1-Bit Compressed Sensing Of Positive Semi-Definite Matrices Via Rank-1 Measurement Matrices

Xiyuan Wang*, Kun Wang†, Zhongshan Zhang*, and Keping Long*

*School of Computer & Communication Engineering University of Science and Technology Beijing, China

†National Computer Network Emergency Response Technical Team/Coordination Center of China, Beijing, China

北京科技大学 University of Science and Technology Beijing

Shanghai, China

ICASSP·2016

1-bit Compressed Sensing (CS) of Positive Semi-definite (PSD) Matrix

- Frugal sensing (Mehanna and Sidiropoulos 2013): 1-bit sensing of Toeplitz PSD matrix for spectral estimation and parameter estimation.
- 1-bit sensing of covariance matrix for principle subspace estimation and online line spectrum estimation (Chi 2014a,b).
- Our work: propose 1-bit CS of PSD matrix via rank-1 measurement matrix and provide more accurate recovery error bound.

Measuring via Rank-1 Matrix

- $n \times n$ rank-r PSD matrix $\mathbf{\Sigma} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^T$:
  - $n \times n$ orthogonal matrix $\mathbf{Q}$.
  - $n \times n$ diagonal matrix $\mathbf{\Lambda} = \text{diag}\{\lambda_1, \ldots, \lambda_n\}$ with $\lambda_1 \geq \cdots \geq \lambda_r > \lambda_{r+1} = \cdots = \lambda_n = 0$.
  - Normalized norm: $\|\mathbf{\Sigma}\|_F^2 = \sum_i^r \lambda_i^2 = 1$.
  - Condition number $\kappa = \lambda_1/\lambda_r \leq \infty$.
- Rank-1 measurement matrix $\mathbf{W}_k = \mathbf{a}_k\mathbf{b}_k^T$ with independent $\mathbf{a}_k$ and $\mathbf{b}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_n)$.
- m measured bits $y_k \in \{+1, -1\}$ ($k = 1, \ldots, m$) satisfy (Plan and Vershynin 2013)

$$\mathbb{E}\left[y_k \middle| \mathbf{W}_k, \mathbf{\Sigma}\right] = \theta(\langle \mathbf{W}_k, \mathbf{\Sigma}\rangle) = \theta(\mathbf{a}_k^T\mathbf{\Sigma}\mathbf{b}_k),$$

- For error-free measurement $\theta(\cdot) = \text{sign}(\cdot)$.

Equivalence Between Rank-1 Measurement and Quadratic Measurement

- Quadratic measurement (Chi 2014a,b):
  - Measurement matrix $\tilde{\mathbf{W}}_k = (\tilde{\mathbf{a}}_k\tilde{\mathbf{a}}_k^T - \tilde{\mathbf{b}}_k\tilde{\mathbf{b}}_k^T)/2$ with independent $\tilde{\mathbf{a}}_k$ and $\tilde{\mathbf{b}}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_n)$.
  - Measured bit y_k satisfies $\mathbb{E}[y_k|\tilde{\mathbf{W}}_k, \mathbf{\Sigma}] = \theta(\langle \tilde{\mathbf{W}}_k, \mathbf{\Sigma}\rangle) = \theta((\tilde{\mathbf{a}}_k^T\mathbf{\Sigma}\tilde{\mathbf{a}}_k - \tilde{\mathbf{b}}_k^T\mathbf{\Sigma}\tilde{\mathbf{b}}_k)/2)$
- Equivalence: $\tilde{\mathbf{a}}_k = \frac{1}{\sqrt{2}}(\mathbf{a}_k + \mathbf{b}_k)$, $\tilde{\mathbf{b}}_k = \frac{1}{\sqrt{2}}(\mathbf{a}_k - \mathbf{b}_k) \Rightarrow \tilde{\mathbf{W}}_k = (\mathbf{W}_k + \mathbf{W}_k^T)/2 \Rightarrow \langle \tilde{\mathbf{W}}_k, \mathbf{\Sigma}\rangle = \langle \mathbf{W}_k, \mathbf{\Sigma}\rangle$.

The PSD Recovery Problem (Penalized Trace Minimization)

$$\max_{\mathbf{X}} -\text{tr}(\mathbf{X}\mathbf{S}_m) + \gamma\text{tr}(\mathbf{X}), \text{ s.t. } \mathbf{X} \succeq 0, \|\mathbf{X}\|_F \leq 1,$$

- The empirical mean $\mathbf{S}_m = \frac{1}{2m}\sum_{k=1}^m y_k(\mathbf{W}_k + \mathbf{W}_k^T) = \mathbf{U}\mathbf{T}\mathbf{U}^T$ with orthogonal matrix $\mathbf{U}$ and diagonal matrix $\mathbf{T}$.
- The closed-form solution

$$\hat{\mathbf{X}} = \begin{cases} \frac{1}{\|\mathcal{D}_\gamma(\mathbf{S}_m)\|_F}\mathcal{D}_\gamma(\mathbf{S}_m), & \text{if } \|\mathbf{S}_m\| > \gamma, \\ 0, & \text{otherwise,} \end{cases}$$

$\mathcal{D}_\gamma(\mathbf{S}_m) = \mathbf{U}(\mathbf{T} - \gamma\mathbf{I}_n)^+\mathbf{U}^T$ is the singular value thresholding operator.

Lemma 1 (Eigendecomposition of mean $\mathbf{S} = \mathbb{E}[y_k\mathbf{W}_k]$)

$$\mathbf{S} = \mathbf{Q}\mathbf{D}\mathbf{Q}^T$$

- $\mathbf{D} = \text{diag}\{\sigma_1, \ldots, \sigma_r, 0, \ldots, 0\}$ with $\sigma_i = \lambda_i\delta_i$.

$$\delta_i = \mathbb{E}\left[\frac{\lambda(\mathbf{z}^T\mathbf{\Lambda}^2\mathbf{z})}{\mathbf{z}^T\mathbf{\Lambda}^2\mathbf{z}}z_i^2\right], \delta_i = \sqrt{\frac{2}{\pi}}\mathbb{E}\left[\frac{z_i^2}{\sqrt{\mathbf{z}^T\mathbf{\Lambda}^2\mathbf{z}}}\right] \text{ (the error-free case)}$$

where $\mathbf{z} = [z_1, \ldots, z_n]^T \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_n)$.

Interlude: Bound for Eigenvalues of S

- Proposition 1: $\frac{2}{\pi} \geq \sigma_i \geq \frac{2}{\sqrt{\pi}\kappa r}\frac{\Gamma(\frac{r+1}{2})}{\Gamma(\frac{r}{2})}$, $i = 1, \ldots, r$.
- Asymptotic lower bound: $\lim_{r\to\infty}\frac{2}{\sqrt{\pi}\kappa r}\frac{\Gamma(\frac{r+1}{2})}{\Gamma(\frac{r}{2})} = \frac{\sqrt{2}}{\kappa\sqrt{\pi}}\frac{1}{\sqrt{r}}$.

Useful Bounds for Recovery Error Analysis (Error-free Measurement)

- Error in mean (Lemma 2 of (Chi 2014b)): $\|\mathbf{S}_m - \mathbf{S}\| \leq \sqrt{\frac{C_1 n}{m}\log\frac{2n}{\delta}}$ with probability at least $1 - \delta$
- Error between $\bar{\mathbf{S}} = \mathbf{S}/\|\mathbf{S}\|_F$ and $\mathbf{\Sigma}$ (Lemma 3): $\|\bar{\mathbf{S}} - \mathbf{\Sigma}\|_F = \sqrt{\sum_{i=1}^r \lambda_i^2(\rho\delta_i - 1)^2} \leq \frac{C_2\kappa^{10}}{r}$ for $r > 5\kappa^5$ where $\rho = 1/\|\mathbf{S}\|_F$
- Deviation of non-zero δ_i from 1 (Lemma 4): $|\delta_i - 1| \leq \frac{5\kappa^5}{r}$

Theorem 1 (Recovery Error Bound for Error-free Measurement)

If the regularization parameter

$$\gamma = 2\sqrt{\frac{C_1 n\log(2n/\delta)}{m}}$$

with probability at least $1 - \delta$, we have for $r > 5\kappa^5$ the following bound on recovery error:

$$\|\hat{\mathbf{X}} - \mathbf{\Sigma}\|_F \leq 3\sqrt{\frac{C_1 nr\log(2n/\delta)}{m}} + \frac{C_2\kappa^{10}}{r}.$$

Proof idea: triangular inequality $\|\hat{\mathbf{X}} - \mathbf{\Sigma}\|_F \leq \|\hat{\mathbf{X}} - \bar{\mathbf{S}}\|_F + \|\bar{\mathbf{S}} - \mathbf{\Sigma}\|_F$.

- The first error term: proof idea of Theorem 1 in (Zhang, Yi, and Jin 2014) with the aid of Lemma 2 and the decomposition of tangent and normal spaces at $\mathbf{\Sigma}$ (Candès and Recht 2009).
- The second error term: Lemma 3.

Numerical Results

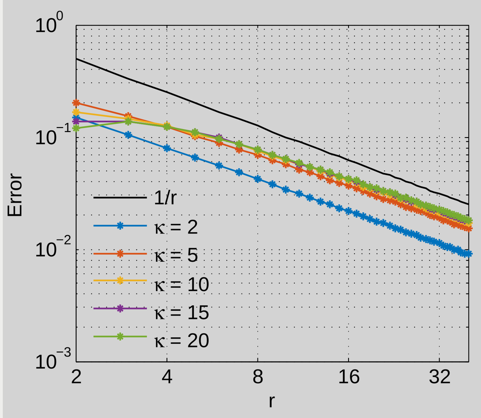


Figure 1: $\mathbb{E}[\|\bar{\mathbf{S}} - \mathbf{\Sigma}\|_F]$ versus r.

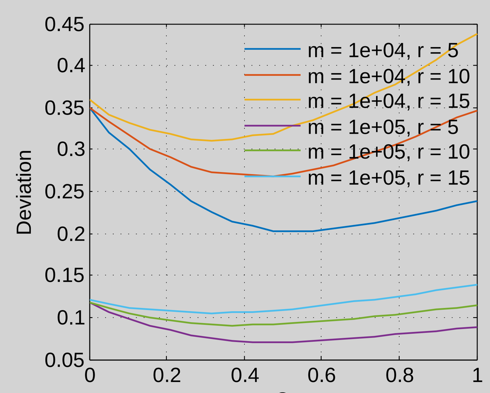
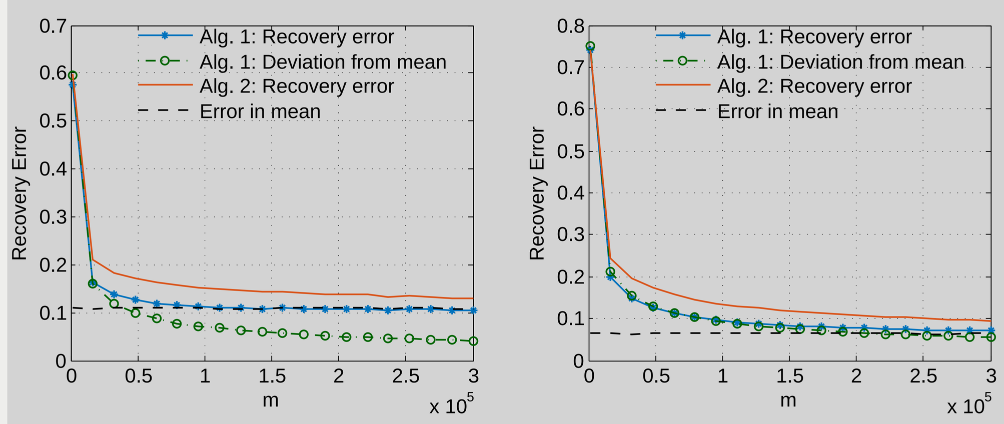
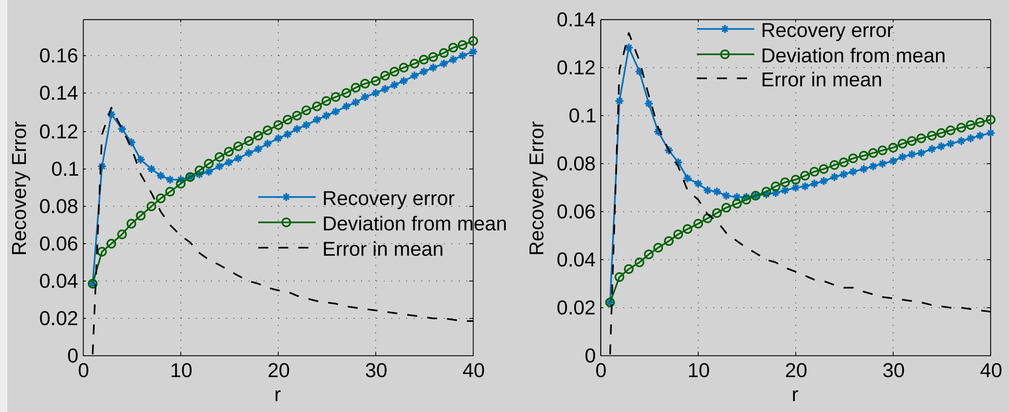


Figure 2: $\mathbb{E}[\|\hat{\mathbf{X}} - \bar{\mathbf{S}}\|_F]$ versus C. ($C^{\text{opt}} \approx 0.5$)



(a) $r = 5$ (b) $r = 10$

Figure 3: Recovery error versus m.



(a) $m = 1 \times 10^5$ (b) $m = 3 \times 10^5$

Figure 4: Recovery error versus r.

- Setup:
  - 1000 trials.
  - Dimension $n = 50$.
  - Eigenvector: orth($\mathbf{X}$) with $\mathbf{X} = [x_{ij}] \in \mathbb{R}^{50\times r}$ and $x_{ij} \sim \mathcal{N}(0, 1)$.
  - Non-zero eigenvalues: $\lambda_i = \tilde{\lambda}_i/\sum_i^r\tilde{\lambda}_i$ with $\tilde{\lambda}_1 = \kappa$, $\tilde{\lambda}_r = 1$, and $\tilde{\lambda}_i \sim \mathcal{U}[1, \kappa]$.
  - $\gamma = C\sqrt{n\log(n)/m}$.
- Algorithm 2:

$$\max_{\mathbf{X}\succeq 0, \|\mathbf{X}\|_F = 1, \|\mathbf{X}\|_* \leq \sqrt{r}} \text{tr}(\mathbf{X}\mathbf{S}_m).$$

- Performance metrics:
  - Recovery error $\|\hat{\mathbf{X}} - \mathbf{\Sigma}\|_F$.
  - Error in mean $\|\bar{\mathbf{S}} - \mathbf{\Sigma}\|_F$.
  - Deviation form mean $\|\hat{\mathbf{X}} - \bar{\mathbf{S}}\|_F$.

Conclusion

- Closed-form solution for PSD matrix recovery problem and more accurate error bound can be obtained with the proposed rank-1 measurement scheme.
- Both analysis and simulation reveal that the recovery is biased but the minimum recovery error is achieved at certainty r.
- Extend the analysis to the case of the noisy measurement and structured PSD matrix (e.g., Toeplitz matrix) recovery in further research.

References

- Candès, E. and B. Recht (2009). Exact Matrix Completion via Convex Optimization. *Foundations of Computational Mathematics*. vol. 9, no. 6, pp. 717–772.
- Chi, Y. (2014a). One-bit principal subspace estimation. *In Proc. IEEE Global Conf. Signal Inf. Process. (GlobalSIP)*. Atlanta, GA, pp. 419–423.
- — (2014b). Subspace Learning From Bits. *arXiv preprint arXiv:1407.6288*.
- Mehanna, O. and N. D. Sidiropoulos (2013). Frugal Sensing: Wideband Power Spectrum Sensing From Few Bits. *IEEE Trans. Signal Processing*. vol. 61, no. 10, pp. 2693–2703.
- Plan, Y. and R. Vershynin (2013). Robust 1-bit Compressed Sensing and Sparse Logistic Regression: A Convex Programming Approach. *IEEE Trans. Inform. Theory*. vol. 59, no. 1, pp. 482–494.
- Zhang, L., J. Yi, and R. Jin (2014). Efficient algorithms for robust one-bit compressive sensing. *In Proc. 31st Int. Conf. on Machine Learning (ICML'14)*. Beijing, China, pp. 820–828.

wangxiyuan@ieee.org