Dimensionality Estimation in Rank Deficient Noise using Profile Likelihood & Sparse Signal Recovery

Vinuthna Vinjamuri

SPC Lab
Indian Institute of Science, Bangalore, India.

16 May 2015

Outline

System Model

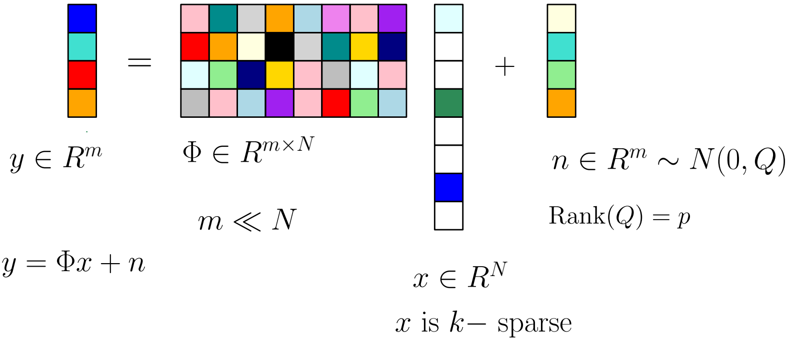

Figure: System Model

Goal: Recover **x** from measurements **y** when, **Q** and $\boldsymbol{\Phi}$ are known

Solution - CoNoSBL

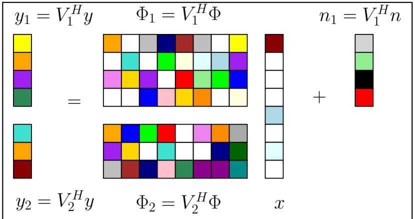

Figure: Equivalent System Model

- Modified EM-SBL to CoNo-SBL that recovers **x** from a mixture of noisy and noiseless measurements
- CRLB on MSE of **x**, when **x** is a compressible signal
- Assumes that **Q** is known
- Analysed time complexity
- Solution when **Q** is unknown?

Non-uniform noise

- EM-SBL suffers from identifiability problem even when $\mathbf{Q} = \sigma^2\mathbb{I}$ is unknown
- Consider estimation of noise statistics when we have non-uniform noise i.e., $\mathbf{Q} = D$ using MMV setup

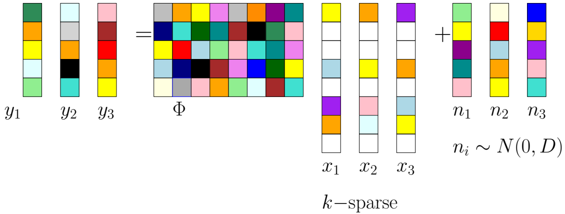

Figure: System Model

Summary - SBL

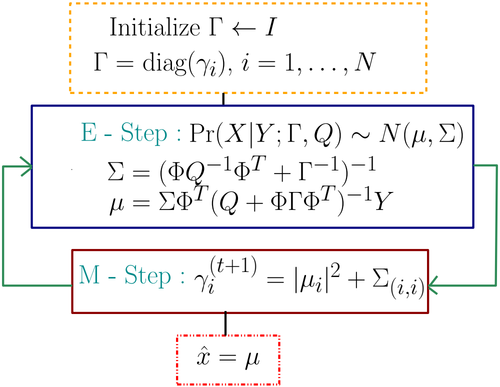

Figure: Summary of SBL

Algorithm

1. $\bar{\mathbf{D}} = \mathcal{I}_m$
2. compute $\hat{\Gamma} = MSBL(\mathbf{Y}, \mathbf{A}, \hat{\mathbf{D}})$
3. Find $S =$ {indices corresponding to k maximum magnitude entries}
4. $P = \mathbf{A}_S(\mathbf{A}_S^H \mathbf{A}_S)^{-1}\mathbf{A}_S^H$
5. $\hat{\mathbf{D}} = \frac{1}{L}\sum_{i=1}^{L}(\mathcal{I} - P)\mathbf{Y}$
6. Repeat 2 to 5 till convergence

Convergence criteria:
$||\Gamma^{t+1} - \Gamma^{t}||_F \leq 10^{-4} \& ||\hat{\mathbf{D}}^{t+1} - \hat{\mathbf{D}}^{t}||_F \leq 10^{-2}$

Simulation Results

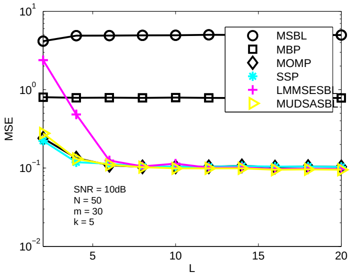

Figure: MSE *vs.* L

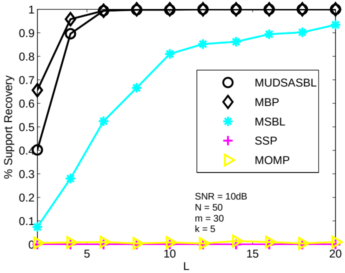

Figure: Support recovery *vs.* L

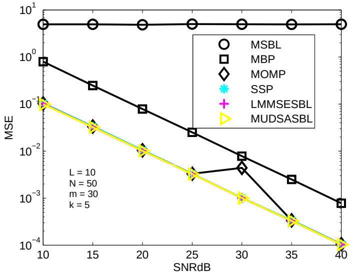

Figure: MSE *vs.* SNR

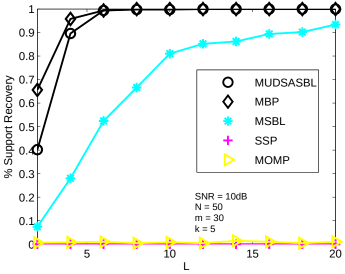

Figure: Support Recovery *vs.* SNR

Analysis

- Let $\mathcal{S} = \{\text{Support of } \mathbf{x}\}$
- $\mathbf{y}_{\mathcal{S}} = \Phi_{\mathcal{S}} \mathbf{x}_{\mathcal{S}} + \mathbf{n}_{\mathcal{S}}$
- $\mathbf{y}_{\mathcal{S}^c} = \mathbf{n}_{\mathcal{S}^c}$
- $\hat{D} = \frac{1}{L} \sum diag(\frac{1}{m} \sum \mathbf{y}^{H}_{\mathcal{S}^c}(:, i)\mathbf{y}^{c}_{\mathcal{S}}(:, i))$
- Estimation of $\hat{\mathbf{D}}$ involves estimating m parameters
- From simulations, it is seen that $L = \mathcal{O}(k)$ measurements are required for successful recovery

Rank-Deficient noise

- Involves estimation of $m^2 + 1$ parameters(m^2 elements and p)
- MUDSASBL gives poor estimate of covariance matrix
- Follow PCA approach:
- $\Sigma_y = \mathbf{\Phi} \Gamma \mathbf{\Phi}^T + \mathbf{Q}$
- $\tilde{\mathbf{Q}} = \Sigma_y - \mathbf{\Phi} \Gamma \mathbf{\Phi}^T$
- Use $\tilde{\mathbf{Q}}$ as an estimate of $\mathbf{Q}$
- To use CoNo-SBL, need to identify the dimensionality of underlying noise subspace ($\{p\}$)

PCA

- Dimensionality reduction - converts a set of observations of possibly correlated variables into a set of values of uncorrelated variables called as principal components
- $\{x_i\}_{i=1}^{n} \in R^m$ be the data matrix
- Goal: To find $\alpha_1, ..\alpha_p$ such that $Var(\alpha_1^T x_i) \geq Var(\alpha_2^T x_i) \geq \cdots Var(\alpha_p^T x_i)$ and $Cov(\alpha_k^T x_i, \alpha_l^T x_i) = 0 \forall k \neq l$
- Mathematically, if $d_j = Var(\alpha_j^T x_i)$ and S is the sample covariance matrix of the data, then PCA: $R^m \rightarrow R^p$ and $\alpha_1, \cdots \alpha_p$ are eigen vectors of S and $d_1, \cdots, d_j$ are corresponding eigen values

What value of p to choose from the eigen values!!

- Percent Variance: Find q between 1 and m such that

$$\frac{d_1 + d_2 + \cdots + d_q}{d_1 + d_2 + \cdots + d_m} \geq \gamma$$

, where γ is a pre-determined proportion, say 80% or 90%

- Scree test: Plot the eigenvalues $d_1, d_2, ..., d_p$ in descending order (scree plot) and look for a "big gap" or an "elbow" in such a graph.
- Sequential tests:
  For $j = 1, 2, \cdots, m - 1$, consider a series of null hypotheses:

$$H_{0,j} : d_m = d_{m1} = \cdots = d_{mj}$$

  Start by testing $H_{0,1}, H_{0,2}, \cdots$ until a null hypothesis is first rejected. Suppose $H_{0,q}$ is the first rejected null hypothesis, then the first m q components are retained.

Drawbacks:

- Threshold γ
- No objective function to determine gap or elbow
- Sequential detection assumes that data is generated from multivariate gaussian distribution
- Computationally expensive

Objective: To propose a simple method to find gap in a objective and automated way

- Main idea: assume a distribution on d_j's and find p by maximising profile likelihood
- Profile Likelihood: Suppose $l(\theta, \psi; y)$ be the likelihood function, θ - main parameter and ψ- nuisance parameter, then profile likelihood for θ is defined as

$$l_\theta(\theta; y) = l(\theta, \hat{\psi}_\theta; y)$$

  $\hat{\psi}_\theta$ is MLE of ψ for fixed θ
- Advantages of profile likelihood:
  - Always available
  - Maximum of profile likelihood is always same as MLE of θ

- $d_1 \geq d_2 \geq \cdots d_m$ be the eigen values
- for a fixed number $1 \leq p \leq m$, $\mathcal{S}_1 = \{d_1, \cdots d_p\}$ and $\mathcal{S}_2 = \{d_{p+1}, \cdots d_m\}$
- If an elbow or gap exists at p, then $\mathcal{S}_1 \in f(d;\theta_1)$ and $\mathcal{S}_2 \in f(d;\theta_2)$. Assume $\mathcal{S}_1$ and $\mathcal{S}_2$ are independent
- p is the main parameter of interest

$$l(p, \theta_1, \theta_2) = \sum_{i=1}^{p} \log f(d;\theta_1) + \sum_{j=p+1}^{m} \log f(d;\theta_2)$$

θ_1 and θ_2 are unknown and should be computed from data

Let $\hat{\theta_1}$ and $\hat{\theta_2}$ be MLEs of θ_1, θ_2 respectively

$$l_p(p) = \sum_{i=1}^{p} \log f(d; \theta_1\hat{(p)}) + \sum_{j=p+1}^{m} \log f(d; \theta_2\hat{(p)})$$

$$\hat{p} = argmax_p l_p(k) \forall k = 1, \cdots, m$$

- Choose f to be Gaussian distribution
- The parameters will now be $\theta_1 = \{\mu_1, \sigma^2\}$ and $\theta_2 = \{\mu_2, \sigma^2\}$

- MLEs for μ_1, μ_2 and σ^2 are computed using sample averages,

$$\hat{\mu_1} = \frac{\sum_{i \in \mathcal{S}_1 d_i}}{p}$$

$$\hat{\mu_2} = \frac{\sum_{j \in \mathcal{S}_2 d_j}}{m - p}$$

and

$$\hat{\sigma^2} = \frac{(p-1)s_1^2 + (m-p-1)s_2^2}{m-2}$$

where s_j^2 is sample variance of $\mathcal{S}_j$

Numerical results

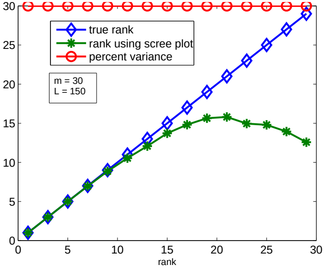

Figure: Rank estimation using screeplot *vs.* percent variance

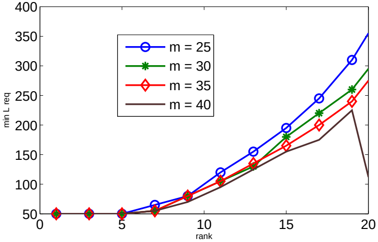

Figure: Phase transition diagram for estimation of rank using scree plot

UCoNo-SBL

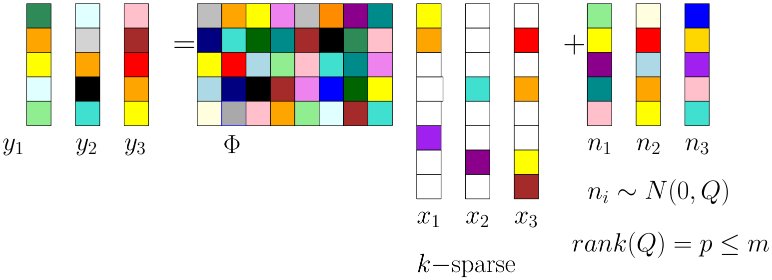

Figure: System model - Unknown noise covariance matrix

Proposed method

Algorithm:

1. Let $x_i \sim \mathcal{N}(0, \Gamma_i)$
2. $\Gamma_i \leftarrow SBL(y_i, \Phi, \mathbb{I})$
3. Compute $\hat{\mathbf{Q}} = \frac{1}{L}\sum_{i=1}^{L}(y_i y_i^T - \Phi\Gamma_i\Phi^T)$
4. find $\hat{p}$ using eigen values of $\mathbf{Q}$, let V_p be the corresponding eigen vectors
5. $Y_1 = V_p^H Y,\ Y_2 = V_{m-p}^H Y,\ \Phi_1 = V_p^H\Phi,\ \Phi_2 = V_{m-p}^H\Phi$ and $\hat{D} = \Lambda(1:p)$

$$(X_i, \Gamma_i) = CoNoSBL(Y_1(i), Y_2(i), \Phi_1, \Phi_2, \hat{D})\forall i$$

6. Repeat 3 to 5 until convergence

Analysis

- Convergence criteria considered is convergence of $\mathbf{Q}$ in terms of frobenius norm on error of covariance matrix
- Since, we are using sample covariance matrix as an estimate of Covariance matrix, requires $L \geq m$

- Consider some training samples to estimate the noise subspace
- Compute $\hat{p}$ from $\hat{\mathbf{Q}}$
- Use CoNoSBL with this estimate of $\mathbf{Q}$
- Identical to using an estimated $\mathbf{Q}$ instead of true $\mathbf{Q}$
- Relatively faster compared to previous approach, however naive in implementation

- Use some other metrics for convergence like chordal distance, spectral distance etc
- Propose a method to recover $\mathbf{Q}$ using lesser number of measurements
- Theoretical guarantees on convergence of Γ and $\mathbf{Q}$
- Analysis on Phase transition of CoNo-SBL

Thank You