Journal Watch:

Transanctions on Signal Processing, September 15, 2016

Saurabh Khanna,
Signal Processing for Communication, ECE, IISc

Super-Resolution Compressed Sensing for Line Spectral Est.: An Iterative Reweighted Approach

Authors: Jun Fang, Feiyu Wang, Yanning Shen, Hongbin Li, and Rick S. Blum

- Line spectral estimation problem in compressive sensing framework.
- Generative model for signals with line spectrum:

$$y_m = \sum_{k=1}^{K} \alpha_k e^{-j\omega_k m} + e_m \qquad m = 1, \dots, M$$

  where $\omega_k \in [0, 2\pi]$ and α_k is complex amplitude.
- Vectorized form:

$$\mathbf{y} = \mathbf{A}(\omega)\alpha + \mathbf{e}$$

  where $\mathbf{A}(\omega) \triangleq [\mathbf{a}(\omega_1)\mathbf{a}(\omega_2)\dots\mathbf{a}(\omega_K)]$]), a matrix with steering vectors as columns.
- How to handle the case when the frequencies ω_i belong to a continuous set.
- Typical solution: Discretize the search space using a finite sized grid and construct a finite sized dictionary.
  - Erroneous estimates of present frequencies due to grid mismatch
- **Question**
  How to estimate continuous frequencies accurately while still using a finite sized dictionary ?

Super-Resolution Compressed Sensing for Line Spectral Est.: An Iterative Reweighted Approach

- Proposed solution: Joint estimation of dictionary and sparse coefficient vector α.

$$\min_{\mathbf{z},\theta} G(\mathbf{z},\theta) = \sum_{n=1}^{N} \log\left(|z_n|^2 + \epsilon\right) + \lambda\,||\mathbf{y} - \mathbf{A}(\theta)\mathbf{z}||_2^2$$

- Majorization-minimization approach used to minimize $G(\mathbf{z},\theta)$.
- Log-sum penalty replaced by its convex upper bound obtained by its Taylor's expansion around previous iterate.
- In each iteration, for fixed θ, $\mathbf{z}$ is updated as a solution to a weighted least squares problem.
- Dictionary parameters θ are updated by moving along the gradient descent direction.
- Cost $G(\mathbf{x},\theta)$ is shown to reduce with each iteration.
- Connection between SBL and Iter. Reweight. approach is exploited to update λ as a function of noise variance estimate in each iteration.
- As $\epsilon \to 0$, algorithm is shown to converge to the correct frequencies.

Efficient Algorithms on Robust Low Rank Matrix Completion Against Outliers

Authors: Licheng Zhao, Prabhu Babu and Daniel P. Palomer

- Robust Matrix completion in presence of outliers.
- Data model:
$$\tilde{\mathbf{M}} = \Omega \circ (\mathbf{M} + \mathbf{N})$$
  where $\mathbf{N}$ is the noise matrix and Ω is the sparse binary sampling matrix.
- $\mathbf{M}$ is the low rank matrix to be estimated.
- $\mathbf{N}$ models two kinds of outliers:
  1. Dense outliers from certain elliptical distributions.
  2. Sparse spike like outliers with small additive Gaussian noise.
- Low rankness of $\mathbf{M}$ is promoted by imposing a bilinear factorization form.
$$\underbrace{\mathbf{M}}_{m \times n} = \underbrace{\mathbf{X}^T}_{m \times r} \underbrace{\mathbf{Y}}_{r \times n} \qquad \text{where } r < \min(m, n)$$
- Matrix Factorization Formulation:
$$\min_{\mathbf{X},\mathbf{Y}} J(\mathbf{X}, \mathbf{Y}) \triangleq \sum_{i=1}^{m} \sum_{j=1}^{n} \Omega_{i,j} f\left(\tilde{\mathbf{M}}_{i,j} - \mathbf{x}_i^T \mathbf{y}_j\right) + \gamma\left(||\mathbf{X}||_F^2 + ||\mathbf{Y}||_F^2\right)$$

Efficient Algorithms on Robust Low Rank Matrix Completion Against Outliers

- Matrix Factorization Formulation:

$$\min_{\mathbf{X},\mathbf{Y}} J(\mathbf{X},\mathbf{Y}) \triangleq \sum_{i=1}^{m}\sum_{j=1}^{n} \Omega_{i,j} f\left(\tilde{\mathbf{M}}_{i,j} - \mathbf{x}_i^T \mathbf{y}_j\right) + \gamma\left(||\mathbf{X}||_F^2 + ||\mathbf{Y}||_F^2\right)$$

- For dense outliers from certain elliptical distributions.

$$f(x) = \log\left(1 + \frac{x^2}{\nu}\right) \qquad \nu > 0$$

- For sparse spike like outliers with small additive Gaussian noise.

$$f(x) = \frac{1}{\beta}\log\left(\frac{e^{\beta x} + e^{-\beta x}}{2}\right) \qquad \beta > 0$$

- Parallel minimization algorithm is proposed.
  - Do a second order convex approximation of $J(\mathbf{X},\mathbf{Y})$.
  - Compute gradient descent direction which reduces the convex approximation.
  - Execute optimal stepsize based update of $\mathbf{X}$ and $\mathbf{Y}$ such that $J(\mathbf{X},\mathbf{Y})$ decreases in every iteration.

Robust Hypothesis Testing with α-Divergence

Authors: Gokhan Gul and Abdelhak M. Zoubir (TU Darmstadt, Germany).

- Robust Hypothesis Testing problem is considered in which probability distribution F_0 and F_1 of observations under $\mathcal{H}_0$ and $\mathcal{H}_1$ hypothesis is partially known.
- Due to imprecise knowledge of F_0, F_1, one possible approach is to extend the model by accepting a set of distributions $\mathcal{G}_j$ under hypothesis $\mathcal{H}_j$, where

$$\mathcal{G}_j = \{G : D_\alpha(G, F_j) \le \epsilon_j\} \qquad j \in \{0,1\}.$$

- Distance metric D_α is chosen to be α-divergence.
- Tradeoff between robustness and detection performance via choice of ϵ_0 and ϵ_1.
- Modified binary (composite) hypothesis testing problem:

$$\begin{aligned} \mathcal{H}_0 : &\quad G \in \mathcal{G}_0 \\ \mathcal{H}_1 : &\quad G \in \mathcal{G}_1 \end{aligned}$$

- Given observations $\mathbf{y} \in \Omega$, a detector output is given by $\delta : \Omega \to [0,1]$.

$$\begin{aligned} P_{FA}(\delta, F_0) &= \int_\Omega \delta(\mathbf{y}) F_0(\mathbf{y}) d\mathbf{y} \\ P_{MD}(\delta, F_1) &= \int_\Omega (1-\delta(\mathbf{y})) F_1(\mathbf{y}) d\mathbf{y} \\ P_E(\delta, F_0, F_1) &= P(\mathcal{H}_0) P_{FA} + P(\mathcal{H}_0) P_{MD} \end{aligned}$$

Robust Hypothesis Testing with α-Divergence

- Saddle Point Specification (Game Theoretic Approach)
- By Sion's minimax theorem

$$\sup_{(G_0,G_1)\in\mathcal{G}_0\times\mathcal{G}_1} \min_{\delta} P_E(\delta, G_0, G_1) = \min_{\delta} \sup_{(G_0,G_1)\in\mathcal{G}_0\times\mathcal{G}_1} P_E(\delta, G_0, G_1)$$

- Based on Sion's minimax theorem, there exists a saddle value for the objective P_E, i.e.,

$$P_E(\delta, \hat{G}_0, \hat{G}_1) \geq P_E(\hat{\delta}, \hat{G}_0, \hat{G}_1) \geq P_E(\hat{\delta}, G_0, G_1)$$

- $(\hat{\delta}, \hat{G}_0, \hat{G}_1)$ are found by solving below constrained optimizations simultaneously.

$$\hat{G}_0 = \arg\sup_{G_0\in\mathcal{G}_0} P_{FA}(\delta, G_0)$$

$$\hat{G}_0 = \arg\sup_{G_0\in\mathcal{G}_0} P_{FA}(\delta, G_0)$$

$$\hat{\delta} = \arg\min_{\delta} P_E(\delta, \hat{G}_0, \hat{G}_1)$$

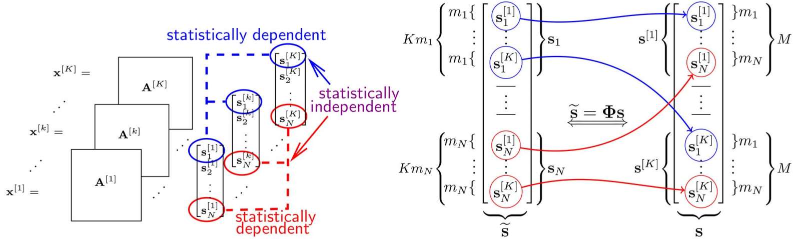
- Optimal closed form expressions derived via KKT for $(\hat{\delta}, \hat{G}_0, \hat{G}_1)$.

Joint Independent Subspace Analysis Using Second-Order Statistics

Authors: Dana Lahat and Christian Jutten

(GIPSA-Lab, Gernoble, France)

- Joint Independent Subspace Analysis:



- Consider T observations of K vectors $\mathbf{x}^{[k]}(t)$, modeled as

$$\mathbf{x}^{[k]}(t) = \mathbf{A}^{[k]}\mathbf{s}^{[k]}(t) \qquad 1 \leq t \leq T, \;\; 1 \leq k \leq K.$$

- Each $\mathbf{s}^{[k]}(t)$ can be partitioned into N blocks of known sizes. $\mathbf{s}_i^{[k]}(t)$ denotes the i^{th} block/partition.
- The blocks $\mathbf{s}_i^{[k]}(t)$ and $\mathbf{s}_j^{[k]}(t)$ are statistically independent for $i \neq j$.
- The blocks $\mathbf{s}_i^{[k]}(t)$ and $\mathbf{s}_i^{[l]}(t)$ are statistically dependent (modelled by same covariance matrix).

Joint Independent Subspace Analysis Using Second-Order Statistics

- By stacking all data sets as one vector, we get

$$\begin{bmatrix} \mathbf{x}^{[1]} \\ \vdots \\ \mathbf{x}^{[K]} \end{bmatrix} = \begin{bmatrix} \mathbf{A}^{[1]} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{A}^{[K]} \end{bmatrix} \begin{bmatrix} \mathbf{s}^{[1]} \\ \vdots \\ \mathbf{s}^{[K]} \end{bmatrix}$$

$$\mathbf{x}(t) = \mathbf{A}\mathbf{x}(t)$$

- An interesting viewpoint obtained by rearranging terms.

$$\mathbf{x}(t) = \sum_{i=1}^{N} \begin{bmatrix} \mathbf{A}_i^{[1]} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{A}_i^{[K]} \end{bmatrix} \begin{bmatrix} \mathbf{s}_i^{[1]}(t) \\ \vdots \\ \mathbf{s}_i^{[K]}(t) \end{bmatrix}$$
$$= \sum_{i=1}^{N} (\mathbf{I}_K \boxtimes \mathbf{A}_i)\mathbf{s}_i(t) = \sum_{i=1}^{N} \mathbf{x}_i(t),$$

where N is the number of blocks/partitions of $\mathbf{s}^{[k]}(t)$, and $\mathbf{A}_i \triangleq \left[\mathbf{A}_i^{[1]} | \mathbf{A}_i^{[2]} | \ldots | \mathbf{A}_i^{[K]}\right]$, and $\mathbf{x}_i \triangleq \left[\mathbf{x}_i^{[1]\,T} | \mathbf{x}_i^{[2]\,T} | \ldots | \mathbf{x}_i^{[K]\,T}\right]^T$.

Other Interesting Papers:

- An Iterative Reweighted Method for Tucker Decomposition of Incomplete Sensors.
- Steady-State Statistical Performance Analysis of Subspace Tracking Methods.
- An Opportunistic Sensor Scheduling Solution to Remote State Estimation Over Multiple Channels.
- Closed Form and Near Closed Form Solutions for TOA based Joint Source and and Sensor Localization.