Classifying time series data via frequency decomposition and manifold techniques

Adam Guo, supervised by Prof. Hrushikesh Mhaskar

Pomona College ('22), Claremont Graduate University

Objectives

Classification of ime series using:

- dynamic models to represent time series data as points on a Grassmann manifold
- kernel methods on the manifold to perform classification while taking advantage of the full geometric information

Introduction

- Time series classification is a problem with many applications in signal processing and machine learning
  - Identifying sources of sound or radar signals, detecting algorithmically-generated "deepfake" video and audio, and many other applications
- We investigate a method for feature extraction: representing signals as points on a Grassmann manifold via dynamic model parametrisation

Dynamic model parametrisation

The autoregressive-moving-average (ARMA) is a well-known dynamic model for time series data that parametrises a signal $f(t)$ by the equations

$$f(t) = Cz(t) + w(t), \quad w(t) \sim \mathcal{N}(0, R) \tag{1}$$

$$z(t+1) = Az(t) + v(t), \quad v(t) \sim \mathcal{N}(0, Q) \tag{2}$$

where $z \in \mathbb{R}^d$ is the hidden state vector, $f : \mathbb{R} \to \mathbb{R}^p$, and $d \leq p$ is the hidden state dimension [1]. There are widely-used closed form solutions for estimating the parameters A and C. It can be shown [1] that the expected observation sequence is given by

$$O_\infty = \mathbb{E}\begin{bmatrix} f(0) \\ f(1) \\ f(2) \\ \vdots \end{bmatrix} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \end{bmatrix} z(0) \tag{3}$$

Representation on Grassmann manifold

- (3) shows that the expected observations of $f(t)$ lie in the column space of the observability matrix O_∞
- Approximate O_∞ by truncating at the mth block, to form $O_m \in \mathsf{M}_{mp \times d}$
- Hence, ARMA model yields a representation of a signal as a Euclidean subspace (given by the column space of O_m), and thus a point on the Grassmann manifold $\mathbf{Gr}(d, mp)$

Kernel methods on Grassmann manifold

- Kernel based methods are used widely in manifold learning.
- Jayasumana et al. [2] define the analogue of the Gaussian RBF kernel for the Grassmann manifold:

$$d_P([Y_1], [Y_2]) = 2^{-1/2} \| Y_1 Y_1^T - Y_2 Y_2^T \|_F \tag{4}$$

$$k_P([Y_1, Y_2]) = \exp(-\gamma d_P^2([Y_1], [Y_2])) \tag{5}$$

where $[Y_i]$ is the subspace spanned by the columns of Y_i, Y_1 and Y_2 are matrices with orthonormal columns, and γ is a hyperparameter

- We use the representation of a time series as a point on the Grassman manifold to do classification based on Support Vector Machines (SVMs) using the above kernel

Algorithm

Input: list of train signals $\{X_i\}_{i=1}^n$, list of train labels $\{y_i\}_{i=1}^n$, list of test signals $\{Y_i\}_{i=1}^m$
Output: list of predicted test labels

- For $i = 1, 2, \ldots, n$:
  - Compute parameters C and A for X_i
  - Compute O_m for X_i
  - Orthonormalise O_m and store as U_i
- Using kernel k_P, fit SVM on $\{U_i\}_{i=1}^n$, $\{y_i\}_{i=1}^n$
- Predict SVM on $\{Y_i\}_{i=1}^m$, return predicted labels

Experiments

Experiments were performed on four sets of data. For each dataset, we performed the supervised learning algorithm detailed above using the raw data as input, and compared its classification accuracy to a simple SVM that flattens the input data and uses a Gaussian kernel in Euclidean space. We also include the best results found in the relevant literature for reference. Note that these results extensive preprocessing on the data before classification, while our results are based on raw data.

SUNY EEG Database [3]:

- EEG tests of alcoholic and non-alcoholic subjects ($p = 64$)
- Train/test split is predefined (at ~48% test data), one trial is performed
- Parameters used: $d = m = 10, \gamma = 0.2$

Vehicle audio recordings [4]:

- Audio recordings of different vehicles moving through a parking lot at around 15 mph ($p = 2$)
- 50% of data used for testing, 20 trials performed
- Last 6 seconds of each recording is used (only the part where the car is near the microphone)
- Parameters used: $d = 2, m = 10, \gamma = 10$

Lip videos [5]:

- Video recordings of a person speaking the digits 1-5 ($p = 3850$)
- 50% of data used for testing, 20 trials performed
- Parameters used: $d = m = 10, \gamma = 0.2$
- Since the videos are not equally long, for Euclidean SVM we perform principal component analysis across the frames to extract 30 principal vectors

Results

Datasets	Grass.	Eucl.	Literature
Alcohol EEG	**99.8**%	80.8%	97.1% [6]
Vehicle audio	62.8%	51.8%	**88.2**% [4]
Video digits	**97.0**%	76.6%	94.7% [5]

Table 1: Classification accuracies of different algorithms per dataset (best performer in bold)

Conclusions

- Grassmann SVM performs significantly better (11-21% more) than straightforward Euclidean SVM, demonstrating effectiveness of parametrising signal in the Grassmann manifold
- For alcohol EEG and video digit datasets, Grassmann SVM approach on raw data performs better than literature results that use extensive preprocessing (orthogonal wavelet filter bank in [6], PCA using earth mover's distance in [5])

Further work

- Implementing preprocessing techniques used with success in literature may improve performance, particuarly for vehicle audio dataset
- Fourier or wavelet transforms, instantaneous frequency decompositions [7]

References

[1] P. Turaga, A. Veeraraghavan, A. Srivastava, and R. Chellappa. Statistical Computations on Grassmann and Stiefel Manifolds for Image and Video-Based Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(11):2273–2286, November 2011.

[2] Sadeep Jayasumana, Richard Hartley, Mathieu Salzmann, Hongdong Li, and Mehrtash Harandi. Kernel Methods on Riemannian Manifolds with Gaussian RBF Kernels. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(12):2464–2477, December 2015. arXiv: 1412.0265.

[3] Xiao Lei Zhang, Henri Begleiter, Bernice Porjesz, Wenyu Wang, and Ann Litke. Event related potentials during object recognition tasks. *Brain Research Bulletin*, 38(6):531–538, January 1995.

[4] Justin Sunu and Allon G. Percus. Dimensionality reduction for acoustic vehicle classification with spectral embedding. *arXiv:1705.09869 [physics, stat]*, February 2018. arXiv: 1705.09869.

[5] Linh Lieu and Naoki Saito. Signal Ensemble Classification Using Low-Dimensional Embeddings and Earth Mover's Distance. In Jonathan Cohen and Ahmed I. Zayed, editors, *Wavelets and Multiscale Analysis: Theory and Applications*, Applied and Numerical Harmonic Analysis, pages 227–256. Birkhäuser, Boston, 2011.

[6] Manish Sharma, Dipankar Deb, and U. Rajendra Acharya. A novel three-band orthogonal wavelet filter bank method for an automated identification of alcoholic EEG signals. *Applied Intelligence*, August 2017.

[7] Charles K. Chui and H.N. Mhaskar. Signal decomposition and analysis via extraction of frequencies. *Applied and Computational Harmonic Analysis*, 40(1):97–136, January 2016.