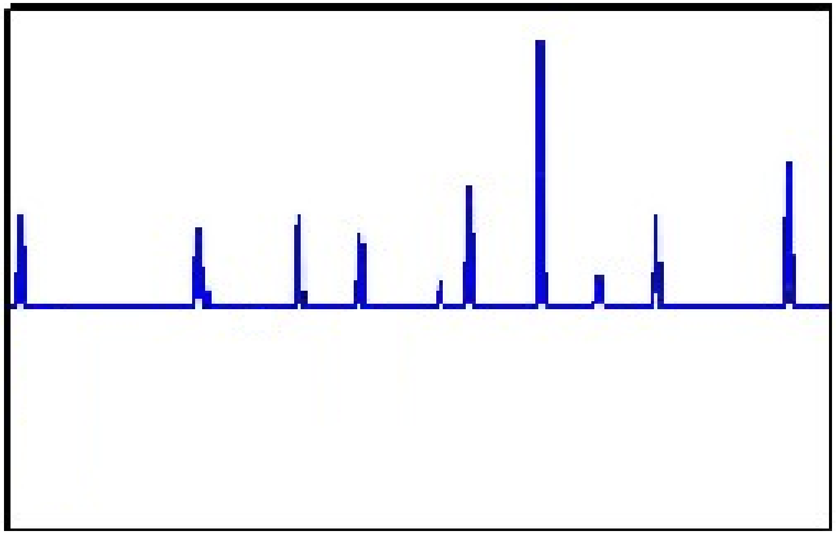
Novel Variations of Sparse Representation Techniques with Applications

Qun Feng Tan

Advisor: Prof. Shrikanth Narayanan

Contributions



Sparse Representation involves:
Estimation of a sparse multi-dimensional vector, satisfying a linear system of equations given high-dimensional observed data and a design matrix (dictionary).

Novel Contributions of this work:

- New Variations of Group Regularization Techniques, applied to denoising in Automatic Speech Recognition systems
- New methods for combining predictions from different windows

Novel Variations of Group Regularization Algorithms

$$\min ||y - Ax||_2^2 + \lambda_0 ||x||_0$$

We propose solutions to deal with the case where we have dictionary partitions:

$$A = [A_1 | A_2 | \cdots | A_n]$$

We also additionally take into account the situation where we have a highly collinear dictionary:

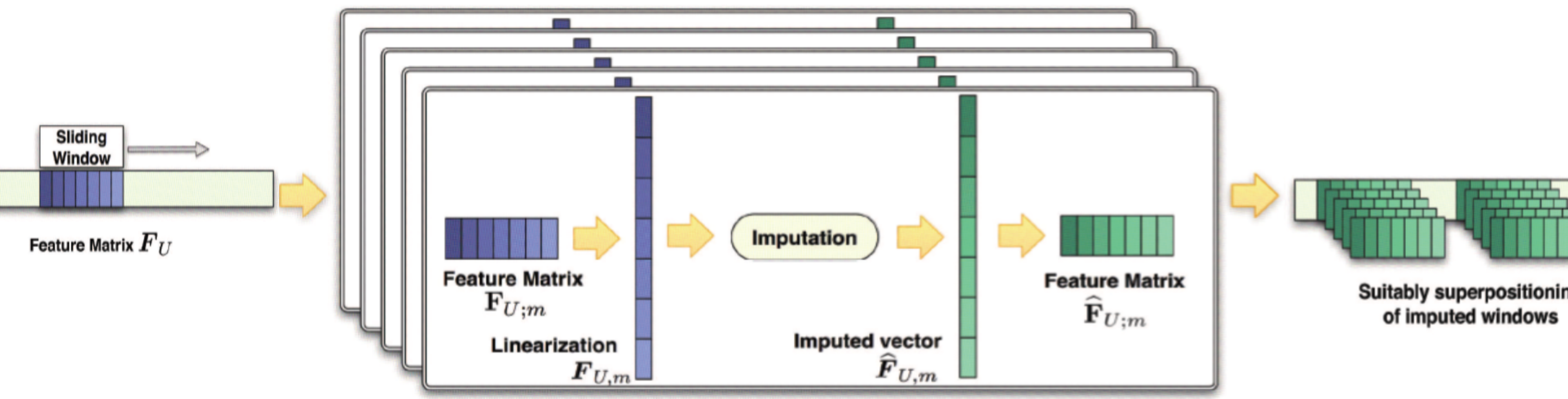
- **Group Elastic Net Algorithm**: Extension of the Elastic Net to the multi-group case by an iterative approach.

$$\min ||y - [A_1 | A_2 | \cdots | A_n] x||_2^2 + \lambda_1 ||x||_1 + \lambda_2 ||x||_2^2$$

Novel Variations of Group Regularization Algorithms

- **Group Sparse Bayesian Learning Algorithm:** Extension of the Sparse Bayesian Learning (SBL) Algorithm to both grouping + collinear dictionaries (*Algorithm too involved, left out of presentation*)

Application to Denoising in Automatic Speech Recognition



Novel Variations of Group Regularization Algorithms

RESULTS FOR AURORA 3.0 NOISY DATASET

Algorithm	Accuracy (%)
Before denoising	62.87
LARS-EN	75.26
Group EN	77.50
Group SBL	76.46
ETSI AFE	77.21
CMN	67.09

Qun Feng Tan, Shrikanth S. Narayanan, “Novel Variations of Group Sparse Regularization Techniques with Applications to Noise Robust Automatic Speech Recognition” – IEEE Transactions on Audio, Speech, and Language Processing, vol. 20, no. 4, pp. 1337 –1346, May 2012

Novel Variations of Sparse Representation Algorithms

Better recombination of predictions from windowing scheme

- Typical strategy is to do averaging of the predictions
- What we propose is an iterative scheme to combine the predictions in a more effective way (experimentally justified)

$$Y_s = Y R_i$$

$$Y - \text{Original Feature Matrix}$$
$$Y_s - \text{Extracted Window}$$
$$R_i - \text{Window Extraction Matrix}$$

Novel Variations of Sparse Representation Algorithms

Proposed Algorithm:

1) For each window, regularize:

$$\min ||y_s - Ax||_2^2 + \lambda_1 ||x||_1 + \lambda_2 ||x||_2^2$$

2) Combine predictions by doing one more optimization:

$$\min_{\widehat{Y}} ||Y - \widehat{Y}||_2^2 + \sum_{i=1}^{N_w} ||Y_{N_w} - Y R_i||_2^2$$

Novel Variations of Sparse Representation Algorithms

Algorithm	Accuracies (%)	Runtimes	Significant?
SNR 0 dB			
Unimputed	9.63	NA	NA
CMN	27.78	NA	NA
EN Averaging	26.64	0.0158	NA
EN Coupled	40.39	0.0197	Yes
SNR 5 dB			
Unimputed	36.78	NA	NA
CMN	55.91	NA	NA
EN Averaging	64.38	0.0209	NA
EN Coupled	72.57	0.0215	Yes
SNR 10 dB			
Unimputed	61.41	NA	NA
CMN	87.82	NA	NA
EN Averaging	83.85	0.0296	NA
EN Coupled	89.99	0.0364	Yes
SNR 15 dB			
Unimputed	81.99	NA	NA
CMN	95.67	NA	NA
EN Averaging	93.25	0.0409	NA
EN Coupled	95.84	0.0356	Yes