Application: ADNI

Charlie Song

1 Data

The ADNI data set records five variables: ADAS13, RAVLT_imm, RAVLT_learn, MMSE, FAQ for $n = 808$ subjects. The number of observations m_i for each subject is between 2 and 11 with total number of observation $\sum_i m_i = 4457, 4457, 4456, 4497, 4471$ and average number of observation $\sum_i m_i/n = 5.51, 5.51, 5.51, 5.57, 5.53$ for variable ADAS13, RAVLT_imm, RAVLT_learn, MMSE, FAQ respectively. We highlights three subjects in the study in solid black lines, which shows that many individuals drop out of the study half of the way.

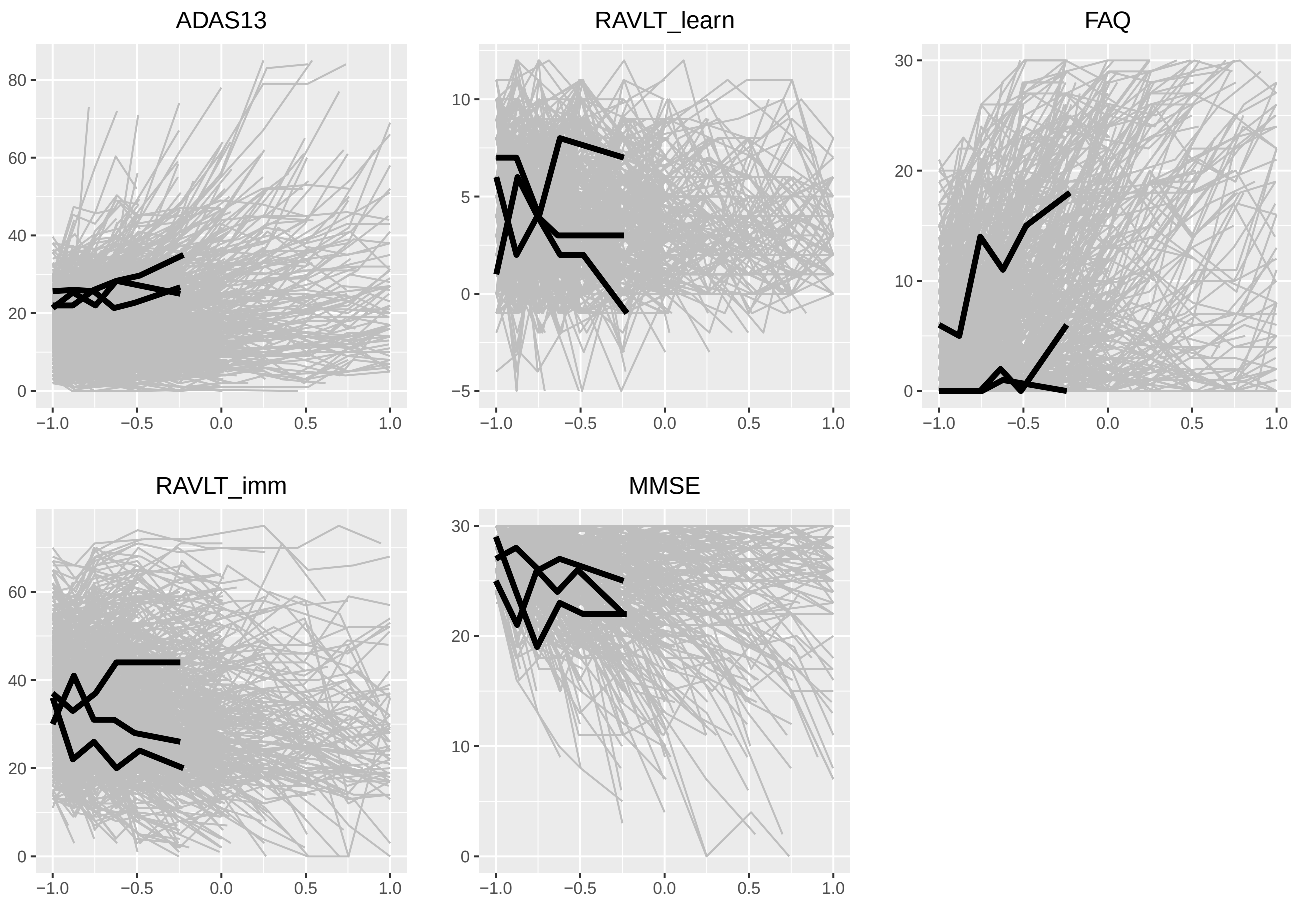

Figure 1: Spaghetti plots of ADAS13, RAVLT_imm, RAVLT_learn, MMSE, FAQ for 808 subjects with average 5.5 observation per subject ranging from 2 to 11. Three subjects' curves are highlighted for each variable.

2 Covariance Test

Table 3 summarizes the p-values testing the goodness-of-fit for each variables using various methods.

The bootstrap method is the original method implemented in Chen et al. (2019). The modified method estimates the alternative covariance matrices using FACE. The proposed methods improves the modified version and performs slightly different. The direct method tests against the alternative of a known parametric model using methods in Crainiceanu and Ruppert (2004). The alternative model can be written as:

$$X_i(t) = b_{0i} + b_{1i}t + b_{2i}t^2$$
$$(b_{0i}, b_{1i}, b_{2i})^\top \sim N(\mathbf{0}, \begin{bmatrix} \sigma_0^2 & \sigma_{01} & 0 \\ \sigma_{01} & \sigma_1^2 & 0 \\ 0 & 0 & \sigma_2^2 \end{bmatrix})$$

In the null hypothesis for all tests, we assume a linear random effects model nested within the alternative model with subject-specific random intercepts b_{0i} and slopes b_{1i}:

$$X_i(t) = b_{0i} + b_{1i}t$$
$$(b_{0i}, b_{1i})^\top \sim N(\mathbf{0}, \begin{bmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{01} & \sigma_1^2 \end{bmatrix})$$

3 Conclusion

The three goodness-of-fit tests of covariance matrix based on nonparametric methods (bootstrap, modified and proposed method) seem to suggest no nonparametric technique required to model the covariance matrices for all five variables of the ADNI dataset. However, the p-values in testing RAVLT_imm are closest to 0.01 among five variables in all three methods, which suggests possible accuracy gain from nonparametric model. Moreover, the bootstrap (original test in Chen et al. (2019)) methods yields smaller p-values than the other two methods for all variables because it suffers from inflated type I error rates when the functional data is sparsely observed.

The direct method testing against the alternative of a known parametric model shows that the alternative parametric model is preferred for ADAS13 and FAQ. The direct tests for RAVLT_imm

	Bootstrap	Modified	Proposed	Direct
ADAS13	0.178	0.314	0.291	$<$ **2.2e-16**
RAVLT_imm	**0.098**	0.161	0.152	NA
RAVLT_learn	0.263	0.511	0.486	NA
MMSE	0.323	0.453	0.421	1
FAQ	0.182	0.343	0.326	$<$ **2.2e-16**

Table 1: P-values calculated from the ADNI data ($m \sim \{2-11\}$ and $\bar{m} \approx 5.5$), the bold p-values are less than 0.1. RAVLT_imm and RAVLT_learn cannot be testes with the direct method because we cannot fit the alternative model.

and RAVLT_learn cannot be carried out because we cannot fit the alternative model.
Based on these test results, we suggest fitting ADAS13 and FAQ with the alternative parametric model is enough. We also recommend fitting RAVLT_imm with nonparametric methods like FACE. And there is not much evidence to support fitting RAVLT_learn and MMSE with nonparametric method or the alternative parametric model will provide better result.
we summarized the limitation of our analysis here. First, we did not consider many alternative parametric models, so there might exist some better models. Second, many subjects drop out of the study but our nonparametric method did not consider the issues.

References

Chen, S. T., Xiao, L. and Staicu, A.-M. (2019), 'A smoothing-based goodness-of-fit test of covariance for functional data', *Biometrics* **75**(2), 562–571.

Crainiceanu, C. M. and Ruppert, D. (2004), 'Likelihood ratio tests in linear mixed models with one variance component', *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **66**(1), 165–185.