MULTIVARIATE ANALYSIS AND KERNEL METHODS FOR MUSIC DATA ANALYSIS

Jerónimo Arenas-García, Anders Meng, Kaare Brandt Petersen and Lars Kai Hansen

Informatics and Mathematical Modelling, Technical University of Denmark

OVERVIEW

- Increasing interest in customizable methods for organizing large digital music collections
- Multivariate Analysis (MVA) methods enjoy some desirable characteristics, such as flexibility and scalability.
- We present two modified versions of Orthonormalized PLS (OPLS), one of them being a kernel extension, that are well-suited for discovering relevant dynamics in large music collections

1 Orthonormalized PLS (OPLS)

- Training data: $\mathbf{X} = [(\mathbf{x}_1), \ldots, (\mathbf{x}_l)]^T$; $\mathbf{Y} = [\mathbf{y}_1, \ldots, \mathbf{y}_l]^T$
- Find projection vectors for feature extraction:

$$\tilde{\mathbf{X}}' = \tilde{\mathbf{X}}\mathbf{U}$$

where each column in $\mathbf{U}$ $(dim(\mathbf{x}) \times n_p)$ is a projection vector, and $\tilde{\mathbf{X}}$ is a centered version of $\mathbf{X}$

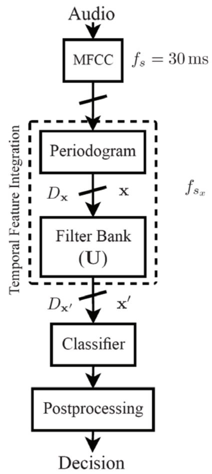
- OPLS:

$$\boxed{\begin{array}{ll} \text{maximize:} & \mathrm{Tr}\{\mathbf{U}^T\tilde{\mathbf{X}}^T\tilde{\mathbf{Y}}\tilde{\mathbf{Y}}^T\tilde{\mathbf{X}}\mathbf{U}\} \\ \text{subject to:} & \mathbf{U}^T\tilde{\mathbf{X}}^T\tilde{\mathbf{X}}\mathbf{U} = \mathbf{I} \end{array}}$$

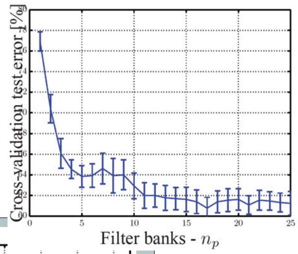
- OPLS properties:
  - Only projections for input data; projected data is white
  - Optimal features for linear prediction (with a bottleneck) of the training labels with square loss

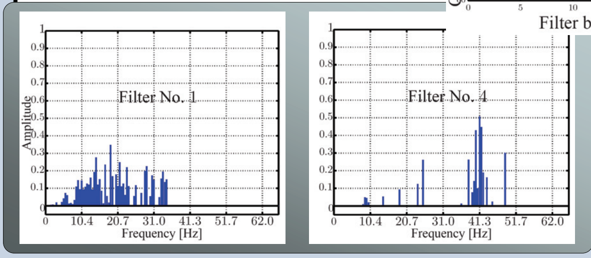
- **Flexibility**: The output space $(\mathbf{Y})$ can encode any kind of information, like genre, user ratings …, including also the multi-label case
- **Scalability**: MVA extract relevant projections using the covariance matrices only, which can be computed as a sum over all patterns. Therefore, MVA can be used with large datasets, and for incremental learning

2 Positive constrained OPLS



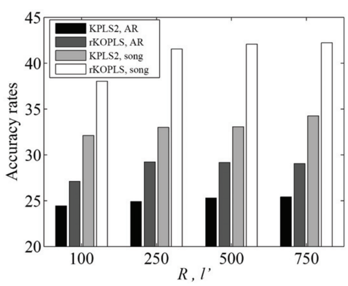
- We use a modified version of OPLS for discovering relevant dynamics in music
- The periodograms of MFCCs are computed; these spectra contain information about how fast MFCCs change
- Periodograms are filtered to calculate the energy in different frequency bands
- Instead of using a fixed filter bank, $\mathbf{U}$ is supervisedly designed using OPLS, with the additional constraint $u_{ij} \geq 0$, so that $\mathbf{U}$ can be interpreted as the frequency response of a filter bank
- The features have an straightforward interpretation, e.g., energy in the modulation frequencies of instruments, etc …
- … While keeping a good discriminative power

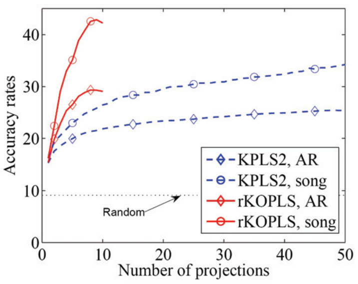




3 Feature Extraction with rKOPLS

- Kernel extension of the OPLS algorithm to increase the discrimination power of the extracted features
- To reduce the complexity of the algorithm (both in the training and test phases), sparsity is imposed in the solution
- rKOPLS can be used with large data sets, e.g., in music analysis applications with large collections

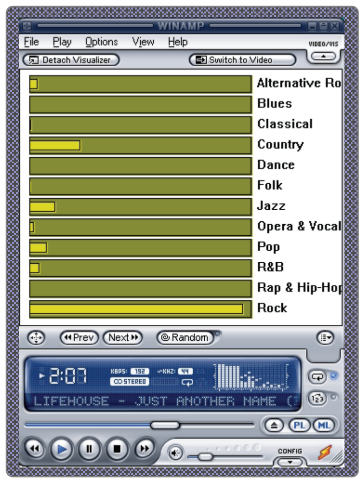




Experiments settings

- Input features: AR models of MFCC coefficients (1.2 sg)
  Each AR model is summarized in a 135 length vector
- Training data: 57,388 AR vectors, approx. evenly distributed
  Test data: 36,556 AR vectors corresponding to 500 songs
- Most kernel MVA methods cannot handle such a training set
- Gaussian kernel; width parameter selected with CV
- Classifier: SLP + softmax network

- rKOPLS significantly outperforms KPLS2 with only 10 proj.
- Accuracy does not increase significantly for $R > 500$
- This system is running on-line inside a plug-in for winamp



rKOPLS provides expressive power + controlled CPU cost