Summer Term Presentation

Geethu Joseph

Indian Institute of Science, Bangalore

August 10, 2013

Overview

Compressive Sensing Problem

- Sparse representation of signals
- Solving a system of linear equations

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} \tag{1}$$

- Sparsest solution
- Requires only $m \sim \mathcal{O}(k \log(N/k))$ measurements for reconstruction

Sparse Signal Recovery: Post-processing Measurements

- A linear transformation on measurements that can improve the performance of a recovery algorithm
- $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$ is converted to an equivalent $\boldsymbol{P}\boldsymbol{y} = \boldsymbol{P}\boldsymbol{A}\boldsymbol{x}$.
- $(\boldsymbol{P}\boldsymbol{y}, \boldsymbol{P}\boldsymbol{A})$ is fed to decoder instead of $(\boldsymbol{y}, \boldsymbol{A})$
- How to design such a $\boldsymbol{P}$?
- Which attributes of $\boldsymbol{A}$ should be improved?

QR decomposition & OMP

- OMP depends on mutual coherence
- Orthogonalize rows via economic QR decomposition

$$\boldsymbol{A}^{\mathsf{T}} = \boldsymbol{Q}^{\mathsf{T}}\boldsymbol{R} \implies \boldsymbol{R}^{-\mathsf{T}}\boldsymbol{A} = \boldsymbol{Q} \tag{2}$$

- $(\boldsymbol{R}^{-\mathsf{T}}\boldsymbol{y}, \boldsymbol{Q})$ instead of $(\boldsymbol{y}, \boldsymbol{A})$
- Merits
  - Mutual coherence has improved
  - Requires lesser number of measurements

Mutual Coherence

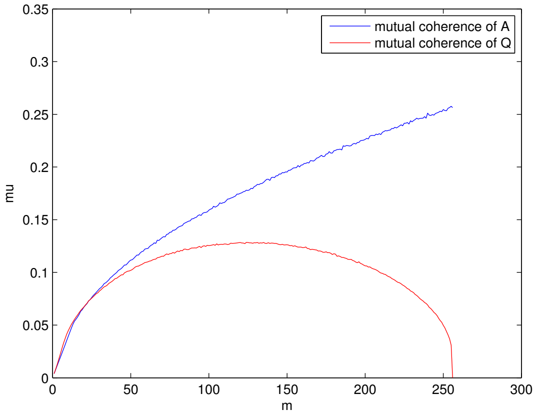

Figure: Improvement in mutual coherence

OMP Performance

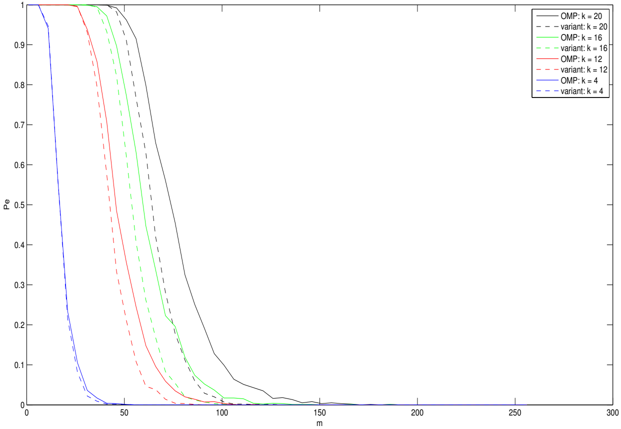

Figure: Improvement in number of measurements

Issues with QR Approach

- Computational complexity for QR decomposition $\sim \mathcal{O}(Nm^2)$
- Operations required for OMP $\sim \mathcal{O}(Nmk)$
- With CoSaMP
    - No improvement in number of measurements
    - Lesser number of iterations required for the same stopping criteria
    - Complexity of one iteration $\sim \mathcal{O}(Nm)$

Two Phase Sparse Recovery Algorithm

- Lossless recovery of strictly sparse signals
- Phase I - Simple and fast algorithm
- Phase II - Sophisticated algorithms for phase II recovery
- Use group testing ideas for phase I recovery
  - Sparse measurement matrix to identify entries of sparse vector
  - Low quality recovery
- Trade-off between accuracy and computational complexity

Recovery using Sparse Matrix

- 'Sudocodes'
- Entries of input corresponding to zero measurement, are all zeros
- If two measurements match, then they corresponds to same set of non-zero entries of input vector
- Entries that corresponds to one measurement and not present the other, are all zeros

A Simple Example - Recovery using Sparse Matrix

$$\begin{bmatrix} 0 \\ 5 \\ 9 \\ 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} ? \\ ? \\ ? \\ ? \\ ? \\ ? \\ ? \end{bmatrix}$$

A Simple Example - Recovery using Sparse Matrix

$$\begin{bmatrix} 0 \\ 5 \\ 9 \\ 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} ? \\ ? \\ ? \\ ? \\ ? \\ ? \\ ? \end{bmatrix} \implies \boldsymbol{x} = \begin{bmatrix} 0 \\ ? \\ 0 \\ ? \\ 0 \\ ? \\ ? \end{bmatrix}$$

A Simple Example - Recovery using Sparse Matrix

$$\begin{bmatrix}0\\5\\9\\5\end{bmatrix} = \begin{bmatrix}1&0&1&0&1&0&0\\0&1&1&0&0&1&0\\1&1&0&1&0&0&1\\0&1&0&0&1&0&1\end{bmatrix}\begin{bmatrix}?\\?\\?\\?\\?\\?\\?\end{bmatrix} \Longrightarrow \boldsymbol{x} = \begin{bmatrix}0\\?\\0\\?\\0\\?\\?\end{bmatrix} \rightarrow \begin{bmatrix}0\\5\\0\\?\\0\\0\\0\end{bmatrix}$$

A Simple Example - Recovery using Sparse Matrix

$$\begin{bmatrix}0\\5\\\color{blue}{9}\\5\end{bmatrix} = \begin{bmatrix}1&0&1&0&1&0&0\\0&1&1&0&0&1&0\\\color{blue}{1}&\color{blue}{1}&\color{blue}{0}&\color{red}{1}&\color{blue}{0}&\color{blue}{0}&\color{blue}{1}\\0&1&0&0&1&0&1\end{bmatrix}\begin{bmatrix}?\\?\\?\\?\\?\\?\\?\end{bmatrix} \implies \boldsymbol{x} = \begin{bmatrix}0\\?\\0\\?\\0\\?\\?\end{bmatrix} \rightarrow \begin{bmatrix}0\\5\\0\\?\\0\\0\\0\end{bmatrix} \rightarrow \begin{bmatrix}0\\5\\0\\\color{red}{4}\\0\\0\\0\end{bmatrix}$$

Some Questions

- Optimal sparsity of sparse matrix?
- Number of measurements are required in each phase?
- Possible recovery techniques in phase II?
- How to handle noisy measurements?

Estimation of Function in Parametric Form

- A continuous and smooth function characterized by an a set of unknown parameters
- Function $f(z;\alpha)$
  - z representing a point in its domain
  - α is vector of parameters
- Given function values at m points, $\{r_i = f(z_i;\alpha)\}_{i=1}^{m}$

Compressive Sensing Problem Formulation

- Form an over-complete basis with discretized values of parameters $\{\alpha_j\}_{j=1}^{N}$

$$\underbrace{\begin{bmatrix} r_1 \\ r_2 \\ . \\ . \\ . \\ r_m \end{bmatrix}}_{\mathbf{r}} = \underbrace{\begin{bmatrix} f(z_1;\alpha_1) & f(z_1;\alpha_2) & . & . & . & f(z_1;\alpha_N) \\ f(z_2;\alpha_1) & f(z_2;\alpha_2) & . & . & . & f(z_2;\alpha_N) \\ . & . & & & & . \\ . & . & & & & . \\ . & . & & & & . \\ f(z_{m,};\alpha_1) & f(z_m;\alpha_2) & . & . & . & f(z_2;\alpha_N) \end{bmatrix}}_{\mathbf{\Phi}} \underbrace{\begin{bmatrix} q_1 \\ q_2 \\ . \\ . \\ . \\ q_N \end{bmatrix}}_{\mathbf{q}} \quad (3)$$

- Sparse solution gives the estimate of solution

Compressive Sensing Approach

- No exhaustive search
- Approximation of function using linear combination of more than one template functions
- Is it always possible to find a suitable set of candidates?
- Applications in Spectrum Cartography

Spectrum Cartography

- Estimating power distribution in frequency and space
- Applications in Wireless Cognitive Radio (CR) network
- Goals
  - Spatial reuse of frequency
  - Transmit power estimation
  - Tracking activities of primary users

Problem

- Sensors collect periodogram samples of received signal at sampling frequencies
- Sensors co-operate to estimate the PSD map
- Unknowns
  - Transmitter location
  - Transmit power
  - Bandwidth

Sparsity

- Sparsity arises due to
  - Narrow-band nature of transmit-PSD
  - Sparsely located transmitters
- Compressed Sensing Techniques?
- How to form a suitable basis?

System Model

- Basis Expansion Models (BEM)
- N_S transmitting sources
  - Stationary
  - Mutually uncorrelated
- N_r sensors located at $\mathcal{X} = \{\mathbf{x}_r\}_{r=1}^{N_r}$
- Bandwidth available B
- Measurements $\phi_r(f_n)$, $r = 1, 2, ..., N_R$ and $n = 1, 2, ..., N$

BEM

- Transmit PSD: For $s = 1, 2, ..., N_s$

$$\phi_s(f) = \sum_{v=1}^{N_b} \theta_{sv} b_v(f) \qquad (4)$$

- $\{b_v(f)\}$ is centered at f_v , $v = 1, 2, ..., N_b$

Basis Functions in Frequency Domain

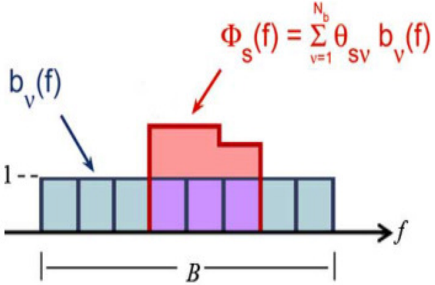

Figure: Basis functions in frequency domain

Finding Transmitter Locations

- Assumption : Channel gains in parametric form
  - Simple choice is path loss model
  - $\gamma_{sr} = min\{1, (||\mathbf{x}_s - \mathbf{x}_r||_2/d_0)^{-\alpha}\}$
- Choice of basis
  - Over-complete basis functions spanning the bandwidth B
  - Various choices of center frequencies, bandwidth,...etc

Modeling PSD

- When source locations are known, linear model

$$\begin{aligned}\phi_r(f_n) &= \sum_{s=1}^{N_s} \gamma_{sr}\phi_s(f_n) + \sigma_r^2 \\ &= \boldsymbol{b}_r^{\mathsf{T}}(f_n)\boldsymbol{\theta} + \sigma_r^2\end{aligned}$$

- Using matrix notations

$$\boldsymbol{\Phi}_r = \boldsymbol{B}_r\boldsymbol{\theta} + \sigma_r^2\mathbf{1} \quad (5)$$

- $\boldsymbol{\theta}$ is common to all sensors
- When source locations are unknown, virtual grid model is considered

Virtual Grid Model

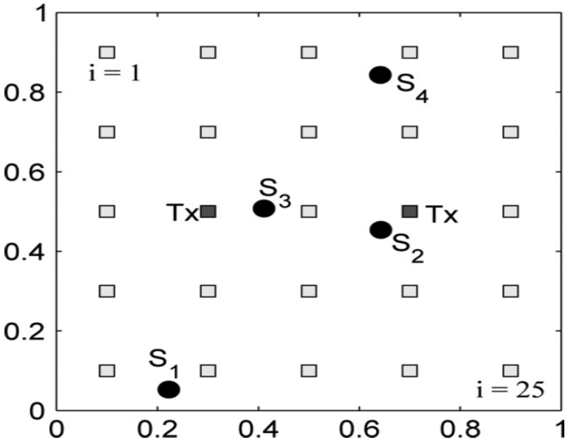

Figure: Virtual network grid with 25 candidate locations, 2 transmitters and 4 sensors

Solution via LASSO

- Basis functions corresponding to all combinations of center frequencies and possible transmitter locations
- LASSO criterion

$$\min_{\boldsymbol{\theta}\geq 0, \sigma_r^2 \geq 0} \sum_{r=1}^{N_r} ||\boldsymbol{\Phi}_r - \boldsymbol{B}_r\boldsymbol{\theta} - \sigma_2^r\mathbf{1}||^2 + \lambda\mathbf{1}^T\boldsymbol{\theta} \quad (6)$$

- Sparse solution reveal the location of sources and their bands

Issues with the approach

- Only path-loss model is studied
- Shadowing and mobility are not included in the model
- Generally unknown spatial loss function : $l_s(\mathbf{x})$
  - Frequency flat channel between source s and receiver at $\mathbf{x}$
- PSD model

$$\begin{aligned}\phi(\mathbf{x}, f) &= \sum_{s=1}^{N_s} l_s(\mathbf{x}) \sum_{v=1}^{N_b} \theta_{sv} b_v(f) \\ &= \sum_{v=1}^{N_b} g_v(\mathbf{x}) b_v(f)\end{aligned}$$

- Solved using spline based technique

Spline-Based Cartography

- Estimation

$$
\begin{aligned}
\{g_v\}_{v=1}^{N_b} \quad = \quad & \arg\min_{g_v} \frac{1}{N} \sum_{r=1}^{N_r} \sum_{n=1}^{N} \left(\phi_{rn} - \sum_{v=1}^{N_b} g_v(\mathbf{x}_r) b_v(f_n) \right)^2 \\
& + \lambda \sum_{v=1}^{N_b} \int_{\mathbb{R}^2} ||\nabla^2 g_v(\mathbf{x})||_F^2 d\mathbf{x} \\
& + \mu \sum_{v=1}^{N_b} || \begin{bmatrix} g_v(\mathbf{x}_1) & . & . & . & g_v(\mathbf{x}_{N_r}) \end{bmatrix}^T ||_2
\end{aligned}
$$

Spline-Based Cartography

Least Square Term

$$\{g_v\}_{v=1}^{N_b} = \arg\min_{\{g_v \in \mathcal{S}\}} \frac{1}{N} \sum_{r=1}^{N_r} \sum_{n=1}^{N} \left(\phi_{rn} - \sum_{v=1}^{N_b} g_v(\mathbf{x}_r) b_v(f_n) \right)^2$$
$$+\lambda \sum_{v=1}^{N_b} \int_{\mathbb{R}^2} ||\nabla^2 g_v(\mathbf{x})||_F^2 d\mathbf{x}$$
$$+\mu \sum_{v=1}^{N_b} || \left[g_v(\mathbf{x}_1) \quad . \quad . \quad . \quad g_v(\mathbf{x}_{N_r}) \right]^T ||_2$$

Spline-Based Cartography

Roughness Regularization Term

$$\{g_v\}_{v=1}^{N_b} = \arg\min_{\{g_v \in \mathcal{S}\}} \frac{1}{N} \sum_{r=1}^{N_r} \sum_{n=1}^{N} \left(\phi_{rn} - \sum_{v=1}^{N_b} g_v(\mathbf{x}_r) b_v(f_n) \right)^2 + \lambda \sum_{v=1}^{N_b} \int_{\mathbb{R}^2} ||\nabla^2 g_v(\mathbf{x})||_F^2 d\mathbf{x} + \mu \sum_{v=1}^{N_b} || \left[g_v(\mathbf{x}_1) \quad . \quad . \quad . \quad g_v(\mathbf{x}_{N_r}) \right]^T ||_2$$

Spline-Based Cartography

Sparsity Inducing Term

$$\{g_v\}_{v=1}^{N_b} = \arg\min_{\{g_v \in \mathcal{S}\}} \frac{1}{N} \sum_{r=1}^{N_r} \sum_{n=1}^{N} \left(\phi_{rn} - \sum_{v=1}^{N_b} g_v(\mathbf{x}_r) b_v(f_n) \right)^2 + \lambda \sum_{v=1}^{N_b} \int_{\mathbb{R}^2} ||\nabla^2 g_v(\mathbf{x})||_F^2 d\mathbf{x} + \mu \sum_{v=1}^{N_b} || \left[g_v(\mathbf{x}_1) \quad . \quad . \quad . \quad g_v(\mathbf{x}_{N_r}) \right]^T ||_2$$

Thin-spline solution

- Closed form expression for thin-plate splines

$$\hat{g}_v(\mathbf{x}) = \sum_{r=1}^{N_r} \beta_{vr} K(||\mathbf{x} - \mathbf{x}_r||_2) + \boldsymbol{\alpha}_{v1}^T \mathbf{x} + \alpha_{v0} \qquad (7)$$

  - $K(\rho) = \rho^2 log(\rho)$ and $\boldsymbol{\beta}_v = \begin{bmatrix} \beta_{v1} & . & . & . & \beta_{vN_r} \end{bmatrix}^T \in \mathcal{B}, v = 1, 2, ..., N_b$
  - $\mathcal{B} = \{\boldsymbol{\beta} \in \mathbb{R}^{N_r} : \sum_{r=1}^{Nr} \beta_r = 0, \sum_{r=1}^{Nr} \beta_r \mathbf{x}_r = \mathbf{0} : \mathbf{x}_r \in \mathcal{X}\}$

- Parameters $\boldsymbol{\alpha} = \begin{bmatrix} \alpha_{10} & \boldsymbol{\alpha}_{11}^T & . & . & . & \alpha_{N_b0} & \boldsymbol{\alpha}_{N_b1}^T \end{bmatrix}$ and $\boldsymbol{\beta} = \begin{bmatrix} \boldsymbol{\beta}_1^T & . & . & . & \boldsymbol{\beta}^T{}_{N_b} \end{bmatrix}$ are to be estimated

Equivalent Group LASSO

- Equivalent LASSO problem

$$\min_{\zeta} \frac{1}{2}||\boldsymbol{y} - \boldsymbol{X}\zeta||_2^2 + \mu \sum_{v=1}^{N_b} ||\zeta_v||^2 \tag{8}$$

  - Solution: $\zeta = \begin{bmatrix} \zeta_1 & \dots & \zeta_{N_b} \end{bmatrix}^T$
  - $\boldsymbol{x}$ - function of $\{b_v(f)\}_{v=1}^{N_b}, \{\boldsymbol{x}_r\}_{r=1}^{N_r}$
  - $\boldsymbol{y}$ - function of $\{\phi_{rm}\}_{r=1}^{N_r}, m = 1, 2, ..., N$
- Optimal parameters are obtained by change of variables : $\begin{bmatrix} \beta_v^T & \alpha_v^T \end{bmatrix}^T = \boldsymbol{T}\zeta_{\boldsymbol{v}}$

Issues & Project Plan

- Issues
  - Location of sensors
  - Accuracy of model
  - Candidate set for parameters
- Plans
  - Algorithm, analysis and comparative study with existing literature
  - Reduced Complexity: Distributed Algorithm
  - PSD Tracker

Summary

- Various problems in Compressed Sensing studied during summer
- Spectrum cartography models
- Compressed Sensing techniques in estimating PSD map

Equivalent LASSO

- Three matrix are involved in obtaining optimal solution:

1. $\boldsymbol{T} \in \mathbb{R}^{N_r \times 3}$ with r^{th} row $\boldsymbol{t}_r = \begin{bmatrix} 1 & \boldsymbol{x}_r^T \end{bmatrix}$, for $r = 1, 2, ..., N_r$
    1. $\boldsymbol{B} \in \mathbb{R}^{N \times N_r}$ with n^{th} row $\boldsymbol{b}_n = \begin{bmatrix} b_1(f_n) & . & . & . & b_{N_b}(f_n) \end{bmatrix}$
    2. $\boldsymbol{K} \in \mathbb{R}^{N_r \times N_r}$ with $(i,j)^{\text{th}}$ entry $[\boldsymbol{K}]_{ij} = K(||\boldsymbol{x}_i - \boldsymbol{x}_j||)$

- QR decomposition of $\boldsymbol{T} = \begin{bmatrix} \boldsymbol{Q}_1 & \boldsymbol{Q}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{R} \\ \boldsymbol{0} \end{bmatrix}$

Equivalent LASSO

- $\boldsymbol{y} = \frac{1}{\sqrt{N_r N}}[\boldsymbol{\Phi}^{\mathsf{T}} \boldsymbol{0}]^{\mathsf{T}}$
- $\boldsymbol{X} = \frac{1}{\sqrt{N_r N}} \begin{bmatrix} \boldsymbol{B} \otimes \boldsymbol{I}_{N_r} \\ \boldsymbol{I}_{N_b} \otimes \{\text{bdaig}\{(N_r N \lambda \boldsymbol{Q}_2^{\mathsf{T}} \boldsymbol{K} \boldsymbol{Q}_2)^{1/2}, \boldsymbol{0}\}[\boldsymbol{K}\boldsymbol{Q}_2 \boldsymbol{T}]^{-1} \end{bmatrix}$
- Optimal parameters are obtained by change of variables : $\begin{bmatrix} \beta_v^{\mathsf{T}} & \alpha_v^{\mathsf{T}} \end{bmatrix}^{\mathsf{T}} = \text{bdaig}\{\boldsymbol{Q}_2, \boldsymbol{I}_3\}[\boldsymbol{K}\boldsymbol{Q}_2 \boldsymbol{T}]^{-1} \zeta_{\boldsymbol{v}}$