Motivation

• Observing latent structure in a signal of interest allows for data analysis and interpretation.
• NMF is a widely-used method for observing latent structure, but is sensitive to noise.
• Noise in data prevents proper recovery of latent structure.
• **Goal:** Improve recovery of latent structure by jointly filtering the data to remove noise and factoring it to expose its latent structure.

Joint Filtering and Factorization

• Non-negative matrix factorization (NMF) provides a parts-based representation of a signal.
• Minimum variance distortionless response (MVDR) filter computes a power spectrum that estimates the spectral envelope of a signal without distorting the spectrum.
• Combine the cost functions of NMF and MVDR to perform joint filtering and factorization.

\[ J = \| G \otimes V - WH \|_F^2 + \lambda_1 \| G \otimes (WH) \|_F^2 + \lambda_2 \| G \otimes (WH) - A \otimes (WH) \|_F^2 \]

• Update equations:
  \[ G = \frac{W^T W + \lambda_1 A \otimes (WH) \otimes (W^T)^2}{\| W^T W \|_F^2 + \lambda_1 \| A \otimes (WH) \|_F^2} \]
  \[ W = W \otimes W^T + \lambda_1 C \otimes D + \lambda_2 \otimes F \otimes H \]
  \[ H = H \otimes G \]

• Filtering operation can be viewed as computing the optimal filter \( g_m[n] \) for each frame in the STFT of the noisy signal.

\[ X(m, \omega) = \sum_{n=-\infty}^{\infty} x[n] w[n+m] e^{-j\omega n} \]
\[ Y(m, \omega) = \sum_{n=-\infty}^{\infty} \sum_{\omega} w[n+m] e^{-j\omega n} \]
\[ = X(m, \omega) G(m, \omega) \]

Basis Recovery Experiment

• Correlation metric:
  \[ \rho = \frac{1}{K} \sum_{k=1}^{K} \frac{\| W_{\text{clean}}(k) W_{\text{noisy}}(k) \|_2}{\| W_{\text{clean}}(k) \|_2 \| W_{\text{noisy}}(k) \|_2} \]

• Added white, pink, speech babble, and factory noises to 100 TIMIT sentences at 5 dB and 10 dB SNR.

Conclusion & Future Work

• Proposed joint filtering and factorization approach by combining NMF and MVDR into a single cost function.
• Correlation scores show better basis recovery of noisy signals.
• PESQ scores show denoising performance is comparable to other denoising methods.
• Use more generalized divergence metrics in cost function.
• Extend approach to other analysis methods, eg. PCA.
• Evaluate proposed algorithm’s performance on other tasks, eg. phoneme recognition.

Work supported by NIH Grant DC007124