Meeting: Robust TFBoost

July 11, 2021

- Competitors:
  - **Maronna and Yohai (2013)**: MM-estimator for functional linear model (code?)
  - Shin and Lee (2016): M-estimator for functional linear model (slope function in the RKHS, M estimator using MAD as the residual scale.) Compared to Maronna and Yohai (2013), they studied asymptotic properties of the estimator. Why want to assume RKHS?
  - Qingguo (2017): FPC scores + M-estimator (tuning constant selected to minimize a criterion)
  - **Huang et al. (2015)**: sieve M-estimators for semi-functional model. (have code)
  - Cai et al. (2020): a new Huber loss with data-driven tuning parameters. "enables us to reach better robustness and efficiency than other robust methods in the presence of outliers or heavy-tailed error distribution"
  - **Boente et al. (2020)**: MM-estimator for semi-functional linear model. (have code)
  - Wang et al. (2017): functional sliced inverse regression (trimmed estimation of the mean and covariance) (no reply)
  - Kalogridis and Van Aelst (2019): robust FPC scores + MM functional linear model.
  - functional quantile regression papers (no reply)
  - **FLM1**, **FLM2**, **FAM**, **FAME**, **FPPR**, **FGAM**
- Contaminated Settings:
  - **Boente et al. (2020)**: regression outliers and high-leverage outliers (extreme values in the scores), compared with Huang
  - **Huang et al. (2015)** only vertical outliers, t(3) and Cauchy distributed errors (comparing OLS, LAD, Huber and T type estimators)

- **Maronna and Yohai (2013)** glass data, compared with partial least squares and Crambes et al. (2008). considered high-leverage outliers.
- Kalogridis and Van Aelst (2019) considered bad laverage points in the same way as Maronna and Yohai (2013) competitors: Maronna and Yohai (2013), FPCA regression, and Shin and Lee (2016).
- Qingguo (2017) vertical outliers
- Shin and Lee (2016) vertical outliers
- Wang et al. (2017) add errors to curves only (strange)

- High-leverage outliers

$$y_i = \int \beta(t)x_i(t)dt + \sigma_0\epsilon_i$$

$$\beta(t) = \sum_{j-1}^{50} b_j\phi_j,$$

where $\phi_1(t) = 1$ and $\phi_j(t) = \sqrt{2}\cos((j-1)\pi t), j \geq 2.$

$$b_1 = 0.3, b_j = 4(-1)^{j+1}j^{-2}, j \geq 2$$

x_i is generated from Gaussian process with mean 0 and eigenfunctions $\phi_j(t)$, scores $\xi_{ij} \sim N(0, j^{-2})$, $\sigma_0 = 1$, and $\epsilon_i \sim N(0,1)$.

We consider four types of outliers ($C = 10$)

1. Shape outliers
   10% data, $\epsilon_{i_{\text{out}}} \sim N(C, 0.25)$, $\xi_{i_{\text{out}},2} \sim N(C/2, 0.25)$, the other parameters stay the same.
2. Magnitude outliers
   10% data $x_{i_{\text{out}}} = 2x_{i_{\text{out}}}, y_{i_{\text{out}}} = 1.5y_{i_{\text{out}}}$
3. Measurement error outliers
   10% data, 10% time points $t_{i,\text{out}}$
   $\epsilon_{i_{\text{out}}} \sim N(C, 0.25)$, $x_{i_{\text{out}}}(t_{i,\text{out}}) = x_i(t_{i,\text{out}}) + \eta_{i_{\text{out}}}$, where $\eta_{i_{\text{out}}} \sim 0.5N(10, 0.25) + 0.5N(-10, 0.25)$
4. Pure vertical outliers
   10% data, $\epsilon_{i_{\text{out}}} \sim N(C, 0.25)$

- Type A tree? not considering it may cause issues in the linear setting

- Questions about Thesis writing
  - Background:
  - Timeline?

- Action items:
  - RobustFPLM code
  - Schedule for a committee meeting?
  - Ask Kalogridis and Van Aelst (2019) for the code? ...
  - Ask Maronna and Yohai (2013) for the code of Crambes et al. (2008).

References

Boente, G., Salibian-Barrera, M., and Vena, P. (2020). Robust estimation for semi-functional linear regression models. *Computational Statistics & Data Analysis*, 152:107041.

Cai, X., Xue, L., and Wang, Z. (2020). Robust estimation with modified huber's function for functional linear models. *Statistics*, pages 1–11.

Crambes, C., Delsol, L., and Laksaci, A. (2008). Robust nonparametric estimation for functional data. *Journal of Nonparametric Statistics*, 20(7):573–598.

Huang, L., Wang, H., Cui, H., and Wang, S. (2015). Sieve m-estimator for a semi-functional linear model. *Science China Mathematics*, 58(11):2421–2434.

Kalogridis, I. and Van Aelst, S. (2019). Robust functional regression based on principal components. *Journal of Multivariate Analysis*, 173:393–415.

Maronna, R. A. and Yohai, V. J. (2013). Robust functional linear regression based on splines. *Computational Statistics & Data Analysis*, 65:46–55.

Qingguo, T. (2017). M-estimation for functional linear regression. *Communications in Statistics-Theory and Methods*, 46(8):3782–3800.

Shin, H. and Lee, S. (2016). An rkhs approach to robust functional linear regression. *Statistica Sinica*, pages 255–272.

Wang, G., Zhou, J., Wu, W., and Chen, M. (2017). Robust functional sliced inverse regression. *Statistical papers*, 58(1):227–245.