Results

2020-06-05

Outline

- 1. Trajectories with different $w = \alpha' x$ valuse
- 2. Different α_1 and α_2 values
- 3. Add L1 penalty in purity calculation function

Trajectory

Data generation

If the true outcome generation models are (with different true α):

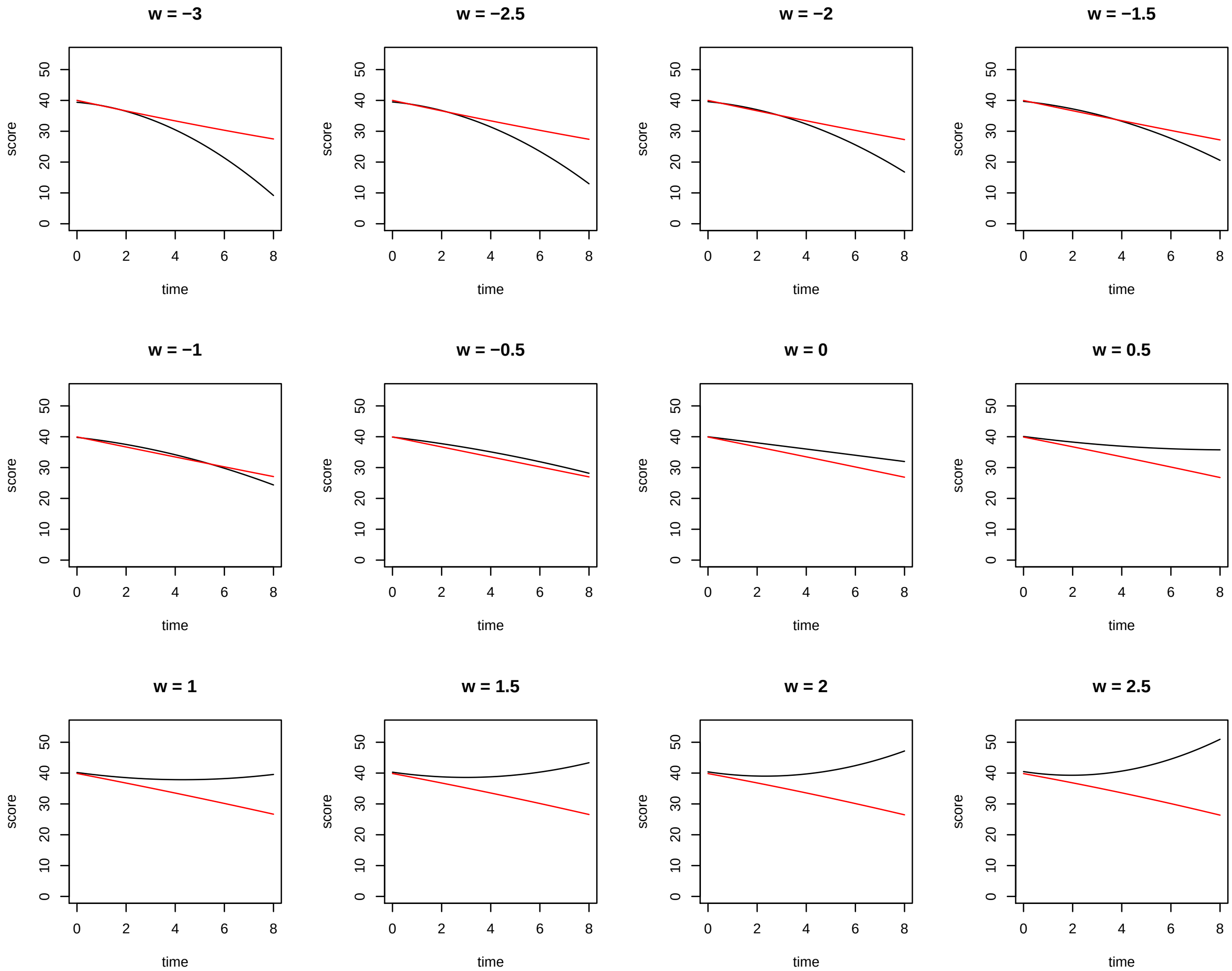
$$Y_{drg} = S(\beta_{drg} + \Gamma_{drg}(\alpha'_{drg}x)) + Sb_{drg} + \epsilon_{drg}$$

$$Y_{pbo} = S(\beta_{pbo} + \Gamma_{pbo}(\alpha'_{pbo}x)) + Sb_{pbo} + \epsilon_{pbo}$$

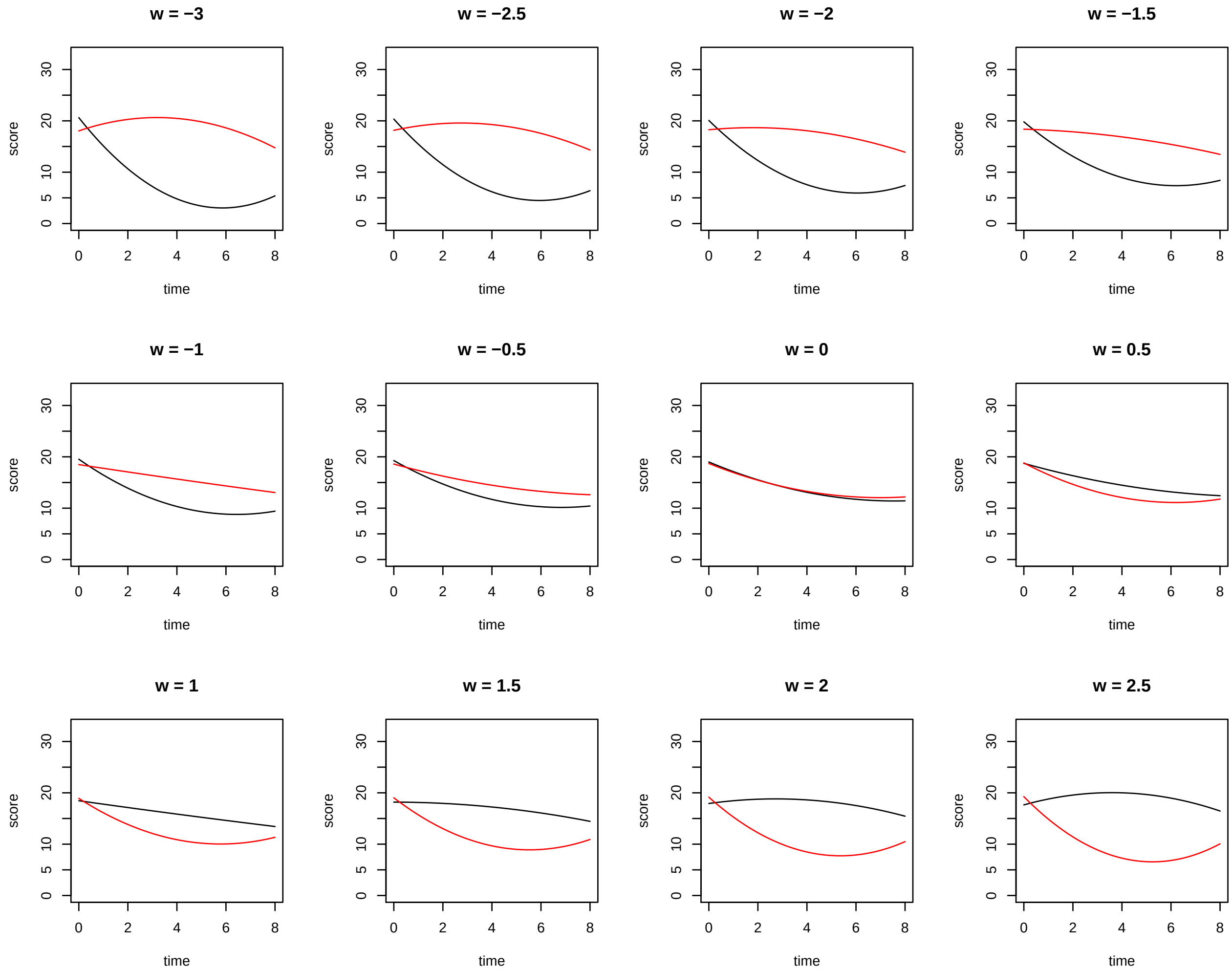
where

- $S = [1, t, t^2]$, $t = [0, 1, 2, 3, 4, 6, 8]$ is the design matrix for fixed effect and random effect
- $x_1 \sim MVN(\mu_x, \Sigma_x)$, $\mu_x = \mathbf{0}_p$, Σ_x has diagonal equals to 1 and 0.5 anywhere else. The dimension of x is set to be $p = 20$.
- 80 noises are also added in the dataset, each of them are generated from $x_{2,j} \sim N(0, 1)$. Therefore, there are $np = 100$ covariates in total.
- $\beta_{drg} = [40, -1, -0.02], \beta_{pbo} = [40, -1.1, -0.02]$
- $\Gamma_{drg} = [0, 0.1\cos(\theta), 0.1\sin(\theta)], \Gamma_{pbo} = [0, 0.01, 0], \theta = \frac{\pi}{2}$
- $b_{drg} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.01 & 0 \\ 0 & 0 & 0.02 \end{bmatrix}, b_{pbo} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.01 & 0.005 \\ 0 & 0.005 & 0.01 \end{bmatrix}$
- $\epsilon_{drg}, \epsilon_{pbo} \sim N(0, 3^2)$
- $\alpha = 1_p$

The trajectory plot with different w value, black line is the drug group while the red line presents the placebo group.

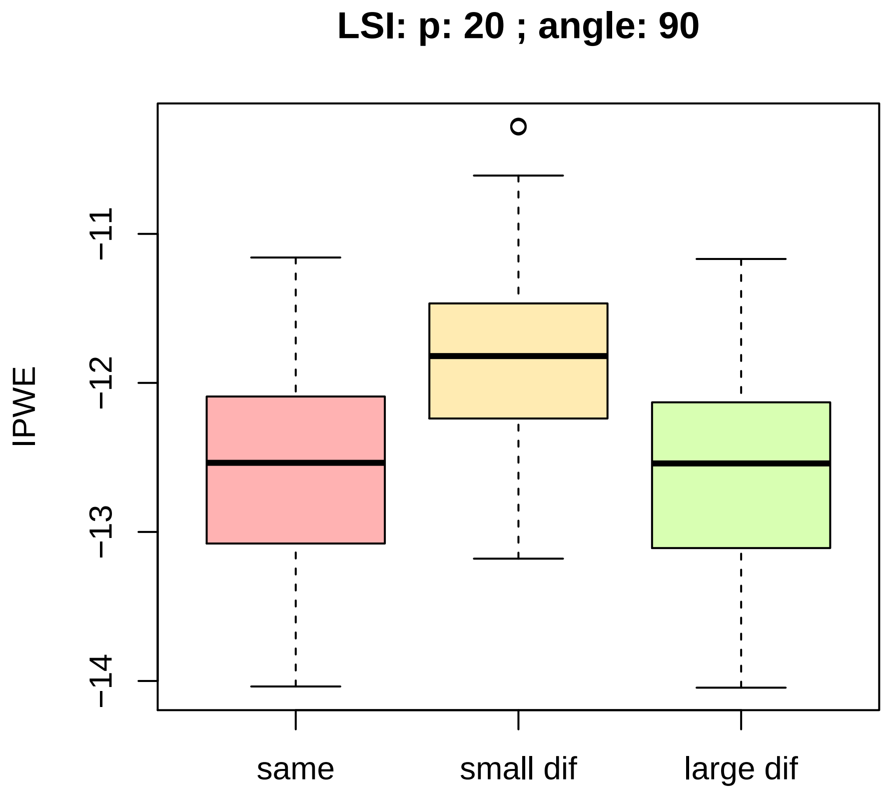


The trajectory plot for EMBARC dataset, used the covariates selected by the forward and backward selection.



Diff alpah

- same alpha: $\alpha_1 = \alpha_2 = I_p$
- small difference:
- $\alpha_1 = I_p$
- $\alpha_2 = [I_{\frac{p}{2}}, 1 : \frac{p}{2}]$
- big difference:
- $\alpha_1 = I_p$
- $\alpha_2 = 1 : p$



LASSO

- Penalty added in the Purity function $G(\alpha)$

$$G_p(\alpha) = A_0 + A_1 \hat{\mu}_x' \alpha + \frac{A_2}{2}[\alpha' \hat{\Sigma}_x \alpha + \alpha' \hat{\mu}_x \hat{\mu}_x' \alpha] - \lambda ||\alpha||_1$$

  where

$$A_0 = -q + \frac{1}{2} tr(\hat{D}_2^{-1} \hat{D}_1) + \frac{1}{2} tr(\hat{D}_1^{-1} \hat{D}_2) + \frac{1}{2}(\hat{\beta}_1 - \hat{\beta}_2)'(\hat{D}_1^{-1} + \hat{D}_2^{-1})(\hat{\beta}_1 - \hat{\beta}_2)$$
$$A_1 = (\hat{\beta}_1 - \hat{\beta}_2)'(\hat{D}_1^{-1} + \hat{D}_2^{-1})(\hat{\Gamma}_1 - \hat{\Gamma}_2)$$
$$A_2 = (\hat{\Gamma}_1 - \hat{\Gamma}_2)'(\hat{D}_1^{-1} + \hat{D}_2^{-1})(\hat{\Gamma}_1 - \hat{\Gamma}_2)$$

  q is the dimension of D matrix.

- Calculation derivation:
  - approximation
  - sign function

- Data set
  - Simulation (100 covariates, 80 noises)
  - EMBARC (215 covariates)

Attempt 1

1. Initial $\alpha^{(0)}$ and λ
2. Fit linear mixed effect model and estimate $\hat{\beta}_1^{(0)}, \hat{\beta}_2^{(0)}, \hat{\Gamma}_1^{(0)}, \hat{\Gamma}_2^{(0)}, \hat{D}_1^{(0)}, \hat{D}_2^{(0)}$ and then associated $\hat{A}_0^{(0)}, \hat{A}_1^{(0)}, \hat{A}_2^{(0)}$

3. Update α:

- Maximize the penalized purity function
- Since no class from, an approximation of $||\alpha||_1$ is conducted

$$G_p(\alpha) = A_0 + A_1\hat{\mu}_x'\alpha + \frac{A_2}{2}[\alpha'\hat{\Sigma}_x\alpha + \alpha'\hat{\mu}_x\hat{\mu}_x'\alpha] - \lambda||\alpha||_1$$

4. Repeat 2-3 until converge.

5. Tune parameter

- For a given sequence of $\lambda \in \{0.001, 0.01, 0.1, 1, 10, 100\}$, find the $\hat{\alpha}$ that maximizes the penalized purity function.
- Do 10 fold cross validation. Use 9 of the part to train the data, fit LME with $\hat{\alpha}'x$ and calculate $\hat{\beta}_1, \hat{\beta}_2, \hat{\Gamma}_1, \hat{\Gamma}_2, \hat{D}_1, \hat{D}_2$.
- Calculate the IPWE for each of the 10 fold cross validation
- Repeat IPWE calculation for 100 times
- Choose the λ that achieve the max IPWE.

6.

- $\lambda \in \{0.001, 0.01, 0.1, 1, 10, 100\}$, the one with the largest IPWE l_1
- $\lambda \in seq(l_1/10, 1_1 * 10, l_1)$, the one with the largest IPWE l_2
- $\lambda \in seq(l_2/10, 1_2 * 10, l_2)$, the one with the largest IPWE l_3, which is the tuned parameter $\hat{\lambda}$

For the penalty, we use the approximation

$$|\alpha_i^{(j+1)}| = |\alpha_i^{(j)}| + \frac{1}{2|\alpha_i^{(j)}|}\{(\alpha_i^{(j+1)})^2 - (\alpha_i^{(j)})^2\}$$

And

$$||\alpha^{(j+1)}||_1 \approx ||\alpha^{(j)}||_1 + \frac{1}{2}\sum_{i=1}^{p}\frac{(\alpha_i^{(j+1)})^2}{|\alpha_i^{(j)}|} - \frac{1}{2}\sum_{j=1}^{p}\frac{(\alpha_i^{(j)})^2}{|\alpha_i^{(j)}|} \propto \frac{1}{2}\sum_{i=1}^{p}\frac{(\alpha_i^{(j+1)})^2}{|\alpha_i^{(j)}|}$$

Therefore,

$$\begin{aligned}
argmaxG_p(\alpha) &\approx argmax\big[A_0 + A_1\hat{\mu}_x'\alpha^{(j+1)} + \frac{A_2}{2}[(\alpha^{(j+1)})'\hat{\Sigma}_x\alpha^{(j+1)} + (\alpha^{(j+1)})'\hat{\mu}_x\hat{\mu}_x'\alpha^{(j+1)}] - \frac{\lambda}{2}\sum_{i=1}^{p}\frac{(\alpha_i^{(j+1)})^2}{|\alpha_i^{(j)}|}\big] \\
&= argmax\big[A_0 + A_1\hat{\mu}_x'\alpha^{(j+1)} + \frac{A_2}{2}[(\alpha^{(j+1)})'\hat{\Sigma}_x\alpha^{(j+1)} + (\alpha^{(j+1)})'\hat{\mu}_x\hat{\mu}_x'\alpha^{(j+1)}] - \frac{\lambda}{2}(\alpha^{(j+1)})'\Psi\alpha^{(j+1)}\big]
\end{aligned}$$

Calculation the derivation:

$$\frac{\partial A_1\hat{\mu}_x'\alpha^{(j+1)}}{\partial\alpha^{(j+1)}} = A_1\hat{\mu}_x'$$

$$\frac{\partial(\frac{A_2}{2}[(\alpha^{(j+1)})'(\hat{\Sigma}_x + \hat{\mu}_x\hat{\mu}_x')\alpha^{(j+1)})}{\partial\alpha^{(j+1)}} = A_2(\alpha^{(j+1)})'(\hat{\Sigma}_x + \hat{\mu}_x\hat{\mu}_x')$$

$$\frac{\partial\frac{\lambda}{2}(\alpha^{(j+1)})'\Psi\alpha^{(j+1)}}{\partial\alpha^{(j+1)}} = \lambda(\alpha^{(j+1)})'\Psi$$

$$\Psi = diag\{\frac{1}{\alpha_1^{(j)}}, ..., \frac{1}{\alpha_p^{(j)}}\}$$

Therefore,

$$A_1\hat{\mu}_x' + A_2(\alpha^{(j+1)})'(\hat{\Sigma}_x + \hat{\mu}_x\hat{\mu}_x') - \lambda(\alpha^{(j+1)})'\Psi \equiv 0$$

$$\Longrightarrow \alpha^{(j+1)} = A_1\big(\lambda\Psi - A_2(\hat{\Sigma}_x + \hat{\mu}_x\hat{\mu}_x')\big)^{-1}\hat{\mu}_x$$

Attempt 2

1. Initial $\alpha^{(0)}$ and λ
2. Fit LME, and estimate $\beta_1^{(0)}, \beta_2^{(0)}, \Gamma_1^{(0)}, \Gamma_2^{(0)}, D_1^{(0)}, D_2^{(0)}$
3. Calculate $A_1^{(0)}, A_2^{(0)}, B_1^{(0)}, B_2^{(0)}, B^{(0)}$
4. Update α: for the ith iteration ($i = 0, 1, 2, ...$),
   - for each $j \in \{1, 2, ..., p\}$, calculate $\alpha_j^{(i+1)}$ and update $\alpha_j^{(i+1)}$
5. Calculate $w = \alpha_j^{(i+1)} x$ and fit new LME, update $\beta_1^{(i+1)}, \beta_2^{(i+1)}, \Gamma_1^{(i+1)}, \Gamma_2^{(i+1)}, D_1^{(i+1)}, D_2^{(i+1)}$
6. Repeat 3-5 until convergence.

$$\hat{\alpha}_j^{(i+1)} = \begin{cases} -\frac{1}{(B_j^{(i)})' B_j^{(i)}} (B_j^{(i)})' B_{-j}^{(i)} \alpha_{-j}^{(i)} - \frac{B_{1,j}^{(i)}}{2(B_j^{(i)})' B_j} + \frac{\lambda}{2(B_j^{(i)})' B_j^{(i)}} & (B_j^{(i)})' B_{-j}^{(i)} \alpha_{-j}^{(i)} + \frac{B_{1,j}^{(i)}}{2} > \frac{\lambda}{2} \\ 0 & |(B_j^{(i)})' B_{-j}^{(i)} \alpha_{-j}^{(i)} + \frac{B_{1,j}^{(i)}}{2}| \leq \frac{\lambda}{2} \\ -\frac{1}{(B_j^{(i)})' B_j^{(i)}} (B_j^{(i)})' B_{-j}^{(i)} \alpha_{-j}^{(i)} - \frac{B_{1,j}^{(i)}}{2(B_j^{(i)})' B_j^{(i)}} - \frac{\lambda}{2(B_j^{(i)})' B_j^{(i)}} & (B_j^{(i)})' B_{-j}^{(i)} \alpha_{-j}^{(i)} + \frac{B_{1,j}^{(i)}}{2} < -\frac{\lambda}{2} \end{cases}$$

Parameter tunning

- For a given sequence of $\lambda \in \{0.001, 0.01, 0.1, 1, 10, 100\}$, find the $\hat{\alpha}$ that maximizes the penalized purity function.
- Do 10 fold cross validation. Use 9 of the part to train the data, fit LME with $\hat{\alpha}' x$ and calculate $\hat{\beta}_1, \hat{\beta}_2, \hat{\Gamma}_1, \hat{\Gamma}_2, \hat{D}_1, \hat{D}_2$.
- Calculate the IPWE for each of the 10 fold cross validation
- Repeat IPWE calculation for 100 times
- Choose the λ that achieve the max IPWE.

Update α calculation

Maximum value of

$$G_p(\alpha) = A_0 + A_1 \hat{\mu}_x' \alpha + \frac{A_2}{2} [\alpha' \hat{\Sigma}_x \alpha + \alpha' \hat{\mu}_x \hat{\mu}_x' \alpha] - \lambda ||\alpha||_1, \quad \lambda > 0$$

Or Minimum value of

$$G_p(\alpha) = -\alpha' B_2 \alpha - B_1 \alpha + \lambda ||\alpha||_1$$

where

- $B_1 = A_1 \mu_x'$
- $B_2 = \frac{A_2}{2} (\hat{\Sigma}_x + \hat{\mu}_x \hat{\mu}_x')$. And let $B_2 = B' B$

Then

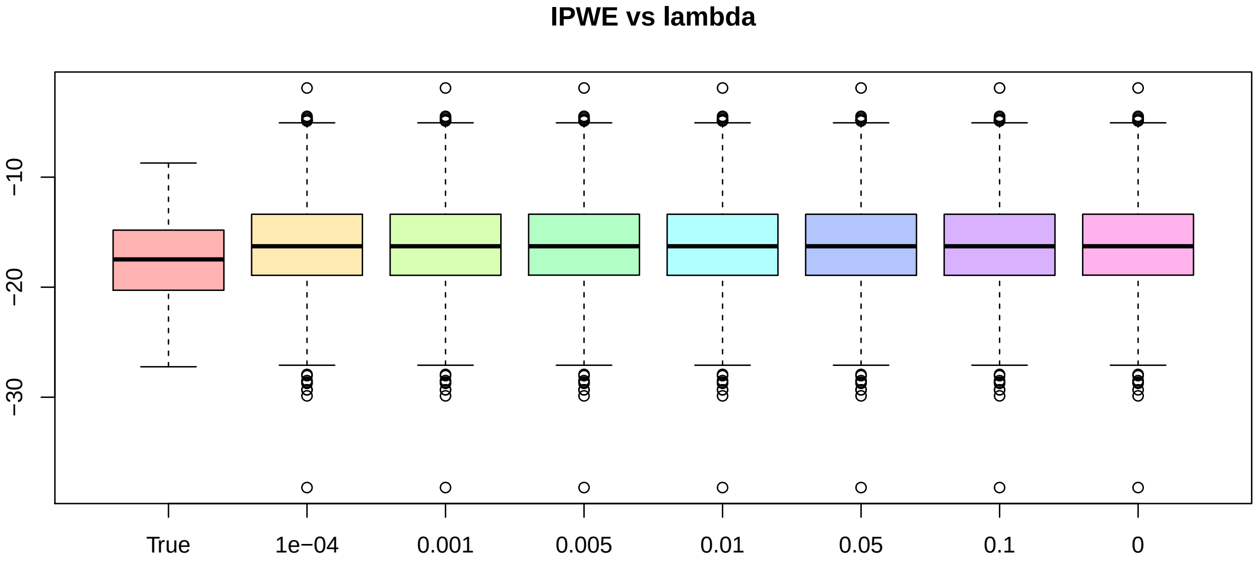
$$\frac{\partial G_p(\alpha)}{\partial \alpha_j} = -2B_j' B_j \alpha_j - 2B_j' B_{-j} \alpha_{-j} - B_{1,j} + \lambda \; sign(\hat{\alpha}_j) \equiv 0$$

$$\alpha_j = -\frac{1}{B_j' B_j} B_j' B_{-j} \alpha_{-j} - \frac{B_{1,j}}{2B_j' B_j} + \frac{\lambda}{2B_j' B_j} \; sign(\hat{\alpha}_j)$$

$\Longrightarrow$

$$\hat{\alpha}_j = \begin{cases} -\frac{1}{B_j' B_j} B_j' B_{-j} \alpha_{-j} - \frac{B_{1,j}}{2B_j' B_j} + \frac{\lambda}{2B_j' B_j} & B_j' B_{-j} \alpha_{-j} + \frac{B_{1,j}}{2} > \frac{\lambda}{2} \\ 0 & |B_j' B_{-j} \alpha_{-j} + \frac{B_{1,j}}{2}| \leq \frac{\lambda}{2} \\ -\frac{1}{B_j' B_j} B_j' B_{-j} \alpha_{-j} - \frac{B_{1,j}}{2B_j' B_j} - \frac{\lambda}{2B_j' B_j} & B_j' B_{-j} \alpha_{-j} + \frac{B_{1,j}}{2} < -\frac{\lambda}{2} \end{cases}$$

Boxplot of IPWE estimation vs different λ values



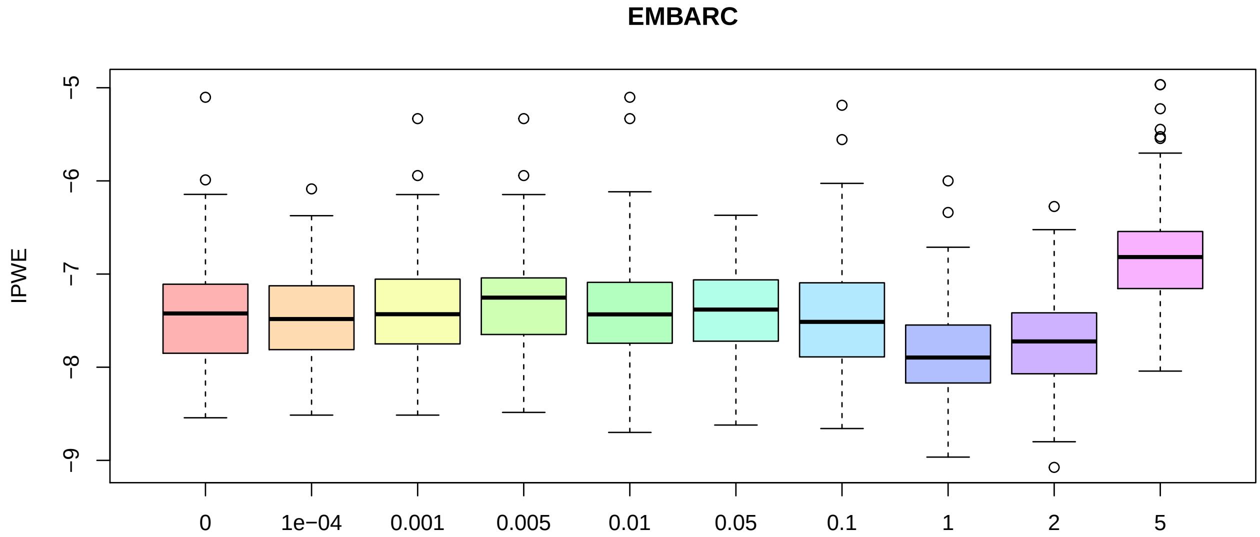
IPWE boxplot, with 100 covariates

If we input all the 100 covariates in the three models:

IPWE boxplot for EMBARC dataset

For example, the $\hat{\alpha}$ value

```
##   [1]  0.082 -0.085  0.023 -0.104  0.000  0.000 -0.047 -0.103 -0.054 -0.035
##  [11] -0.028  0.000 -0.074 -0.119  0.000  0.005 -0.126 -0.134 -0.139 -0.020
##  [21] -0.028 -0.068 -0.060 -0.061 -0.135 -0.124 -0.032 -0.095 -0.126 -0.059
##  [31] -0.004 -0.022 -0.049 -0.097  0.000 -0.103  0.000  0.000 -0.013 -0.119
##  [41] -0.043 -0.040 -0.028 -0.074  0.018 -0.063 -0.136  0.000 -0.011 -0.102
##  [51] -0.106 -0.131 -0.023  0.006 -0.086 -0.060 -0.067 -0.041 -0.137 -0.134
##  [61] -0.030 -0.116 -0.101 -0.048 -0.030 -0.031 -0.052 -0.089 -0.021  0.000
##  [71] -0.101 -0.090 -0.110 -0.031 -0.060 -0.064 -0.019  0.000 -0.105 -0.080
##  [81] -0.128 -0.127 -0.123 -0.033 -0.028 -0.061  0.003 -0.010  0.000  0.000
##  [91] -0.078 -0.072 -0.114 -0.024  0.005 -0.061 -0.055 -0.025  0.000 -0.024
## [101] -0.088 -0.030 -0.076 -0.021 -0.112 -0.066 -0.098 -0.065  0.023 -0.040
## [111]  0.000  0.000 -0.025 -0.076 -0.017 -0.084 -0.035 -0.003 -0.127 -0.094
## [121] -0.074 -0.033 -0.045 -0.032 -0.030 -0.055 -0.043  0.000 -0.059 -0.014
## [131] -0.049 -0.009  0.020 -0.062 -0.103 -0.077 -0.060 -0.046 -0.062 -0.113
## [141] -0.021 -0.084 -0.020  0.026 -0.074 -0.103 -0.104 -0.063 -0.073 -0.072
## [151] -0.007  0.000 -0.070 -0.093 -0.108 -0.130 -0.125 -0.023 -0.029 -0.068
## [161]  0.004  0.016  0.000  0.000 -0.042 -0.111 -0.096 -0.027  0.021 -0.063
## [171] -0.040  0.020  0.000 -0.034 -0.102 -0.045 -0.079 -0.015 -0.135 -0.060
## [181] -0.045 -0.069  0.000 -0.011  0.015  0.000 -0.045 -0.103 -0.060 -0.081
## [191]  0.005 -0.040 -0.108 -0.098 -0.081 -0.035 -0.037 -0.018 -0.021 -0.068
## [201] -0.067 -0.016 -0.066 -0.014 -0.093 -0.038 -0.003  0.003 -0.085 -0.040
## [211] -0.031 -0.041 -0.065 -0.124 -0.030  0.057  0.022  0.000  0.000  0.000
## [221]  0.009  0.000  0.000 -0.015 -0.005 -0.024 -0.013  0.007  0.015  0.000
## [231]  0.000  0.000  0.000
```


EMBARC
IPWE
−5
−6
−7
−8
−9
0
1e−04
0.001
0.005
0.01
0.05
0.1
1
2
5