# Bootstrap aggregated sparse FPCA for classification

Hyunsung Kim

January 20, 2020

Department of Statistics
Chung-Ang University

# Simulations with 2 datasets

- The growth data
  - 93 observations(60 training set, 33 test set)
  - Dense data $\Rightarrow$ randomly sparsify with 2~6 observations
- The bone mineral density
  - 160 observations(100 training set and 60 test set)
    - On the robust bagging, 60 training set, 40 validation set and 60 test set
  - Sparse data

# Simulation 1

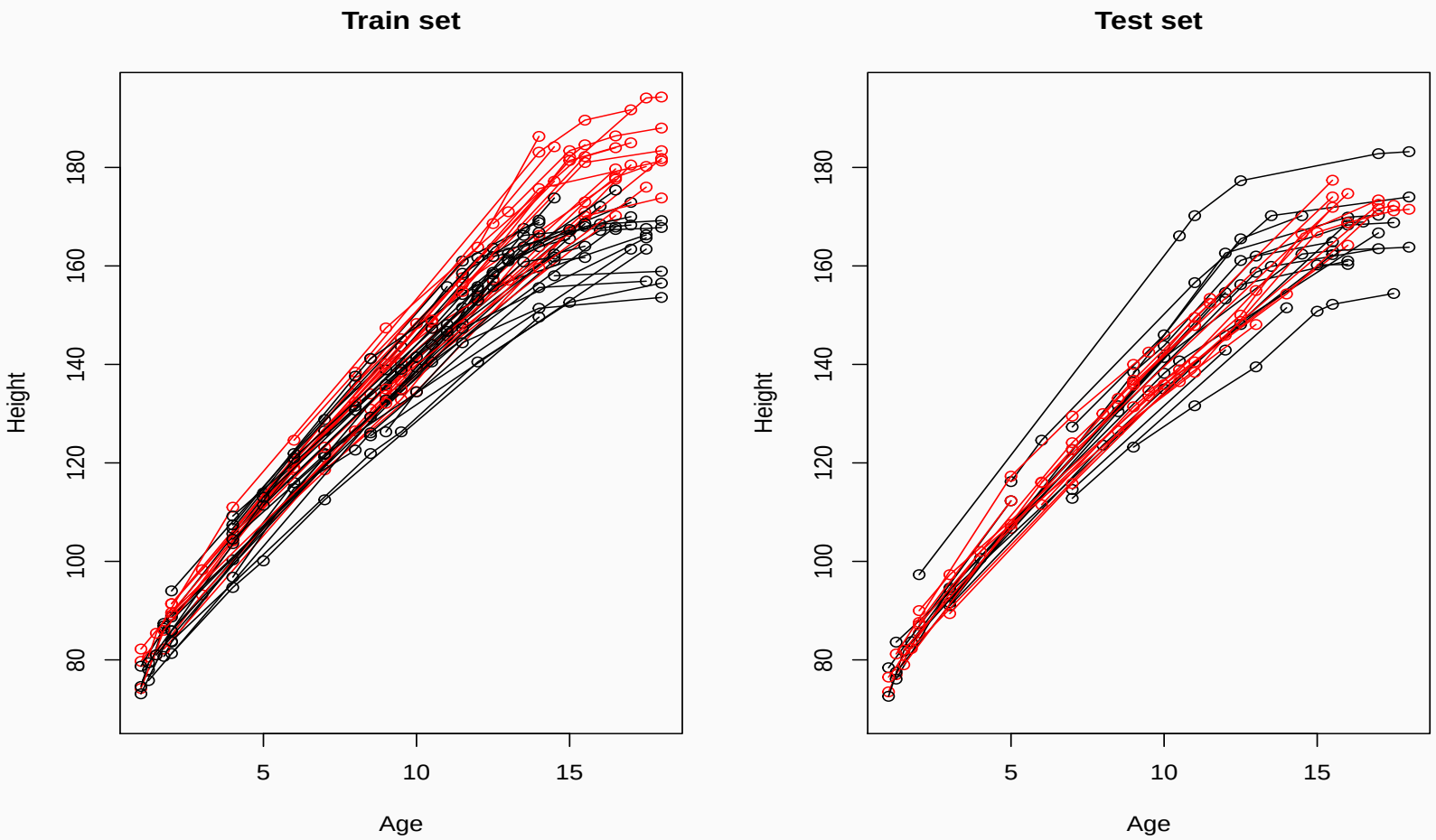


**Figure 1:** The growth data

# Simulation results

**Table 1:** The accuracy of classifiers

| Method | Logistic Regression | SVM (Linear) | SVM (Gaussian) | SVM (Sigmoid) | KNN | LDA | QDA | Naive Bayes |
|---|---|---|---|---|---|---|---|---|
| Single | 0.758 | 0.727 | 0.727 | 0.727 | 0.606 | 0.727 | 0.697 | 0.727 |
| Majority | 0.818 | 0.788 | 0.758 | 0.818 | 0.576 | 0.818 | 0.758 | 0.758 |

# Simulation 2

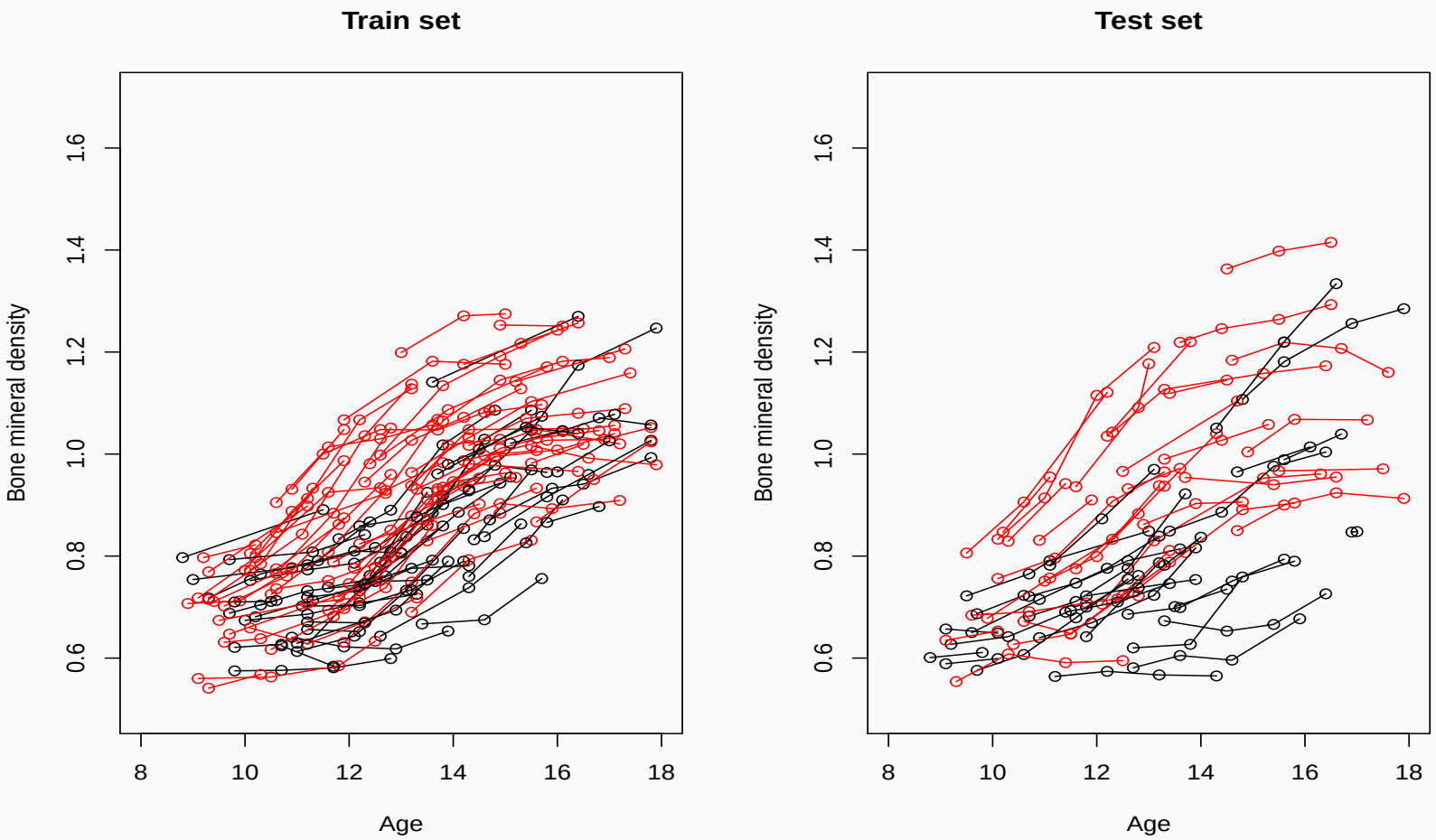


**Figure 2:** The bone mineral density data

# Simulation results

**Table 2:** The accuracy of classifiers with each tuned hyperparameters

| Method | Logistic Regression | SVM (Linear) | SVM (Gaussian) | SVM (Sigmoid) | KNN | LDA | QDA | Naive Bayes |
|---|---|---|---|---|---|---|---|---|
| Single | 0.767 | 0.717 | 0.717 | 0.733 | 1.00 | 0.767 | 0.750 | 0.767 |
| Majority | 0.750 | 0.683 | 0.567 | 0.717 | 1.00 | 0.750 | 0.717 | 0.717 |
| OLS | 0.717 | 0.700 | 0.533 | 0.583 | 0.00 | 0.233 | 0.417 | 0.383 |
| Logit | 0.517 | 0.750 | 0.567 | 0.683 | 1.00 | 0.433 | 0.733 | 0.667 |

# Simulation results

**Table 3:** The accuracy of classifiers with overall tuned hyperparameters

| Method | Logistic Regression | SVM (Linear) | SVM (Gaussian) | SVM (Sigmoid) | KNN | LDA | QDA | Naive Bayes |
|---|---|---|---|---|---|---|---|---|
| Single | 0.767 | 0.717 | 0.717 | 0.733 | 1.00 | 0.767 | 0.750 | 0.767 |
| Majority | 0.733 | 0.700 | 0.700 | 0.683 | 1.00 | 0.717 | 0.733 | 0.733 |
| OLS | 0.650 | 0.717 | 0.617 | 0.700 | 1.00 | 0.650 | 0.567 | 0.533 |
| Logit | 0.617 | 0.367 | 0.700 | 0.600 | 1.00 | 0.600 | 0.583 | 0.500 |
| Ridge-logit | 0.667 | 0.650 | 0.667 | 0.667 | 1.00 | 0.667 | 0.733 | 0.700 |
| OOB error | 0.733 | 0.700 | 0.700 | 0.683 | 1.00 | 0.717 | 0.733 | 0.733 |

# Simulation results

**Table 4:** The accuracy of classifiers with overall tuned hyperparameters and random subspace method

| Method | Logistic Regression | SVM (Linear) | SVM (Gaussian) | SVM (Sigmoid) | KNN | LDA | QDA | Naive Bayes |
|---|---|---|---|---|---|---|---|---|
| Single | 0.767 | 0.717 | 0.717 | 0.733 | 1.00 | 0.767 | 0.750 | 0.767 |
| Majority | 0.717 | 0.650 | 0.650 | 0.650 | 1.00 | 0.700 | 0.750 | 0.717 |
| OOB error | 0.717 | 0.667 | 0.683 | 0.683 | 1.00 | 0.717 | 0.717 | 0.750 |

# Simulation results

**Table 5:** The accuracy of classifiers with overall tuned hyperparameters and BIC

| Method | Logistic Regression | SVM (Linear) | SVM (Gaussian) | SVM (Sigmoid) | KNN | LDA | QDA | Naive Bayes |
|---|---|---|---|---|---|---|---|---|
| Single | 0.767 | 0.717 | 0.717 | 0.733 | 1.00 | 0.767 | 0.750 | 0.767 |
| Majority | 0.717 | 0.667 | 0.700 | 0.700 | 1.00 | 0.717 | 0.717 | 0.767 |
| OOB error | 0.733 | 0.667 | 0.667 | 0.683 | 1.00 | 0.733 | 0.717 | 0.717 |

# Simulation results of robust bagging

**Table 6:** The accuracy of classifiers with each tuned hyperparameters

| Method | Logistic Regression | SVM (Linear) | SVM (Gaussian) | SVM (Sigmoid) | KNN | LDA | QDA | Naive Bayes |
|---|---|---|---|---|---|---|---|---|
| Single | 0.733 | 0.767 | 0.700 | 0.567 | 1.00 | 0.783 | 0.667 | 0.750 |
| Majority | 0.800 | 0.767 | 0.650 | 0.767 | 1.00 | 0.783 | 0.733 | 0.750 |
| OOB error | 0.800 | 0.767 | 0.650 | 0.767 | 1.00 | 0.783 | 0.733 | 0.750 |
| Robust | 0.733 | 0.733 | NA | 0.750 | NA | 0.733 | 0.733 | 0.717 |

# Simulation results of robust bagging

**Table 7:** The accuracy of classifiers with overall tuned hyperparameters

| Method | Logistic Regression | SVM (Linear) | SVM (Gaussian) | SVM (Sigmoid) | KNN | LDA | QDA | Naive Bayes |
|---|---|---|---|---|---|---|---|---|
| Single | 0.733 | 0.767 | 0.700 | 0.567 | 1.00 | 0.783 | 0.667 | 0.750 |
| Majority | 0.817 | 0.767 | 0.733 | 0.767 | 1.00 | 0.783 | 0.733 | 0.767 |
| OOB error | 0.817 | 0.750 | 0.733 | 0.800 | 1.00 | 0.817 | 0.750 | 0.750 |
| Robust | 0.750 | 0.750 | 0.733 | 0.783 | NA | 0.750 | 0.783 | 0.733 |