Results

2020-05-28

Outline

1. Simulation if we have different α_{drg} and α_{pbo} in data generation
2. EMBARC data analysis with orthogonal ploynomial transformation
3. Variable selection in linear mixed effect model

Different α_{drg} and α_{pbo}

i. 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
- $x \sim MVN(\mu_x, \Sigma_x)$, $\mu_x = \mathbf{0}_p$, Σ_x has diagonal equals to 1 and 0.5 anywhere else. The dimension of x is set to be $p = 2, 4, 8, 16$.
- $\beta_{drg} = [40, -1, -0.02], \beta_{pbo} = [40, -1.1, -0.02]$
- $\Gamma_{drg} = [0, 0.1\cos(\theta), 0.1\sin(\theta)], \Gamma_{pbo} = [0, 0.01, 0], \theta \in \{0, \frac{\pi}{3}, \frac{\pi}{2}, \frac{2\pi}{3}, \pi\}$
- $b_{drg} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.01 & 0 \\ 0 & 0 & 0.02 \end{bmatrix}, b_{pbo} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.01 & 0.005 \\ 0 & 0.005 & 0.01 \end{bmatrix}$
- $\epsilon_{drg}, \epsilon_{pbo} \sim N(0, 3^2)$
- $\alpha_{drg} = \mathbf{1}_p$, $\alpha_{pbo} = [1, 2, .., p]$, both of them are then transformed to have norm equals to 1.

ii. Estimation

We use the old approach, which assumes the α is the same in both drug group and placebo group.

The dataset purity is calculated based on the shared α as

$$G(\alpha) = E_w(g(w)) = A_0 + A_1\mu'_x\alpha + \frac{A_2}{2}\left[\alpha'\Sigma_x\alpha + \alpha'\mu_x\mu'_x\alpha\right] \tag{1}$$

where

$$A_0 = -q + \frac{1}{2}tr(D^{-1}_{pbo}D_{drg}) + \frac{1}{2}tr(D^{-1}_{drg}D_{pbo}) + \frac{1}{2}(\beta_{drg} - \beta_{pbo})'(D^{-1}_{drg} + D^{-1}_{pbo})(\beta_{drg} - \beta_{pbo})$$

$$A_1 = (\beta_{drg} - \beta_{pbo})'(D^{-1}_{drg} + D^{-1}_{pbo})(\gamma_{drg} - \gamma_{pbo})$$

$$A_2 = (\gamma_{drg} - \gamma_{pbo})'(D^{-1}_{drg} + D^{-1}_{pbo})(\gamma_{drg} - \gamma_{pbo})$$

q is the dimension of D matrix.

iii. Simulation

Scenarios: $p = 2, 4, 8, 16$, $\theta = 0, \frac{\pi}{3}, \frac{\pi}{2}, \frac{2\pi}{3}, \pi$.

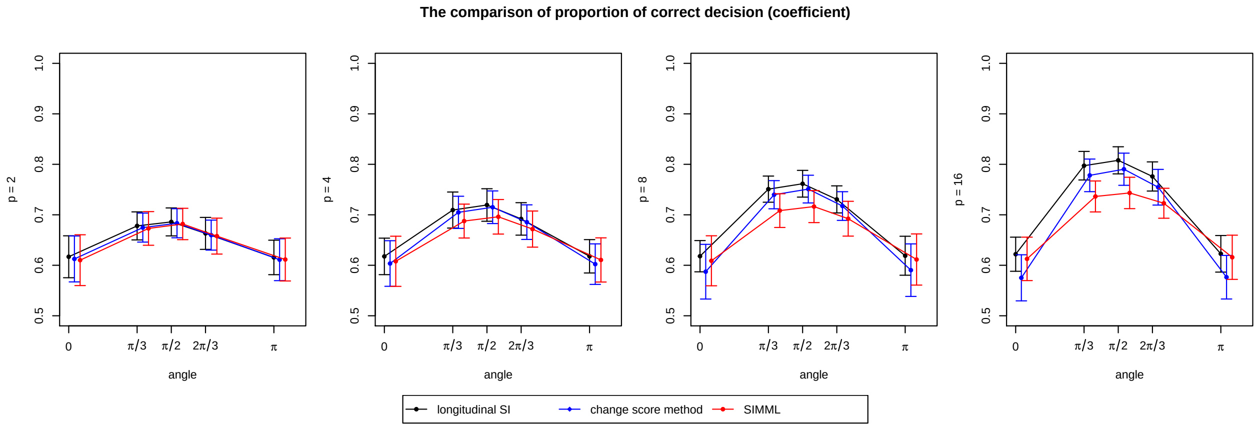
Repetition: 100

Metric: proportion of correct decision, IPWE (10 CV).

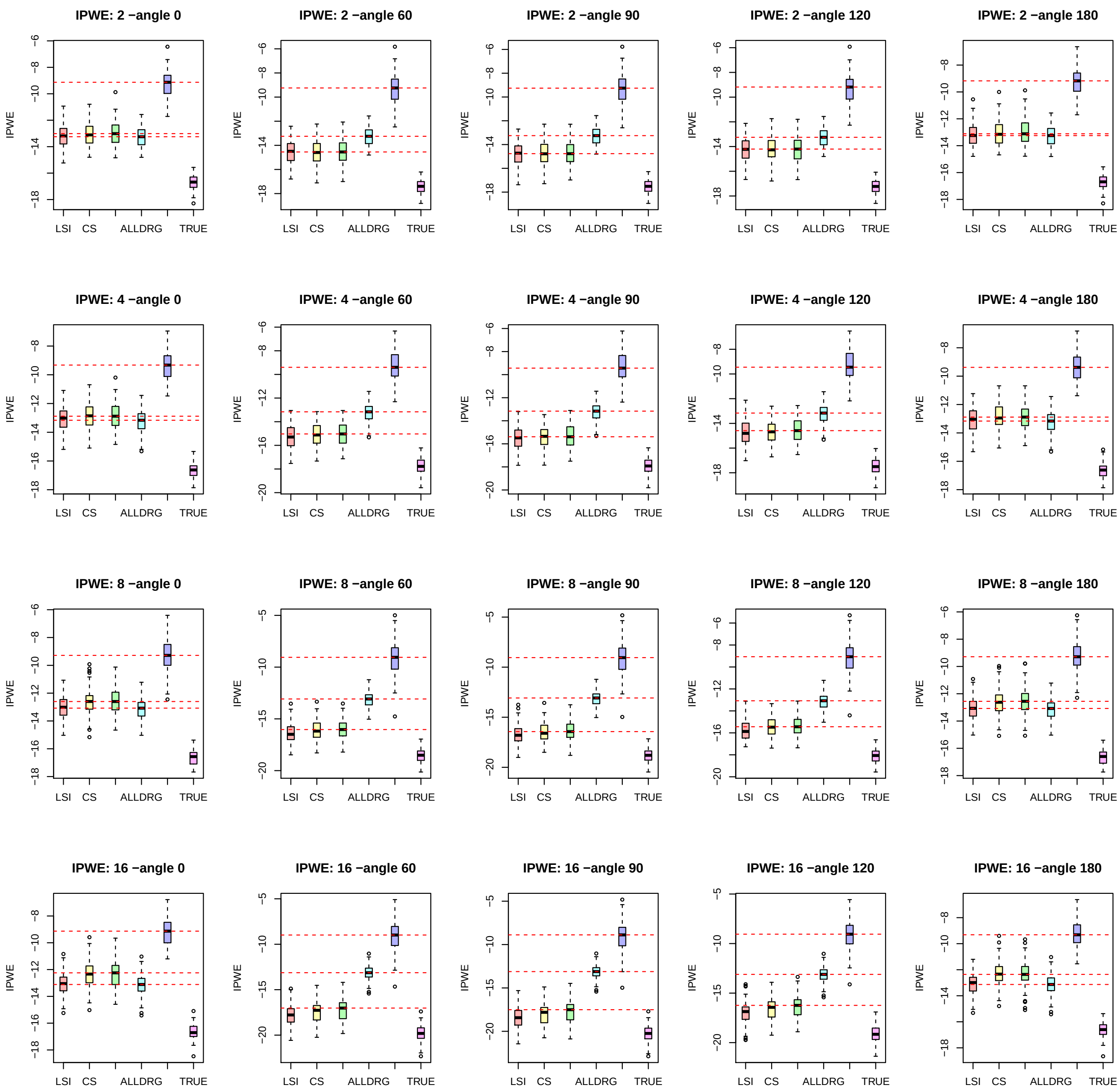
iv. Results

PCD (coefficient)

		Longitudinal		Change score		SIMML	
p	theta	mean	sd	mean	sd	mean	sd
2	0	0.617	0.021	0.613	0.023	0.610	0.026
	60	0.678	0.014	0.675	0.015	0.673	0.017
	90	0.686	0.014	0.683	0.015	0.682	0.016
	120	0.663	0.016	0.660	0.015	0.658	0.018
	180	0.616	0.017	0.611	0.021	0.611	0.022
4	0	0.618	0.018	0.604	0.023	0.608	0.025
	60	0.709	0.018	0.705	0.016	0.688	0.017
	90	0.719	0.016	0.715	0.016	0.696	0.017
	120	0.692	0.016	0.685	0.017	0.672	0.018
	180	0.618	0.017	0.602	0.020	0.611	0.022
8	0	0.618	0.016	0.587	0.028	0.609	0.025
	60	0.751	0.013	0.740	0.014	0.708	0.017
	90	0.762	0.014	0.751	0.014	0.716	0.016
	120	0.731	0.014	0.717	0.014	0.692	0.018
	180	0.619	0.020	0.590	0.027	0.611	0.026
16	0	0.622	0.017	0.575	0.023	0.613	0.022
	60	0.797	0.014	0.778	0.016	0.736	0.016
	90	0.808	0.014	0.790	0.016	0.743	0.016
	120	0.776	0.015	0.755	0.018	0.723	0.015
	180	0.623	0.018	0.576	0.022	0.616	0.022



IPWE (coefficient)



PCD (outcome)

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

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

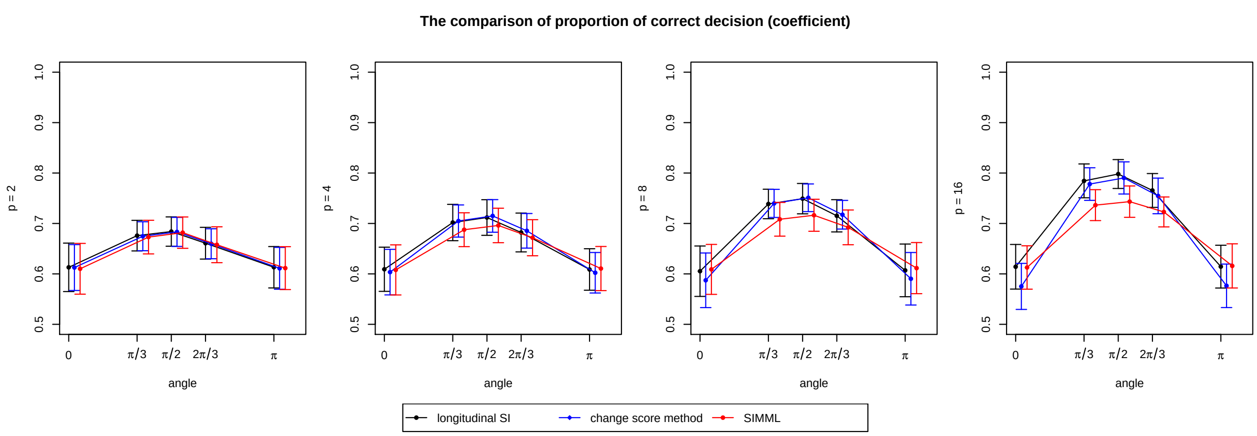
$$Y_{pbo} = S(\beta_{pbo} + \Gamma_{pbo}(\alpha'_{pbo}x)) + Zb_{pbo} + \epsilon_{pbo}$$

where

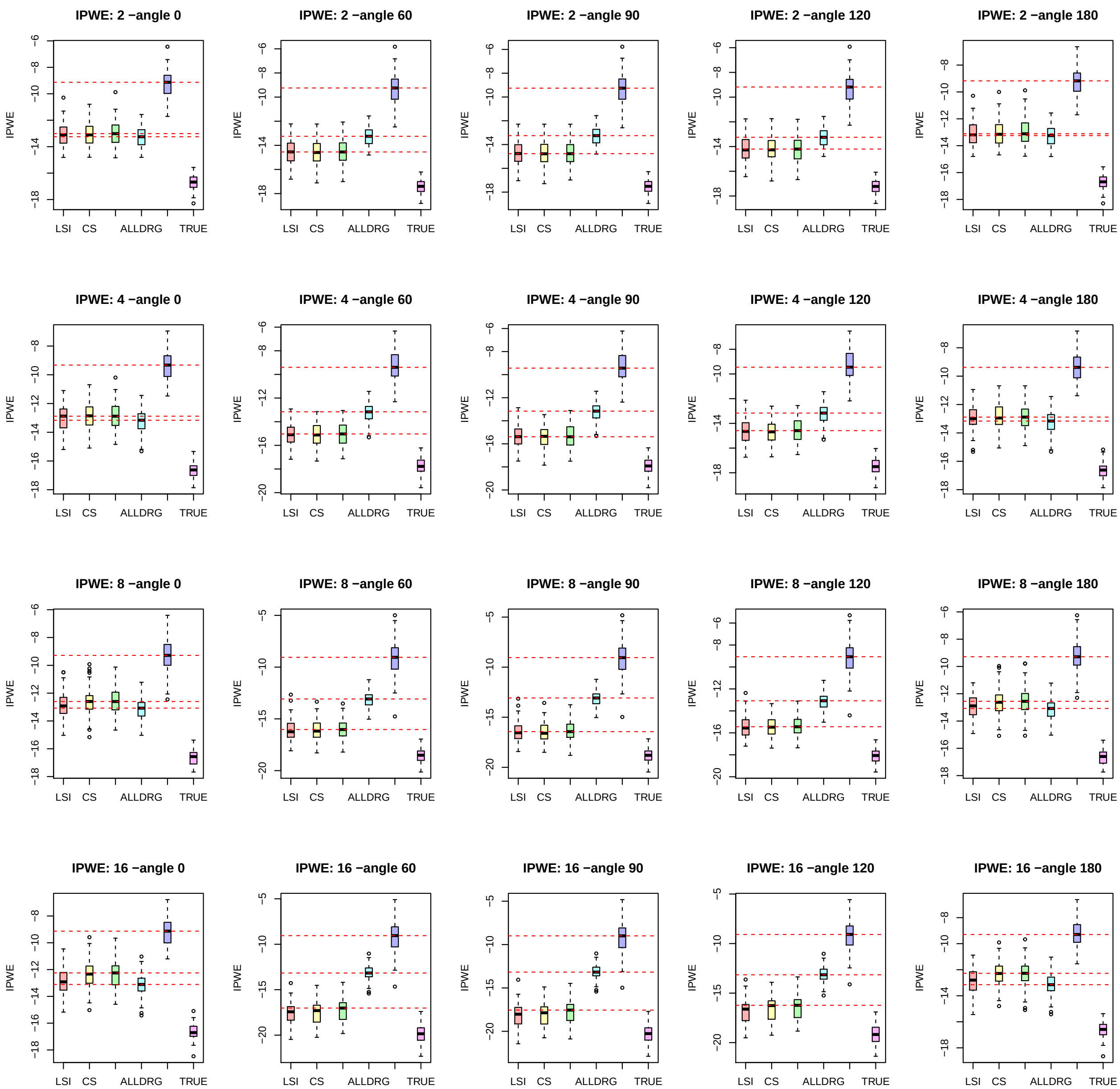
- $S = [1, t, t^2]$, $t = [0, 1, 2, 3, 4, 6, 8]$ is the design matrix for fixed effect
- $Z = [1, t]$, $t = [0, 1, 2, 3, 4, 6, 8]$ is the design matrix for random effect

		Longitudinal		Change score		SIMML	
p	theta	mean	sd	mean	sd	mean	sd

2	0	0.613	0.025	0.613	0.023	0.610	0.026
	60	0.676	0.015	0.675	0.015	0.673	0.017
	90	0.684	0.015	0.683	0.015	0.682	0.016
	120	0.661	0.016	0.660	0.015	0.658	0.018
	180	0.613	0.021	0.611	0.021	0.611	0.022
4	0	0.609	0.022	0.604	0.023	0.608	0.025
	60	0.702	0.018	0.705	0.016	0.688	0.017
	90	0.712	0.018	0.715	0.016	0.696	0.017
	120	0.682	0.020	0.685	0.017	0.672	0.018
	180	0.609	0.021	0.602	0.020	0.611	0.022
8	0	0.605	0.026	0.587	0.028	0.609	0.025
	60	0.739	0.015	0.740	0.014	0.708	0.017
	90	0.749	0.015	0.751	0.014	0.716	0.016
	120	0.715	0.016	0.717	0.014	0.692	0.018
	180	0.607	0.027	0.590	0.027	0.611	0.026
16	0	0.614	0.023	0.575	0.023	0.613	0.022
	60	0.784	0.017	0.778	0.016	0.736	0.016
	90	0.798	0.015	0.790	0.016	0.743	0.016
	120	0.765	0.017	0.755	0.018	0.723	0.015
	180	0.614	0.022	0.576	0.022	0.616	0.022



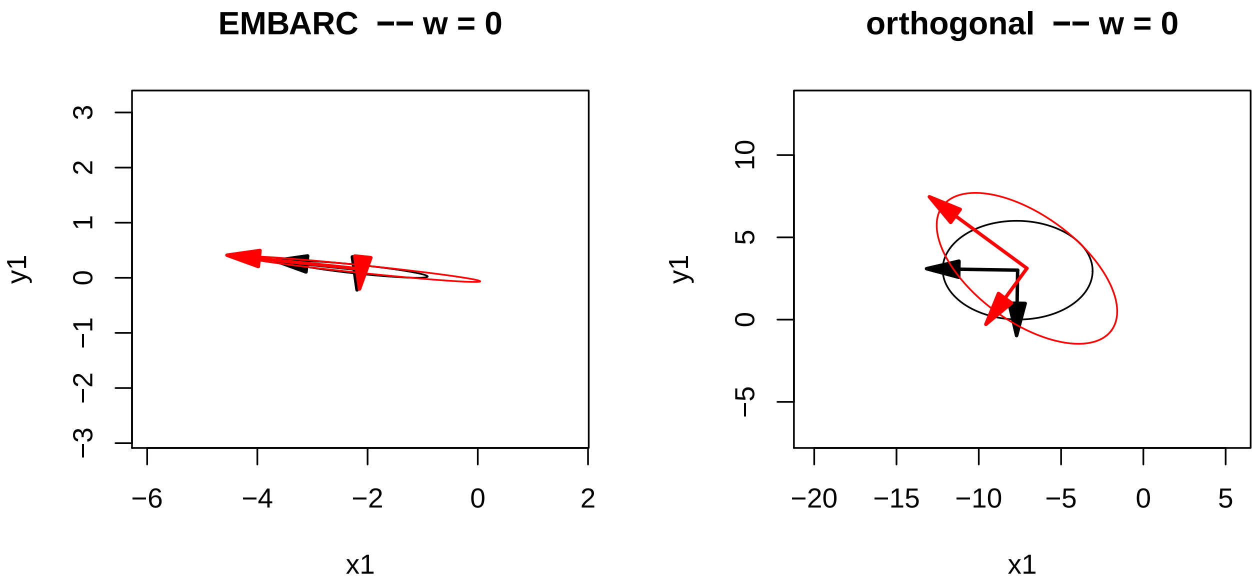
IPWE (outcome)



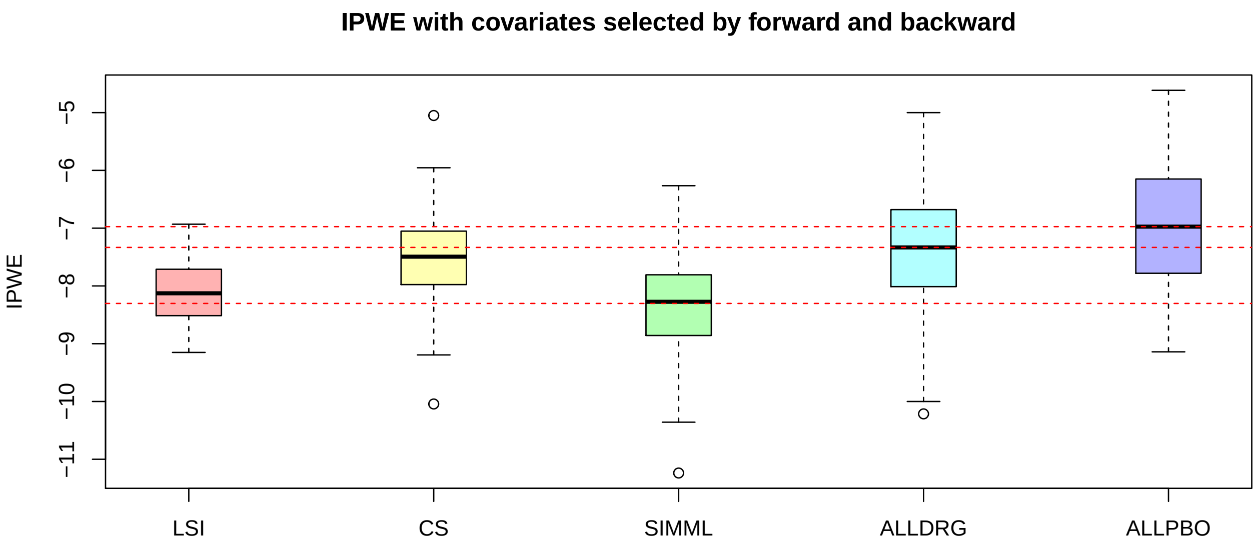
Orthogonal polynomial transforms

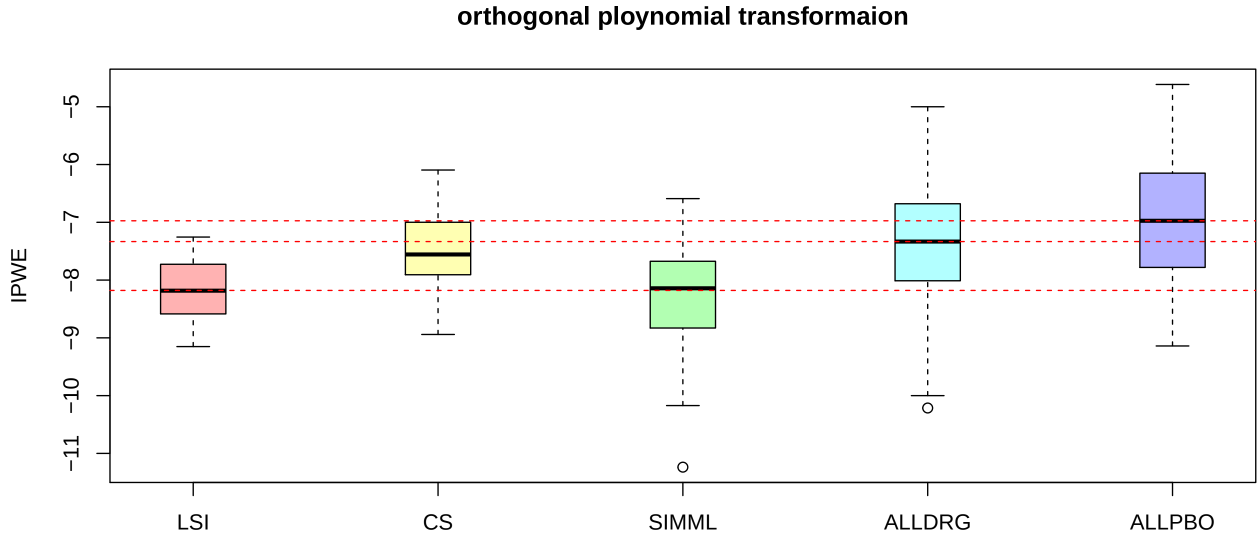
Apply the orthogonal polynomial transformation for the estimation of β, Γ, D when optimize the α and calculate the IPWE.

ellipse


EMBARC -- w = 0
orthogonal -- w = 0
y1
x1
−3
−2
−1
0
1
2
3
−6
−4
−2
0
2
−5
0
5
10
−20
−15
−10
−5
0
5

IPWE


IPWE with covariates selected by forward and backward
IPWE
−5
−6
−7
−8
−9
−10
−11
LSI
CS
SIMML
ALLDRG
ALLPBO



LASSO for linear mixed effect model

For a linear mixed effect, with $N = \sum_{i=1}^{n} n_i$ subjects and n clusters. For the ith cluster, the linear mixed effect model can be fitted as

$$y_i = X_i\beta + Z_i b_i + \epsilon_i \tag{2}$$

where

- y_i is an $n_i \times 1$ vector of responses for cluster i.
- β is a $p \times 1$ vector of fixed effects
- X_i is the design matrix for fixed effect with dimension $n_i \times p$
- Z_i is the design matrix for random effect with dimension $n_i \times q$
- b_i is the random effect, $b_i \sim N(0, \sigma D^2)$, D is positive definite.
- ϵ_i is the ranodm error $\sim N(0, \sigma I_{n_i})$
- Thus y_i has a distribution as $y_i \sim N(X_i\beta, \sigma^2 V_i(\theta))$, $V_i(\theta) = I_{n_i} + Z_i D Z_i'$, θ denotes the $k = \frac{q(q+1)}{2}$ covariance parameters in D.

Therefore, the parameter space is $\Omega = (\beta, \theta, \sigma)$. The log-lielihood function for equation 1 can be writen as:

$$l_F(\beta, \theta, \sigma) = -\frac{1}{2}\sum_{i=1}^{n} \log|\sigma^2 V_i| - \frac{1}{2}\sigma^{-2}\sum_{i=1}^{n} r_i' V_i^{-1} r_i,\ \ r_i = y_i - X_i\beta \tag{3}$$

When θ is known, the MLE of β is

$$\hat{\beta} = (\sum_{i=1}^{n} X_i' V_i^{-1} X_i)^{-1} (\sum_{i=1}^{n} X_i' V_i^{-1} y) \tag{4}$$

To get the accurate estimates of the variance components, REML estimation is preferred. The restricted log-likelihood function can be writen as,

$$l_R(\theta, \sigma) = l_F(\hat{\beta}, \theta, \sigma) - \frac{1}{2}\log|\sigma^2 \sum_{i=1}^{n} X_i' V_i^{-1} X_i| \tag{5}$$

where $\hat{\beta}$ has the form $\hat{\beta} = (\sum_{i=1}^{n} X_i' V_i^{-1} X_i)^{-1} (\sum_{i=1}^{n} X_i' V_i^{-1} y)$.

Then we can obtain V_i from the REML as $\hat{V}_{Ri}$, and the REML estimator of β is

$$\hat{\beta}_R = (\sum_{i=1}^{n} X_i' \hat{V}_{Ri}^{-1} X_i)^{-1} (\sum_{i=1}^{n} X_i' \hat{V}_{Ri}^{-1} y) \tag{6}$$

The estimation of σ^2 by ML and REML are

$$\sigma^2_{ML} = \frac{1}{N} \sum_{i=1}^{n} r_i' V_i^{-1} r_i \tag{7}$$

$$\sigma^2_{REML} = \frac{1}{N-p} \sum_{i=1}^{n} r_i' V_i^{-1} r_i \tag{8}$$

substitue them in the equation and we can get the profile log-lielihood function

$$p_F(\beta, \theta) = -\frac{1}{2} \sum_{i=1}^{n} \log |\sigma^2 V_i| - \frac{N}{2} \log(\sum_{i=1}^{n} r_i' V_i^{-1} r_i) \tag{9}$$

and restricted profile log-likehood function

$$p_R(\theta) = -\frac{1}{2} \log |\sum_{i=1}^{n} X_i' V_i^{-1} X_i| - \frac{1}{2} \sum_{i=1}^{n} \log |V_i| - \frac{1}{2}(N-p) \log \big[\sum_{i=1}^{n} (y_i - X_i\hat{\beta})' V_i^{-1} (y_i - X_i\hat{\beta})\big] \tag{10}$$

Variable selection

i. Selection of random effect

Maximize the penalized restricted profile log-likelihood

$$p_R(\theta) = -\frac{1}{2} \log |\sum_{i=1}^{n} X_i' V_i^{-1} X_i| - \frac{1}{2} \sum_{i=1}^{n} \log |V_i| - \frac{1}{2}(N-p) \log \big[\sum_{i=1}^{n} (y_i - X_i\hat{\beta})' V_i^{-1} (y_i - X_i\hat{\beta})\big]$$

$$Q_R(\theta) = p_R(\theta) - \lambda_{1n} \sum_{j=1}^{q} w_{1j} |d_j| \tag{11}$$

where $d = (d_1, .., d_q)$, the vector consisting the diagonal elements of D.

Apply Newton-Raphson algorithm.

ii. Selection of fixed effect

$$p_F(\beta, \theta) = -\frac{1}{2} \sum_{i=1}^{n} \log |\sigma^2 V_i| - \frac{N}{2} \log(\sum_{i=1}^{n} r_i' V_i^{-1} r_i)$$

The covariance matrix of random effects V is estimated by $\hat{V}$ in the above step, then drop the constant terms and the profile likelihood function can be writen as

$$p_F(\beta, \theta) = -\frac{N}{2} \log(\sum_{i=1}^{n} r_i' V_i^{-1} r_i) \tag{12}$$

And we could have the first and second derivatives of $p_F(\beta,\theta)$.

To determine the set of covariates for fixed effects, we can maximize the penalized profile log-likelihood fucntion,

$$Q_F(\beta) = p_F(\beta) - \lambda_{2n}\sum_{j=1}^{p} w_{2j}|\beta_j| \tag{13}$$

We can then apply Newton-Raphson algorithm as

$$\beta_{b+1} = \beta_b - M_{\beta\beta}^{-1} SC_\beta, b = 0, 1, ... \tag{14}$$

where β_b is the current step value, SC_β is the $p \times 1$ vector of the first derivative, $M_{\beta\beta}$ is a $p \times p$ matrix of the second derivative. And we can use the approximiation

$$|\beta_j| \approx \frac{1}{2}|\beta_j^{(0)}| + \frac{1}{2}\frac{\beta_j^2}{|\beta_j^{(0)}|} \tag{15}$$

to make it derivable.

iii. Selection tuning parameter by minimize AIC, or BIC.

$$BIC_R = -2p_R(\hat{\theta}) + \log(N) \times df_R \tag{16}$$

$$BIC_F = -2p_F(\hat{\theta}) + \log(N) \times df_F \tag{17}$$

Variable selection for longitudinal single index method

1. initial α, intial the tunning parameter λ

for the jth iteration, $j = 0, 1, ...$

2. Fit the model
$$Y = S(\beta + b + \Gamma(\alpha^{(j)}x)) + \epsilon$$
3. Estimate the random effect by maximize the restrict profile log-likelihood function with Newton-Raphson algorithm, $\hat{D}_{drg}, \hat{D}_{pbo}$
4. Estimate the fixed effect, plug in the estimated random effect into the model, and maximize the penalized profile log-likelihood fucntion,
$$Q_F(\beta) = p_F(\beta) - \lambda_{2n}\sum_{j=1}^{p} w_{2j}|\beta_j|$$
5. Select the tunning parameter by minimize the BIC_F
6. Estimate the α in the purity function.

$$G(\alpha) = E_w(g(w)) = A_0 + A_1\mu_x'\alpha + \frac{A_2}{2}\left[\alpha'\Sigma_x\alpha + \alpha'\mu_x\mu_x'\alpha\right] \tag{18}$$

where

$$A_0 = -q + \frac{1}{2}tr(\hat{D}_{pbo}^{-1}\hat{D}_{drg}) + \frac{1}{2}tr(\hat{D}_{drg}^{-1}\hat{D}_{pbo}) + \frac{1}{2}(\hat{\beta}_{drg} - \hat{\beta}_{pbo})'(\hat{D}_{drg}^{-1} + \hat{D}_{pbo}^{-1})(\hat{\beta}_{drg} - \hat{\beta}_{pbo})$$

$$A_1 = (\hat{\beta}_{drg} - \hat{\beta}_{pbo})'(\hat{D}_{drg}^{-1} + \hat{D}_{pbo}^{-1})(\hat{\Gamma}_{drg} - \hat{\Gamma}_{pbo})$$

$$A_2 = (\hat{\Gamma}_{drg} - \hat{\Gamma}_{pbo})'(\hat{D}_{drg}^{-1} + \hat{D}_{pbo}^{-1})(\hat{\Gamma}_{drg} - \hat{\Gamma}_{pbo})$$

q is the dimension of D matrix.

Repeat the above process under converge.