Prediction

2020-02-02

How does the method work in terms of prediction?

We simulate a train dataset with $n = 100$. Then estimate the $\hat{\beta}_{drg}, \hat{\beta}_{pbo}, \hat{\Gamma}_{drg}, \hat{\Gamma}_{pbo}, \hat{D}_{drg}, \hat{D}_{pbo}$. We then simulate a test dataset with $n = 1000$. For each subject with covariates $\alpha' x_i$, we calculate the associated estimated parabolas as:

$$\hat{z}_{i,drg} = X[\hat{\beta}_{drg} + \hat{\Gamma}_{drg}\alpha' x_i]$$

$$\hat{z}_{i,pbo} = X[\hat{\beta}_{pbo} + \hat{\Gamma}_{pbo}\alpha' x_i]$$

Since the data is simulated, we also know subject's true parabolas as:

$$z_{i,drg} = X[\beta_{drg} + \Gamma_{drg}\alpha' x_i]$$

$$z_{i,pbo} = X[\beta_{pbo} + \Gamma_{pbo}\alpha' x_i]$$

We then calculate the slope summation of the tangent lines within interval [week_0, week_7] as the scalar measure for the trajectories: $s_{drg}, s_{pbo}, \hat{s}_{drg}, \hat{s}_{pbo}$.

If $s_{drg} < s_{pbo}$, the drug works better than the placebo and the subject should be assigned to the drug group. In the meantime, if $\hat{s}_{drg} < \hat{s}_{pbo}$, the subject can be assigned to the drug group based on our estimation. Then the estimation is agreed with the true assignment.

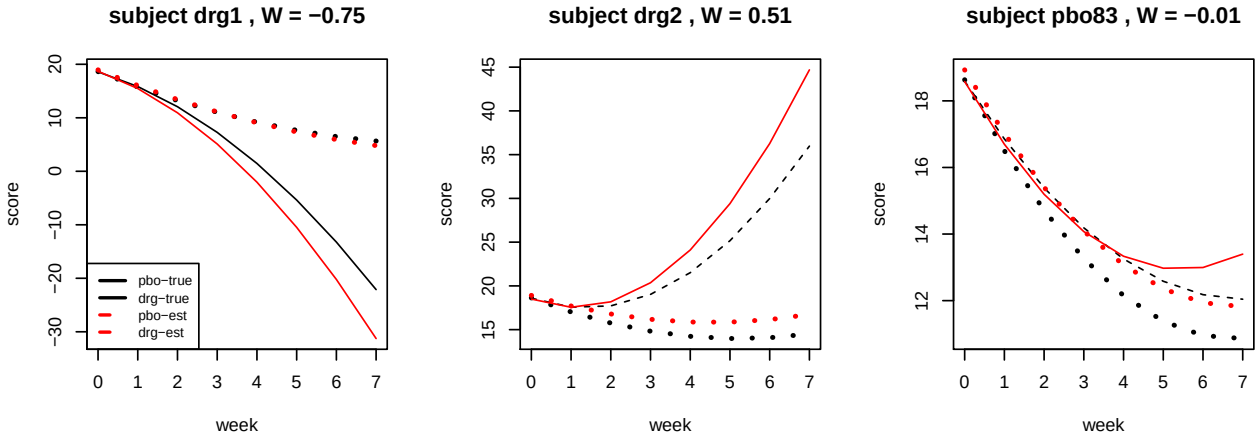
If $s_{drg} > s_{pbo}$, the placebo works better than the drug and the subject should not be assigned to the drug group. In the meantime, if $\hat{s}_{drg} > \hat{s}_{pbo}$, the subject can be assigned to the placebo group based on our estimation. Then the estimation is agreed with the true assignment.

Otherwise, the estimation is not agreed with the true assignment.

1000 repetitions are run, with dimension of covariates varys from 2 to 16. The percentage of agreement is calculated and shown at the last column.

		Purity			Cosine similarity		
Dimension	Angle	True	Mean	SD	Mean	SD	Agreement
2	0	0.542	1.590	0.591	0.964	0.072	0.944
	60	239.225	267.687	55.729	0.999	0.001	0.993
	120	184.880	206.153	42.513	0.999	0.001	0.992
	180	19.363	21.978	4.110	0.997	0.004	0.983
4	0	1.458	1.684	0.610	0.929	0.093	0.941
	60	396.347	448.381	89.557	0.999	0.001	0.993
	120	305.771	347.099	69.841	0.999	0.001	0.992
	180	29.910	36.460	6.647	0.992	0.007	0.980
8	0	5.458	1.815	0.618	0.902	0.099	0.928
	60	710.592	821.330	168.627	0.999	0.001	0.994
	120	547.555	654.124	147.529	0.999	0.001	0.994
	180	51.005	66.120	11.770	0.984	0.010	0.980
16	0	13.458	1.819	0.685	0.941	0.093	0.941
	60	1339.080	1658.016	412.442	0.999	0.001	0.996
	120	1031.121	1304.064	383.379	0.999	0.001	0.996
	180	93.194	132.970	28.201	0.976	0.012	0.983

Following are several plots to show the estimated trajectories and the true trajectories.



All the above plots meet the agreement of true assignment.