Journal Watch
IEEE Information Theory- June,2018

Pradip Sasmal

Indian Institute of Science, Bangalore

August 25, 2018

Organization

- Noisy Subspace Clustering via Matching Pursuits
- Weighted Matrix Completion and Recovery With Prior Subspace Information
- Generalized Line Spectral Estimation via Convex Optimization

Noisy Subspace Clustering via Matching Pursuits (M.Tschannen and H. Bolcskei)

- sparse subspace clustering(SSC), Thresholding-SC (TSC), nearest subspace neighbor (NSN)
- SSC-OMP, SSC-MP - proposed, Analytic performance characterization for noisy data. Better clustering performance and running time.
- $Y = Y_1 \cup \cdots \cup Y_L$, $Y_l \in S_\ell \subset \mathbb{R}^m$, $\dim(S_\ell) = d_\ell$, randomly distributed on $S_\ell \cap \mathbb{S}^{m-1}$, perturbed by additive random noise
- $y_i^\ell = x_i^\ell + z_i^\ell = U^{(\ell)} a_i^\ell + z_i^\ell$, $i \in [n_\ell]$, $U^{(\ell)} \in \mathbb{R}^{m \times d_\ell} -$ orthonormal basis for S_ℓ, a_i^ℓ's independently and uniformly distributed on $\mathbb{S}^{d_\ell}$, $z_i^\ell \sim iid\mathcal{N}(0, \frac{\sigma^2}{m})$

- $aff(S_k, S_\ell) =: \frac{1}{\sqrt{\min\{d_k, d_\ell\}}} \| U^{(k)^T} U^{(\ell)} \|_F = \sqrt{\frac{cos^2(\theta_1) + \cdots + cos^2(\theta_{\min\{d_k, d_\ell\}})}{\min\{d_k, d_\ell\}}}$, $\theta_1 \leq \cdots \leq \theta_{\min\{d_k, d_\ell\}}$ principle angles between S_k and S_ℓ.
- No false connections (NFC) property: The graph G satisfies NFC property if, for all $\ell \in [L]$, the nodes corresponding to $_\ell$ are connected to the other nodes corresponding to Y_ℓ only.
- Sampling density $\rho_\ell := \frac{(n_\ell - 1)}{d_\ell}$, $d_{\max} := \max_{\ell \in [L]} d_\ell$, $\rho_{\min} = \min_{\ell \in [L]} \rho_\ell$
- Assume $m \geq 2d_{\max}$, $\rho_{\min} \geq c_\rho$, $\sigma \leq \frac{1}{2}$, $s_{\max} \leq \min_{\ell \in [L]} \{c_s d_\ell / \log(n_\ell - 1)e / s_{\max}\}$, $c_\rho > 1, 0 < c_s \leq 1/10$

Performance guarantee:

- (SSC-OMP with DI-Stopping): The clustering condition

$$\max_{k,\ell:k\neq\ell} \text{aff}(S_k, S_\ell) + \frac{10\sigma}{\sqrt{\log(N^3 s_{max})}}$$

.

$$\left(\frac{\sqrt{d_{\max}}}{\sqrt{m}} c_\rho + \frac{\sqrt{2}}{\sqrt{\rho_{\min}}}(1 + \frac{3}{2}\rho)\right) \le \frac{1}{8\log(N^3 s_{\max})}$$

with $c_\rho = 10 + 13\sigma$ guarantees the graph generated by SSC-OMP satisfies NFC property w.p. at least

$$P^* := 1 - 6/N - 5Ne^{-c_m m} - 6\sum_{\ell\in[L]} n_\ell e^{c_d d_\ell}$$

for $0 < c_d \le 1/18$ and $0 < c_m \le 1/8$.

Weighted Matrix Completion and Recovery With Prior Subspace Information (A. Eftekhari, D. Yang, M. Wakin)

- Goal: Use of prior knowledge on matrix rcovery and completion problem
- $M_{n\times n} = U \sum V^*$, for $r \leq n$ define $U_r, V_r, \sum_r, M_r = U_r \sum_r V_r^*, M_{r^+} = M - M_r$
- Matrix recovery: $y = R_m(M+E)$, $\|R_m(E)\|_2 \leq e$, R_m linear operator collects m measurements of M. Recover M from y.
- Matrix completion: $R_p(M) = \sum_{i,j=1}^{n} \frac{\epsilon_{ij}}{p} M[i,j] C_{ij}$, $\{\epsilon_{ij}\}_{i,j}$ is a sequence of independent Bernoulli random variables taking one with probability p and zero elsewhere, $R(\cdot) : M \rightarrow R_M$ observe each entry of M with probability $p = m/n^2$, $R_p(M)$ contains m entries of M. $Y = R_p(M+E)$, $\|R_p(E)\|_F \leq e$, recover M from Y

Existing results:

- (r, δ_r)–RIP: $(1-\delta_r)\|X\|_F \leq \|R_m(X)\|_2 \leq (1+\delta_r)\|X\|_F$, $rank(X) \leq r$
- $\min_X \|X\|_*, \quad \text{subject to } \|R_m(X) - y\|_2 \leq e \quad (1)$
- $\delta_{5r} \leq 0.1$, then

$$\|\tilde{M} - M\|_F \leq \frac{\|M_{r^+}\|_*}{\sqrt{r}} + e,$$

  $\tilde{M}$ solution of (1)
- $\eta(M_r) = \frac{n}{r}\max\{\|U_r\|^2_{2\to\infty}, \|V_r\|^2_{2\to\infty}\}$, $\|X\|^2_{2\to\infty}$ largest ℓ_2 norm of the rows of X
- $\min_X \|X\|_*, \quad \text{subject to } \|R_p(X) - Y\|_F \leq e \quad (2)$

$$\|\tilde{M} - M\|_F \leq \frac{\|M_{r^+}\|_*}{\sqrt{r}} + e\sqrt{pn},\ 1 \geq p \geq \frac{\eta(M_r) r \log^2 n}{n},$$

  $\tilde{M}$ solution of (2)

Prior information: Matrix recovery

- $\tilde{U}_r, \tilde{V}_r$ prior knowledge on column and row spaces of M_r, define $Q_{\tilde{U}_r,\lambda} := \lambda P_{\tilde{U}_r} + P_{\tilde{U}_r^{\perp}}$ and $Q_{\tilde{V}_r,\rho}$
- $\min_X \|Q_{\tilde{U}_r,\lambda}.X.Q_{\tilde{V}_r,\rho}\|$, sub to $\|R_m(X) - y\|_2 \leq e \quad (3)$
- $U_r = span(M_r), V_r = span(M_r^*), u = \angle[U_r, \tilde{U}_r], v = \angle[V_r, \tilde{V}_r]$
  $R_m(\cdot)$ is $(r, \delta_{32r}) - RIP$ with

$$\delta_{32r} \leq \frac{0.9 - \max[\alpha_3, \alpha_4]/\sqrt{30}}{0.9 + \max[\alpha_3, \alpha_4]/\sqrt{30}}$$

  then

$$\|\tilde{M} - M\|_F \leq \frac{\|M_{r^+}\|_*}{\sqrt{r}} + e,$$

  $\tilde{M}$ solution of (3), $\alpha_i = \text{function}(u, v, \lambda, \rho)$

Prior information: Matrix completion

- $\min_X \|Q_{\tilde{U}_r,\lambda}.X.Q_{\tilde{V}_r,\rho}\|$, sub to $\|R_p(X) - Y\|_F \leq e \quad (4)$
- $$\|\tilde{M} - M\|_F \leq \frac{\|M_{r^+}\|_*}{\sqrt{r}} + e\sqrt{pn},\ 1 \geq p \geq \frac{\eta(M_r) r \log^2 n}{n},$$

  except with probability $o(n^{-19})$ and provided

  $$1 \geq p \geq \max[\log(\alpha_5 n), 1].\frac{\eta(M_r) r \log n}{n}.\max[\alpha(1 + \sqrt{\frac{\eta_{\tilde{U}\tilde{V}^*}}{\eta_{U_r V_r^*}}})]$$

  $\tilde{M}$ solution of (4)

Generalized Line Spectral Estimation via Convex Optimization (R. Heckel and M. Soltanolkotabi)

- $z \in \mathbb{C}^{2N+1}$- equispaced samples of a complex sinusoids
  $z = \sum_{k=1}^{S} b_k f(v_k)$,
  $f(v_k) := [e^{-i2\pi N\nu}, e^{-i2\pi(N-1)\nu}, \ldots, e^{i2\pi N\nu}]^T$
  $b_k \in \mathbb{C}-$coefficients, $\{v_k \in [0,1]\}-$frequency parameter
- Recover b_k, v_k from $y = Az$, where $A \in \mathbb{C}^{M \times L}$ and $L = 2N + 1$.
- A invertible - line spectral estimation problem (LSEP), $M < L$ - generalized LSEP
- Sparse recovery problem:recovery of $\{(b_k, v_k)\}$ corresponding to recovery of a sparse signal in the continuous indexed dictionary $\{Af(\nu) : \nu \in [0,1]\}$

continue ....

- Components of the mixtures are sufficiently separated, $|v_k - v_{k'}| \geq \frac{2}{N}, \forall\, k, k' \in [S]\, k \neq k'$
- F corresponds to a distribution that picks a vector $a \in \mathbb{C}^L$ uniformly at random from the rows of A.
- Isotropy property: $a \sim F$ obey $\mathbb{E}_{a \sim F}[aa^H] = \frac{1}{M}\mathbf{I}$.
- Incoherence property: if for all $a \sim F$ $\sup_{f \in \mathbb{C}^L : \|f\|_\infty \leq 1} |\langle f, a\rangle|^2 = \|a\|_{\ell_1}^2 \leq \frac{L}{M}\mu$, μ incoherence parameter
- Subsampled Orthogonal Matrices: $\mu = \max_k \frac{L}{M}\|u_k\|_{\ell_1}^2$, u_k : $k-$th row of U.

Extensions to Higher Dimensions

- $z = \sum_{k=1}^{S} b_k f(r_k)$, $r_k \in [0,1]^d$, $f(r) = [f(r)_p] = e^{i2\pi\langle p,r\rangle}$, p is an interger vector with $\ell-$th entry, $\ell = 1,\ldots,d$, given by $[p]_\ell = -N,\ldots,N$. If $d=2$ then $f(r) = [e^{i2\pi(-Nr_1-Nr_2)}, e^{i2\pi(-Nr_1-(N-1)r_2)},\ldots,e^{i2\pi(Nr_1+Nr_2)}]$
- $y = Az$, where $A \in \mathbb{C}^{M\times L^d}$ and $L = 2N+1$.
- minimum separation condition: $\max_{\ell=1,\ldots,d} |[r_k]_\ell - [r_{k'}]_\ell| \geq \frac{c(d)}{N}$
- Atomic norm: $\mathcal{A} := \{f(r) : r \in [0,1]^d\}$

$$\|z\|_{\mathcal{A}} := \inf_{b_k\in\mathbb{C}, r_k\in[0,1]^d} \{\sum_k |b_k| : z = \sum_k b_k f(r_k)\}$$

- $$AN(y) : \min_{\tilde{z}} \|\tilde{z}\|_{\mathcal{A}} \quad \text{sub to} \quad y = A\tilde{z} \quad (1)$$

Results

Theorem

$A \in \mathbb{C}^{M \times L^3}$ with $L \geq 1024$, be a random matrix with rows a_r chosen independently from a distribution obeying the isometry and incoherence property
$\mathbb{E}[a_r a_r^H] = \frac{1}{M}\mathbf{I}$ *and* $\sup_{f \in \mathbb{C}^{L^3} : \|f\|_\infty \leq 1} |\langle f, a_r \rangle|^2 \leq \frac{L^3}{M}\mu$ *for fixed* $\mu \geq 1$. *$y = Az$ with $z = \sum_{k=1}^{S} b_k f(r_k)$. $(b_k) \sim$iid symmetric distributions on the complex unit circle and $r_k = (\beta_k, \tau_k, v_k)$obeys minimum separation condition*
$\max(|\beta_k - \beta_{k'}|, |\tau_k - \tau_{k'}|, |\nu_k - \nu_{k'}|) \geq \frac{5}{N}$, $k \neq k'$. *Then as long as*

$$m \geq c \log^2(L/\delta),$$

with c a fixed numerical constant, z is the unique minimizer of $AN(y)$ in (1) with probability at least $1 - \delta$.

Thank You