OPTIMAL DYNAMIC TREATMENT RULES

For treatment *sequence* $(\mathsf{z}_1, \mathsf{z}_2)$, we seek the *optimal* treatment sequence $(z_1^{\text{opt}}, z_2^{\text{opt}})$ to *maximize* expected response. We decompose the expected (counterfactual) response as

$$\mathbb{E}[Y(\mathsf{z}_1,\mathsf{z}_2)] = \mathbb{E}[Y(z_1^{\text{opt}}, z_2^{\text{opt}})] - \{\mathbb{E}[Y(z_1^{\text{opt}}, z_2^{\text{opt}}) - Y(\mathsf{z}_1, z_2^{\text{opt}})]\} - \{\mathbb{E}[Y(\mathsf{z}_1, z_2^{\text{opt}}) - Y(\mathsf{z}_1,\mathsf{z}_2)]\}\,.$$

The terms in the decomposition on the right-hand side are

- first term: the response under optimal treatment at both intervals;
- second term: the penalty for suboptimal treatment at the first stage, assuming optimal treatment at the second stage;
- third term: the penalty for suboptimal treatment at the second stage, assuming the observed treatment at the first stage.

This decomposition is the basis of the *Structural Nested Mean Model* (SNMM). Further, we can re-write the second and third terms by comparing to the baseline level (zero):

$$\mathbb{E}[Y(z_1^{\text{opt}}, z_2^{\text{opt}}) - Y(\mathsf{z}_1, z_2^{\text{opt}})] = \mathbb{E}[Y(z_1^{\text{opt}}, z_2^{\text{opt}}) - Y(0, z_2^{\text{opt}})] - \mathbb{E}[Y(\mathsf{z}_1, z_2^{\text{opt}}) - Y(0, z_2^{\text{opt}})]$$

$$\mathbb{E}[Y(\mathsf{z}_1, z_2^{\text{opt}}) - Y(\mathsf{z}_1,\mathsf{z}_2)] = \mathbb{E}[Y(\mathsf{z}_1, z_2^{\text{opt}}) - Y(\mathsf{z}_1, 0)] - \mathbb{E}[Y(\mathsf{z}_1,\mathsf{z}_2) - Y(\mathsf{z}_1, 0)]$$

We then specify *blip* models

$$\mathbb{E}[Y(\mathsf{z}_1,\mathsf{z}_2) - Y(\mathsf{z}_1, 0)] = (\mathbf{x}_2\psi_2)\mathsf{z}_2 \qquad \text{Stage 2}$$

$$\mathbb{E}[Y(\mathsf{z}_1, z_2^{\text{opt}}) - Y(0, z_2^{\text{opt}})] = (\mathbf{x}_1\psi_1)\mathsf{z}_1 \qquad \text{Stage 1}$$

where

- $\mathbf{x}_2$ can depend on all data – including observed treatments – observed up to Stage 2.
- $\mathbf{x}_1$ can depend on all data observed up to Stage 1.

1. Use *G-estimation* on Y conditioning on $(\mathbf{x}_1, z_1, \mathbf{x}_2, z_2)$ at second stage to estimate ψ_2 using mean model
$$Y = \mathbf{x}_{21}\beta_2 + z_2\mathbf{x}_{22}\psi_2 + \varepsilon$$
2. For each individual infer the optimal Stage 2 treatment: $z_2^{\text{opt}} = \mathbb{1}\{\mathbf{x}_{22}\widehat{\psi}_2 > 0\}$
3. Form *pseudo-outcome* Y_1
$$Y_1 = Y - (\mathbf{x}_{22}\widehat{\psi}_2)(z_2^{\text{opt}} - z_2)$$
4. Use G-estimation on Y_1 conditioning on $(\mathbf{x}_1, z_1)$ at first stage to estimate ψ_1 using a proposed mean model
$$Y_1 = \mathbf{x}_{11}\beta_1 + z_1\mathbf{x}_{12}\psi_1 + \varepsilon$$
5. For each individual infer the optimal Stage 1 treatment: $z_1^{\text{opt}} = \mathbb{1}\{\mathbf{x}_{11}\widehat{\psi}_1 > 0\}$.

The method is robust to mis-specification of the *nuisance mean model*, provided the treatment model is correctly specified. We can infer the optimized potential outcome

$$Y + (\mathbf{x}_{12}\widehat{\psi}_1)(z_1^{\text{opt}} - z_1) + (\mathbf{x}_{22}\widehat{\psi}_2)(z_2^{\text{opt}} - z_2)$$

which takes the observed outcome Y and adds

- the additional benefit of optimal treatment at stage 1
$$(\mathbf{x}_{12}\widehat{\psi}_1)(z_1^{\text{opt}} - z_1)$$
- the additional benefit of optimal treatment at stage 2
$$(\mathbf{x}_{22}\widehat{\psi}_2)(z_2^{\text{opt}} - z_2).$$

Simulation study: We simulate the data as follows

- $Y(z_1^{\text{opt}}, z_2^{\text{opt}}) \sim Normal(200, 5^2)$
- $X_1 \sim Normal(1, 1)$
- $Z_1 \sim Bernoulli(\text{expit}(\alpha_{10} + \alpha_{11} X_1))$
- $X_2 | X_1, Z_1 \sim Normal(X_1 + Z_1, 1)$
- $Z_2 | X_1, Z_1, X_2 \sim Bernoulli(\text{expit}(\alpha_{20} + \alpha_{21} X_1 + \alpha_{22} X_2 + \alpha_{23} Z_1))$
- Stage 1: $\psi_1 = (-2, 1.5)$
$$\mathbf{x}_1 \psi_1 = \psi_{10} + \psi_{11} X_1$$
- Stage 2: $\psi_2 = (-2, 1, 1, -1)$
$$\mathbf{x}_2 \psi_2 = \psi_{20} + \psi_{21} X_1 + \psi_{22} X_2 + \psi_{23} Z_1$$
- Outcome:
$$Y = Y(z_1^{\text{opt}}, z_2^{\text{opt}}) - (\mathbf{x}_{12}\psi_1)(z_1^{\text{opt}} - z_1) - (\mathbf{x}_{22}\psi_2)(z_2^{\text{opt}} - z_2)$$

In the two fitting steps 1. and 4., the models

$$Y = \beta_{20} + z_2 \mathbf{x}_{22} \psi_2 + \varepsilon$$

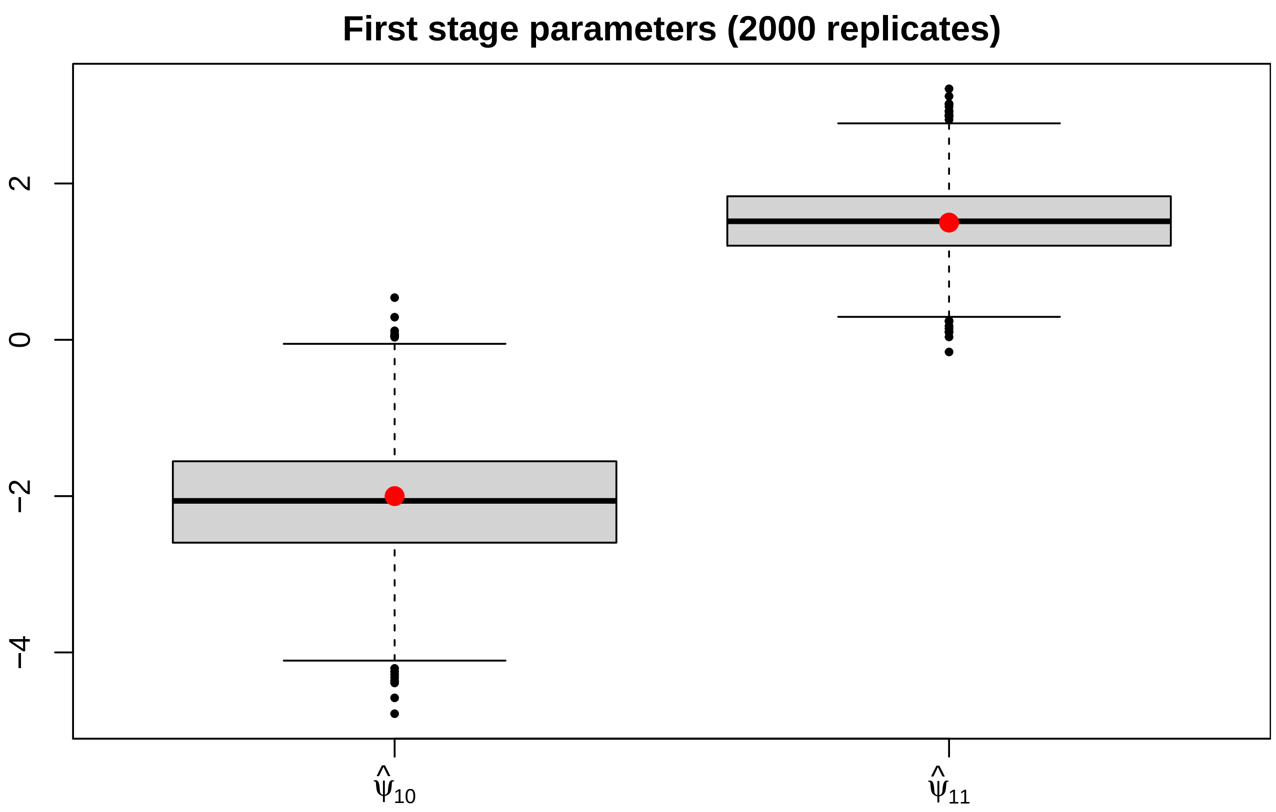
and

$$Y = \beta_{10} + z_1 \mathbf{x}_{12} \psi_1 + \varepsilon$$

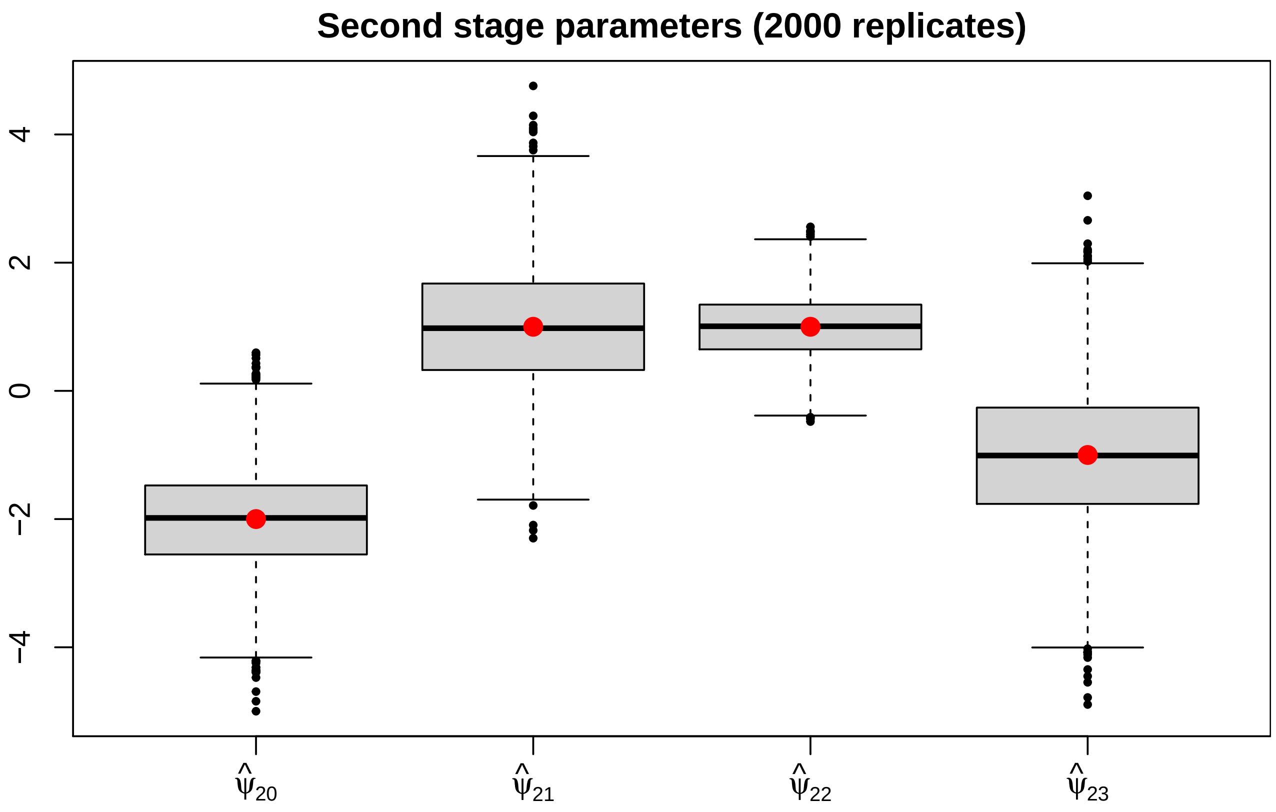
are used, with treatment-free intercept only models. Note that these models are mis-specified. The G-estimation steps are carried out here using propensity score regression.

```
set.seed(34)
expit<-function(x){return(1/(1+exp(-x)))}
n<-1000
al1<-c(-2,1)
al2<-c(-2,3,-1,1)
psi1<-c(-2,1.5)
psi2<-c(-2,1,1,-1)

nreps<-2000
psi1.mat<-matrix(0,nrow=nreps,ncol=length(psi1))
psi2.mat<-matrix(0,nrow=nreps,ncol=length(psi2))
Y.opts<-matrix(0,nrow=nreps,ncol=n)
improved<-rep(0,nreps)

for(irep in 1:nreps){
        X1<-rnorm(n,1,1)
        l1<-al1[1]+al1[2]*X1
        Z1<-rbinom(n,1,expit(l1))
        X2<-rnorm(n)+Z1+X1
        l2<-cbind(1,X1,X2,Z1) %*% al2
        Z2<-rbinom(n,1,expit(l2))

        blip1<-psi1[1]+psi1[2]*X1
        Z1.opt<-as.numeric(blip1>0)

        blip2<-cbind(1,X1,X2,Z1) %*% psi2
        Z2.opt<-as.numeric(blip2>0)

        Y.opt<-rnorm(n,200,5)
        regret2<-(cbind(1,X1,X2,Z1) %*% psi2)*(Z2.opt-Z2)
        regret1<-(cbind(1,X1) %*% psi1)*(Z1.opt-Z1)
        Y<-Y.opt-regret1-regret2

        ##############################
```

```
        #Fit second stage
        fitz2<-glm(Z2~X1+X2+Z1,family=binomial)
        e2<-fitted(fitz2)

    #Use propensity score regression to ensure double robustness
        fity2<-lm(Y~1+(Z2+Z2:X1+Z2:X2+Z2:Z1)+(e2+e2:X1+e2:X2+e2:Z1))
        psi2.hat<-coef(fity2)[c(2,4,5,6)]
        z2.opt<-as.numeric(cbind(1,X1,X2,Z1) %*% psi2.hat >0)

        Y1<-Y-(cbind(1,X1,X2,Z1) %*% psi2.hat)*(z2.opt-Z2)

        #Fit first stage
        fitz1<-glm(Z1~X1,family=binomial)
        e1<-fitted(fitz1)

    #Use propensity score regression to ensure double robustness
        fity1<-lm(Y1~1+(Z1+Z1:X1)+(e1+e1:X1))
        psi1.hat<-coef(fity1)[c(2,4)]
        z1.opt<-as.numeric(cbind(1,X1) %*% psi1.hat >0)

        psi1.mat[irep,]<-psi1.hat
        psi2.mat[irep,]<-psi2.hat

        regret2<-(cbind(1,X1,X2,Z1) %*% psi2.hat)*(z2.opt-Z2)
        regret1<-(cbind(1,X1) %*% psi1.hat)*(z1.opt-Z1)

    Y.opts[irep,]<-Y+regret1+regret2

    #Number who would have improved
    improved[irep]<-sum(Y < Y.opts[irep,])

}
```



First stage parameters (2000 replicates)



Second stage parameters (2000 replicates)

```
#Mean number of subjects who were treated sub-optimally
mean(improved)

+ [1] 545.6025
```