probability at least $1-3le^{-\epsilon}$ for some constant C_l, and

$$\Pr\{Z_{j_1}=-1,...,Z_{j_{l-1}}=-1, \widehat{f}_{j_l}(X)f^*_{j_l}(X)<0\}\leq\{C_l'Q_n(\epsilon)\}^\alpha$$

for a constant C_l'. Hence, with a probability $1-\{3k(k-1)/2\}e^{-\epsilon}$, $\Delta_k\leq CQ_n(\epsilon)^\alpha$ for a constant C.

Similarly, we can examine the difference for Δ_{k-1}. We follow exactly the same arguments as before by considering all possible permutations from $\{1,...,k-2\}$ and $l=1,...,k-2$. The only difference in the argument is that the random set is restricted to subjects with $A\neq k$ and $\widehat{\mathcal{D}}^{(k)}(X)=-1$. However, the probability of the latter differs from the probability $A\neq k$ and $\mathcal{D}^{*(k)}(X)=-1$ by $CQ_n(\epsilon)^\alpha$ from the previous conclusion. Therefore, we obtain that with probability at least $1-\{3k(k-1)/2+3(k-1)(k-2)/2\}e^{-\epsilon}$, $\Delta_{k-1}\leq CQ_n(\epsilon)^\alpha$ for another constant C. Continue the same arguments for $\Delta_l, l = k-2, ..., 1$ so we finally conclude

$$\mathcal{R}(\widehat{\mathcal{D}}) - \mathcal{R}^* \leq CQ_n(\epsilon)^\alpha$$

with probability at least $1 - C'e^{-\epsilon}$ where C' is a constant depending on k. Thus Theorem 3.2 holds.

References

Bartlett, P. L., Jordan, M. I., and McAuliffe, J. D. (2006). Convexity, classification, and risk bounds. *Journal of the American Statistical Association*, 101(473):138–156.

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.