Journal Watch:

Recent papers on ArXiV related to Sparse Signal Recovery

September 19, 2015

Saurabh Khanna,
Signal Processing for Communication, ECE, IISc

Prior Support Knowledge-Aided Sparse Bayesian Learning with Partly Erroneous Support Information

Authors: Jun Fang, Yanning Shen, Fuwei Li and Hongbin Li

- **Problem Statement:**
  - Sparse signal recovery with partial support information.
  - Recover sparse $\mathbf{x} \in \mathbb{R}^n$ from $m(< n)$ measurements $\mathbf{y} \in \mathbb{R}^m$, where

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{w}$$

  - Prior support information available in the form of index set $P \subset T$, where T is the true support. P can be slightly erroneous.
- **Proposed algorithm**
  - Modification to SBL to accomodate prior support information P.
  - A three layer hierarchical model is proposed

$$p(\mathbf{x}|\alpha) = \prod_{i=1}^{n} \mathcal{N}(0, \alpha_i^{-1}) \qquad p(\alpha) = \prod_{i=1}^{n} \text{Gamma}(\alpha_i|a, b_i)$$

$$p(b_i) = \begin{Bmatrix} \text{Gamma}(b_i|p, q) & i \in P \\ \delta(b_i - 10^{-4}) & i \in P^c \end{Bmatrix}$$

  - Variational Bayesian Inference used to obtain posterior distribution of $\mathbf{x}$ and other hidden variables in the latent model.

MAP Support Detection for Greedy Sparse Signal Recovery Algorithms in Compressive Sensing

Authors: Namyoon Lee (Univ. of Texas, Austin)

- **Problem Statement:**
  - Recover sparse vector $\mathbf{x}$ from compressive linear measurements $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{w}$
  - MAIN IDEA: Using MAP support detection techniques in greedy algorithms such as MP, OMP, CoSamp, etc. results in improved performance (at the cost of more computations)
- **MAP support detection technique**
  - Conventional way is to pick columns of $\mathbf{A}$ in a greedy manner which are maximally correlated with the residual.
  - Residual: $\mathbf{r} = \sum_{l \in T \setminus S^{k-1}} \mathbf{a}_l x_l$
  - For index i, generate correlation metric $Z_i^k = \frac{\mathbf{a}_n^T \mathbf{r}^{k-1}}{||\mathbf{a}_n||_2}$
  - Choose index with maximum a posteriori ratio for given observation Z_i^k

$$\Lambda(Z_i^k) = \log \frac{\mathbb{P}(i \in T | Z_i^k)}{\mathbb{P}(i \notin T | Z_i^k)}$$

  - Closed form expression for $\Lambda(Z_i^k)$ for different source distributions.

Improving M-SBL for Joint Sparse Recovery using a Subspace Penalty

Authors: Jong Chul Ye, Jong Min Kim and Yoram Bresler, KAIST, Korea

- Recover joint sparse vectors $\mathbf{X}$ from L compressive linear measurements (MMVs) $\mathbf{Y}_{m\times L} = \mathbf{A}_{m\times N}\mathbf{X} + \mathbf{W}$.
- In M-SBL, we do MAP estimation of $\mathbf{X}$ by assuming parameterized prior on $\mathbf{X}$.

$$p(\mathbf{X};\gamma) = \prod_{j=1}^{L} \mathcal{N}(0,\Gamma), \qquad \Gamma = \mathrm{diag}\boldsymbol{\gamma}$$

Improving M-SBL for Joint Sparse Recovery using a Subspace Penalty

Authors: Jong Chul Ye, Jong Min Kim and Yoram Bresler, KAIST, Korea

- Recover joint sparse vectors $\mathbf{X}$ from L compressive linear measurements (MMVs) $\mathbf{Y}_{m\times L} = \mathbf{A}_{m\times N}\mathbf{X} + \mathbf{W}$.
- In M-SBL, we do MAP estimation of $\mathbf{X}$ by assuming parameterized prior on $\mathbf{X}$.

$$p(\mathbf{X};\gamma) = \prod_{j=1}^{L} \mathcal{N}(0,\Gamma), \qquad \Gamma = \mathrm{diag}\boldsymbol{\gamma}$$

- The prior parameters are found via Type-2 ML estimation, by maximizing

$$L(\gamma) = Tr((\sigma^2\mathbf{I} + \mathbf{A}\Gamma\mathbf{A}^T)^{-1}YY^T) + N\log|\sigma^2 + \mathbf{A}\Gamma\mathbf{A}^T|$$

Improving M-SBL for Joint Sparse Recovery using a Subspace Penalty

Authors: Jong Chul Ye, Jong Min Kim and Yoram Bresler, KAIST, Korea

- Recover joint sparse vectors $\mathbf{X}$ from L compressive linear measurements (MMVs) $\mathbf{Y}_{m\times L} = \mathbf{A}_{m\times N}\mathbf{X} + \mathbf{W}$.
- In M-SBL, we do MAP estimation of $\mathbf{X}$ by assuming parameterized prior on $\mathbf{X}$.

$$p(\mathbf{X};\gamma) = \prod_{j=1}^{L} \mathcal{N}(0,\Gamma), \qquad \Gamma = \mathrm{diag}\boldsymbol{\gamma}$$

- The prior parameters are found via Type-2 ML estimation, by maximizing

$$L(\gamma) = Tr((\sigma^2\mathbf{I} + \mathbf{A}\Gamma\mathbf{A}^T)^{-1}YY^T) + N\log|\sigma^2 + \mathbf{A}\Gamma\mathbf{A}^T|$$

- (Wipf, Rao, Natrajan) An alternate interpretation of M-SBL algorithm:

$$\min_{\mathbf{X}} ||\mathbf{Y} - \mathbf{A}\mathbf{X}||_F^2 + \sigma^2 g_{MSBL}(\mathbf{X}) \qquad \text{(Solve for } X\text{)}$$

$$g_{MSBL}(\mathbf{X}) = \min_{\gamma\geq 0} Tr(\mathbf{X}^T\Gamma^{-1}\mathbf{X}) + N\log|\sigma^2\mathbf{I} + \mathbf{A}\Gamma\mathbf{A}^T| \qquad \text{(Solve for } \gamma\text{)} \qquad (1)$$

Improving M-SBL for Joint Sparse Recovery using a Subspace Penalty

Authors: Jong Chul Ye, Jong Min Kim and Yoram Bresler, KAIST, Korea

- Recover joint sparse vectors $\mathbf{X}$ from L compressive linear measurements (MMVs) $\mathbf{Y}_{m\times L} = \mathbf{A}_{m\times N}\mathbf{X} + \mathbf{W}$.
- In M-SBL, we do MAP estimation of $\mathbf{X}$ by assuming parameterized prior on $\mathbf{X}$.

$$p(\mathbf{X};\gamma) = \prod_{j=1}^{L} \mathcal{N}(0,\Gamma), \qquad \Gamma = \mathrm{diag}\boldsymbol{\gamma}$$

- The prior parameters are found via Type-2 ML estimation, by maximizing

$$L(\gamma) = Tr((\sigma^2\mathbf{I} + \mathbf{A}\Gamma\mathbf{A}^T)^{-1}YY^T) + N\log|\sigma^2 + \mathbf{A}\Gamma\mathbf{A}^T|$$

- (Wipf, Rao, Natrajan) An alternate interpretation of M-SBL algorithm:

$$\min_{\mathbf{X}} ||\mathbf{Y} - \mathbf{A}\mathbf{X}||_F^2 + \sigma^2 g_{MSBL}(\mathbf{X}) \qquad \text{(Solve for } X\text{)}$$

$$g_{MSBL}(\mathbf{X}) = \min_{\gamma\geq 0} Tr(\mathbf{X}^T\Gamma^{-1}\mathbf{X}) + N\log|\sigma^2\mathbf{I} + \mathbf{A}\Gamma\mathbf{A}^T| \qquad \text{(Solve for } \gamma\text{)} \qquad (1)$$

- The first term in (1) behaves like $N||\gamma||_0$.
- The non-separability of $\log|\sigma^2\mathbf{I} + \mathbf{A}\Gamma\mathbf{A}^T|$ term is the main reason why MSBL is able to dodge many local minimizers.

Improving M-SBL for Joint Sparse Recovery using a Subspace Penalty

Authors: Jong Chul Ye, Jong Min Kim and Yoram Bresler, KAIST, Korea

- Recover joint sparse vectors $\mathbf{X}$ from L compressive linear measurements (MMVs) $\mathbf{Y}_{m\times L} = \mathbf{A}_{m\times N}\mathbf{X} + \mathbf{W}$.
- In M-SBL, we do MAP estimation of $\mathbf{X}$ by assuming parameterized prior on $\mathbf{X}$.

$$p(\mathbf{X};\gamma) = \prod_{j=1}^{L} \mathcal{N}(0,\Gamma), \qquad \Gamma = \mathrm{diag}\boldsymbol{\gamma}$$

- The prior parameters are found via Type-2 ML estimation, by maximizing

$$L(\gamma) = Tr((\sigma^2\mathbf{I} + \mathbf{A}\Gamma\mathbf{A}^T)^{-1}YY^T) + N\log|\sigma^2 + \mathbf{A}\Gamma\mathbf{A}^T|$$

- (Wipf, Rao, Natrajan) An alternate interpretation of M-SBL algorithm:

$$\min_{\mathbf{X}} ||\mathbf{Y} - \mathbf{AX}||_F^2 + \sigma^2 g_{MSBL}(\mathbf{X}) \qquad \text{(Solve for } X\text{)}$$

$$g_{MSBL}(\mathbf{X}) = \min_{\gamma\geq 0} Tr(\mathbf{X}^T\Gamma^{-1}\mathbf{X}) + N\log|\sigma^2\mathbf{I} + \mathbf{A}\Gamma\mathbf{A}^T| \qquad \text{(Solve for } \gamma\text{)} \qquad (1)$$

- The first term in (1) behaves like $N||\gamma||_0$.
- The non-separability of $\log|\sigma^2\mathbf{I} + \mathbf{A}\Gamma\mathbf{A}^T|$ term is the main reason why MSBL is able to dodge many local minimizers.
- Looking purely from **cost function point of view**, can we do better ?

Improving M-SBL for Joint Sparse Recovery using a Subspace Penalty

- (Wipf, Rao, Natrajan) An alternate interpretation of M-SBL algorithm:

$$\min_{\mathbf{X}} ||\mathbf{Y} - \mathbf{AX}||_F^2 + \sigma^2 g_{MSBL}(\mathbf{X}) \qquad \text{(Solve for } X\text{)}$$

$$g_{MSBL}(\mathbf{X}) = \min_{\gamma \geq 0} Tr(\mathbf{X}^T \Gamma^{-1} \mathbf{X}) + N \log |\sigma^2 \mathbf{I} + \mathbf{A}\Gamma\mathbf{A}^T| \qquad \text{(Solve for } \gamma\text{)}$$

- **Rank proxy interpretation for log-det term**
- As $\sigma^2 \to 0$,

$$g_{MSBL}(\mathbf{X}) \approx \min_{\gamma \geq 0} N||\gamma||_0 + N \cdot RankProxy(\mathbf{A}\Gamma^{\frac{1}{2}})$$

- Log-det penalty can be replaced with a better penalty: Rank-Proxy($Q^T \mathbf{A}\Gamma^{\frac{1}{2}}$), where Q is a basis of noise subspace, i.e. $R(Q) = R^{\perp}(\mathbf{Y})$

Improving M-SBL for Joint Sparse Recovery using a Subspace Penalty

- (Wipf, Rao, Natrajan) An alternate interpretation of M-SBL algorithm:

$$\min_{\mathbf{X}} ||\mathbf{Y} - \mathbf{AX}||_F^2 + \sigma^2 g_{MSBL}(\mathbf{X}) \qquad \text{(Solve for } X\text{)}$$

$$g_{MSBL}(\mathbf{X}) = \min_{\gamma \geq 0} Tr(\mathbf{X}^T \Gamma^{-1} \mathbf{X}) + N \log |\sigma^2 \mathbf{I} + \mathbf{A} \Gamma \mathbf{A}^T| \qquad \text{(Solve for } \gamma\text{)}$$

- **Rank proxy interpretation for log-det term**
- As $\sigma^2 \to 0$,

$$g_{MSBL}(\mathbf{X}) \approx \min_{\gamma \geq 0} N||\gamma||_0 + N \cdot RankProxy(\mathbf{A}\Gamma^{\frac{1}{2}})$$

- Log-det penalty can be replaced with a better penalty: Rank-Proxy($Q^T \mathbf{A} \Gamma^{\frac{1}{2}}$), where Q is a basis of noise subspace, i.e. $R(Q) = R^{\perp}(\mathbf{Y})$
- **Why** $RankProxy(Q^T \mathbf{A} \Gamma^{\frac{1}{2}})$ **instead of** $RankProxy(\mathbf{A}\Gamma^{\frac{1}{2}})$?

Theorem

If $||\mathbf{X}||_0$, $\mathbf{Y} = \mathbf{AX}$ *and if* $\mathbf{A}$ *satisfies RIP, then we have:*

$$support(\mathbf{X}) = \arg\min_{|I| \geq k} rank(Q^T A_I)$$

- Using *Schatten-p* norm as rank-proxy leads to performance better than M-SBL.

Bayesian Hypothesis Testing for Block Sparse Signal Recovery

Authors: Mehdi Korki, Hadi Zayyani and Jingxin Zhang

- **Problem statement**
  - Recover a **block sparse x** $\in \mathbb{R}^n$ from $m(< n)$ measurements **y** $\in \mathbb{R}^m$, where

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{w}$$

  - Blocks are of unequal sizes and with unknown boundaries
- **Main contributions**
  - Proposes Block Bayesian Hypothesis Testing Algorithm (Block-BHTA)
  - Assumes Bernoulli Gaussian signal prior:

$$p(x_i) = p\delta(x_i) + (1-p)\mathcal{N}(0, \sigma_s^2)$$

  - Transitional probabilities $p_{01} \triangleq Pr(x_{i+1} = 0 | x_i \neq 0)$ and $p_{10} \triangleq Pr(x_{i+1} \neq 0 | x_i = 0)$ are used to encode block structure.
  - In steady state (large n),

$$Pr(x_i = 0) = p = \frac{p_{10}}{p_{01} + p_{10}} \qquad Pr(x_i = 1) = p = \frac{p_{01}}{p_{01} + p_{10}}$$

  - Block start identified at index i if:

$$p_{10}p(\mathbf{y}|x_i = 0, x_{i+1} \neq 0) > p_{00}p(\mathbf{y}|x_i = 0, x_{i+1} = 0)$$

  - Similar test for identification of block termination

Other Interesting Papers:

- Sparse Multinomial Logistic Regression via Approximate Message Passing, Evan Byrne and Philip Schniter
- Bayesian Masking: Sparse Bayesian Estimation with Weaker Shrinkage Bias, Kondo, Hayashi and Maeda
- Recovery of Sparse Positive Signals on the Sphere from Low Resolution Measurements, Tamir Bendory and Yonina C. Eldar
- SAFFRON: A Fast, Efficient and Robust Framework for Group Testing based on Sparse Graph Codes, Kangwook Lee, Ramtin Pedarsani and Kannan Ramachandran
- Bayesian Optimal Approximate Message Passing to Reciver Structured Sparse Signals, Martin Mayer and Nobert Goertz
- Super-Resolution Sparse MIMO-OFDM Channel Estimation Based on SPatial and Temporal Correlations, Zaocheng Wang et al.
- A framework for sparse online learning and its applications, Dayong Wang, Pengchecg Wu, Peilin Zhao, Steven C. H. Hoi
- Type-I and Type-II Bayesian Methods for Sparse Signal Recovery using Scale Mixtures, Ritwik Giri and Bhaskar D. Rao