다중 압축 센싱 시스템 분석

박상준 , 이흥노*
*교신저자
광주과학기술원 정보통신공학과
sjpark1@gist.ac.kr, heungno@gist.ac.kr

초록: 이 논문에서, 우리는 다중 압축 센싱 시스템 (Multiple Compressive Sensing System)의 성능 분석에 대해 공부한다. 우리는 가상의 JT 디코더를 정의하고, JT 디코더가 올바르게 동작 하지 않을 사건에 대해 조사하였다. 그 결과, 우리는 JT 디코더가 올바르게 동작 하지 않을 확률의 upper bound 를 획득 하였다. 또한, Upper bound 를 조사하여, JT 디코더가 올바르게 작동하기 위한 충분 조건을 획득 하였다.

주제어: Compressive Sensing, Underdetermined System

Ⅰ. 서론 및 배경

일반적으로, 선형방정식(linear equation) 해의 형태는 행렬의 특징에 따라 결정된다. 예를 들어서, 행렬의 역행렬이 존재하면, 선형방정식 해는 항상 유일하다. 하지만, 우리의 관심사는 행의 개수가 열의 개수보다 더 적은 경우(underdetermined system)이다. 즉,

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

여기서, $\mathbf{A} \in \Re^{M\times N}$, $\mathbf{x} \in \Re^{N\times 1}$, $\mathbf{y} \in \Re^{M\times 1}$, $\|\mathbf{x}\|_0 = K$, 그리고, $M < N$. 일반적으로, (1)의 해는 무수히 많다. 하지만, 만약 $K < M$ (희소한 해)이면, 단 한개의 해만 (1)을 만족시킬 수 있다. 예를 들어, 행렬 $\mathbf{A}$ 의 무작위로 뽑은 M 개의 열들이 독립이고, $M = 2K$ 이면, (1)의 해는 항상 유일하다[1][2]. 이러한 특징으로 인해, 희소한 해를 가지는 underdetermined system에 관한 연구가 압축센싱(Compressive Sensing)이라는 제목하에 활발하게 진행되고 있다.[1][2]

Support Set 탐지는 압축 센싱의 흥미로운 주제중 하나이다. 먼저 Support set의 정의를 소개한다.

정의 1: Support Set $\mathcal{I}$ 는 희소 $\mathbf{x}$ 의 0이 아닌 인덱스들을 원소로 갖는다.

예를 들면, $\mathbf{x} = [1 \quad 0 \quad 2 \quad 0]^{\mathrm{T}}$ 이면, $\mathbf{x}$ 의 $\mathcal{I}=\{1,3\}$ 이다. $\mathcal{I}$ 을 이용한다면, (1)의 해는 $\mathbf{x} = \left(\mathbf{A}[\mathcal{I}]\right)^{\dagger}\mathbf{y}$, 여기서, $\mathbf{A}^{\dagger}$ 는 $\mathbf{A}$ 의 pseudo inverse, 그리고 $\mathbf{A}[\mathcal{I}]$ 는 $\mathbf{A}$ 의 열들중 $\mathcal{I}$ 의 원소들에 대응하는 것들만 모은 $M \times |\mathcal{I}|$ 행렬이다. 즉, $\mathcal{I}$ 를 정확하게 알면 $\mathbf{x}$ 를 정확하게 찾아낼 수 있다. 이제 좀더 실용적인 모델을 생각해보자. 즉

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n} \tag{2}$$

여기서, $\|\mathbf{n}\|_2 \le \varepsilon$ 이다. 우리의 문제는 $\mathbf{A}$ 및 잡음이 포함된 $\mathbf{y}$ 로부터 $\mathcal{I}$ 를 찾거나 혹은 $\mathbf{x}$ 를 찾는 것이다.

관련된 연구 결과는 다음과 같다. [2]에서는 $M = \mathrm{O}\left(K\log\left(N/K\right)\right)$ 이면, QP(quadratic programming)에 의해 얻어진 신호 $\hat{\mathbf{x}}$ 는 stable한 것을 보였다. 즉 $\|\mathbf{x} - \hat{\mathbf{x}}\|_2 \le C\varepsilon$, 여기서 C 는 양의 상수이다. [3]에서는 Lasso (QP의 일종)의 점근적 성능에 대해서 논하였다. 예를 들어서, Wainwright[3]는 Lasso를 이용하면 $\mathcal{I}$ 을 찾기 위한 필요충분조건이 $M/N \to \infty$ 또는 $x_{\min} \to \infty$ ($x_{\min}$ 는 $\mathbf{x}$ 의 0이 아닌 값들중 절대값이 가장 작은 값)을 보였다. 이 결과들은, 특정 알고리즘들에 의존적이다. 다음으로, 우리는 정보이론관점에서 분석한 결과들에 대해 언급하겠다.

편의를 위해, $\Pr\left(\mathcal{I} - Sucess\right)$ 는 $\mathcal{I}$ 를 올바르게 찾을 확률로 표시한다. Fletcher[5]는, $\mathcal{I}$ 를 올바르게 찾을 조건을 $\left(M,N,K\right)$ 에 대한 함수로 제시하였다. Wainwright[4]는 모든 경우를 다 고려하는 알고리즘을 사용 했을 때의, $\mathcal{I}$ 를 올바르게 탐지하기 위한 필요충분조건을 제시하였다. Gastpar[6]에서는, $\left(M,N,K\right)$ 가 점근적으로 증가를 하고, $x_{\min} < \infty$ 이면, $\mathcal{I}$ 를 올바르게 찾는 것이 불가능하다는 것을 보였다. 즉, $\left(M,N,K\right)$ 가 점근적으로 증가를 할 때, $\mathcal{I}$ 를 올바르게 찾기 위한 충분 조건은 $x_{\min} \to \infty$ 이다. 이 결과는, 앞에서 언급한 Wainwright[3]의 결과와 비슷한 결과이다. Akcakaya[7]는 모든 경우를 조사하는 jointly typical decoder를 제시한후, 이것에 의한 $\Pr\left(\mathcal{I} - Sucess\right)$ 의 upper bound를 제시 하였다.

Ⅱ. 신호 모델 및 연구질문

우리는 (1)을 확장시킨 다중 압축 센싱 시스템 (Multiple Compressive Sensing system)을 고려하고자

한다. 즉

$$\mathbf{y}_i = \mathbf{A}_i \mathbf{x}_i + \mathbf{n}_i \tag{3}$$

여기서, $i = 1, \cdots, S$, $n_i(l) \sim \mathcal{N}(0, \sigma^2_{\text{noise}})$, $\|\mathbf{x}_i\| = K$, 그리고 $A_i(l, j) \sim \mathcal{N}(0,1)$. 또한, 각각의 $\mathbf{x}_i$ 는 동일한 $\mathcal{I}$ 을 가진다. 모든 $\mathbf{A}_i$ 와 $\mathbf{n}_j$ 의 원소들은 서로 상호 독립적이고, $\mathbf{n}_j$ 와 $\mathbf{n}_i$ 의 원소들 또한 서로 상호 독립적이다.

우리의 관심사는, 모든 $\mathbf{y}_i$ 와 $\mathbf{A}_i$ 를 알고 있을 때, $\mathcal{I}$ 를 찾는 것이다.

연구질문들은 다음과 같다. 첫 번째로, M, N, K 들을 고정시킨 반면 S 가 커진다면 $\mathcal{I}$ 를 더 잘 찾을 수 있는가? $x^*_{\min} \to \infty$ ($x^*_{\min}$ 는 모든 $\mathbf{x}_i$ 들의 0이 아닌 값들 중 절대값이 가장 작은 값)이면, $\mathcal{I}$ 를 더 잘 찾을 수 있는가? 등 이다.

Ⅲ. JT 디코더 및 실패 확률

$\mathcal{I}$ 을 검출하기 위해, 본 연구팀은 다음과 같은 jointly typical decoder (JT 디코더)를 제안하고자 한다. 이 JT 디코더는 Akcakaya[7]가 제안한 그들의 디코더를 확장한 것이다.

정의 2: 모든 s 에 대하여, $\text{rank}(\mathbf{F}_{s,\mathcal{J}}) = K$, 이고 (4)를 만족시키면, $SM \times 1$ 벡터 $\mathbf{y} := [\mathbf{y}_1^{\text{T}} \ \cdots \ \mathbf{y}_S^{\text{T}}]^{\text{T}}$ 와 집합 $\mathcal{J}$ 는 δ – jointly typical이다.

$$\left| \sum_s \frac{\|\mathbf{Q}(\mathbf{F}_{s,\mathcal{J}})\mathbf{y}_s\|^2}{SM} - \frac{(M-K)\sigma^2_{\text{noise}}}{M} \right| < \delta \tag{4}$$

여기서, $\|\mathbf{Q}(\mathbf{F}_{s,\mathcal{J}})\mathbf{y}_s\|^2$ 는 랜덤변수이다.

정의 3: JT 디코더는 $SM \times 1$ 벡터 $\mathbf{y}$ 와 δ – jointly typical한 집합 $\mathcal{J}$ 를 찾는다.

이제, 우리는 JT 디코더가 $\mathcal{I}$ 의 검출 실패 사건 $\mathrm{E}(D_{failure})$ 에 대해서 이야기한다. 예를 들면, JT 디코더가 $\mathcal{I}$ 와 $\mathbf{y}$ 가 δ – jointly typical 하지 않다고 선언하는 경우 $\mathrm{E}(\mathbf{y}, \mathcal{I}, \delta)^c$ 와, 임의의 집합 $\mathcal{J}$ 와 $\mathbf{y}$ 가 δ – jointly typical하다고 하는 경우 $\mathrm{E}(\mathbf{y}, \mathcal{J}, \delta)$ 를 생각 할 수 있다. 따라서, 검출 실패 사건은 다음과 같다 .

$$\mathrm{E}(D_{failure}) = \mathrm{E}(\mathbf{y}, \mathcal{I}, \delta)^c \bigcup_{\forall \mathcal{J} \neq \mathcal{I}, |\mathcal{J}| = K} \mathrm{E}(\mathbf{y}, \mathcal{J}, \delta). \tag{5}$$

따라서, JT 디코더의 검출 실패 사건이 일어날 확률은 다음과 같다.

$$\Pr\{\mathrm{E}(D_{failure})\} \le \Pr\{\mathrm{E}(\mathbf{y}, \mathcal{I}, \delta)^c\} + \alpha \Pr\{\mathrm{E}(\mathbf{y}, \mathcal{J}^*, \delta)\}. \tag{6}$$

여기서, $\Pr\{\mathrm{E}(\mathbf{y}, \mathcal{J}^*, \delta)\} := \max_{\forall \mathcal{J} \neq \mathcal{I}, |\mathcal{J}| = K} \left[\Pr\{\mathrm{E}(\mathbf{y}, \mathcal{J}, \delta)\} \right]$ 그리고, $\alpha := \binom{N}{K}$ 이다.

Ⅳ. 결과 및 토론

모든 증명들은 [8]에 수록되어있다. 공간상의 문제로 모든 증명들은 생략한다. 다음 아래 기호들을 토론에 유용하게 사용된다.

$$\begin{aligned} p(\mathcal{T}, \mathcal{J}, \lambda_{\min}) &:= \left(\frac{\sigma^2_{\text{noise}}}{\lambda_{\min}} + \frac{M\delta}{(M-K)\lambda_{\min}} \right)^{\frac{S(M-K)}{2}} \\ &\times \exp\left(-\frac{SM}{2\lambda_{\min}} \left(\frac{M-K}{M} (\sigma^2_{\text{noise}} - \lambda_{\min}) + \delta \right) \right), \end{aligned} \tag{7}$$

$$p(\mathcal{T}, \mathcal{I}) := \exp\left(-\frac{SM\delta}{2\sigma^2_{\text{noise}}} \right) \left(1 + \frac{M\delta}{(M-K)\sigma^2_{\text{noise}}} \right)^{\frac{S(M-K)}{2}}. \tag{8}$$

여기서, $\mathcal{T} = \{S, M, K, \delta, \sigma^2_{\text{noise}}\}$, $\lambda_{\min}$ 은 $\mathbf{R} := \mathbb{E}[\mathbf{L}\mathbf{L}^{\text{T}}]$ 의 가장 작은 Eigen Value, $\mathbf{L} := [c_1(1) \ \cdots \ c_S(M-K)]$, $c_s(j) \sim \mathcal{N}\left(0, \sigma^2_{s,\mathcal{J}} := \sigma^2_{\text{noise}} + \sum_{u \in \mathcal{I} \setminus \mathcal{J}} x_s(u)^2\right)$.

Lemma 1: 모든 s 에 대하여, $\text{rank}(\mathbf{F}_{s,\mathcal{J}}) = K$, $M > K$, $\delta > 0$, 그리고 $\mathcal{J}$ 가 잘못된 Support Set이면,

$$\Pr\{\mathrm{E}(\mathbf{y}, \mathcal{J}, \delta)\} \le p(\mathcal{T}, \mathcal{J}, \lambda_{\min}). \tag{9}$$

Lemma 2: 모든 s 에 대하여, $\text{rank}(\mathbf{F}_{s,\mathcal{J}}) = K$, $M > K$, $\delta > 0$, 그리고 $\mathcal{I}$ 가 support 집합이면,

$$\Pr\{\mathrm{E}(\mathbf{y}, \mathcal{I}, \delta)^c\} \le p(\mathcal{T}, \mathcal{I}). \tag{10}$$

Lemma 3: 모든 s 에 대하여, $\text{rank}(\mathbf{F}_{s,\mathcal{J}}) = K$, $M > K$, $\delta > 0$, 그리고 $\mathcal{J}$ 가 잘못된 Support Set이라 하자. 만약 $\lambda_{\min} \to \infty$ 이면, $p(\mathcal{T}, \mathcal{J}, \lambda_{\min}) \to 0$ 이다.

위 3개의 Lemma들은 핵심연구 결과들이다. 다음 아래의 Proposition들과 정리(Theorem)은 Lemma들로부터 얻어진다. 아래 Proposition들에서, $\mathcal{T}^* = \{S = 1, M, K, \delta, \sigma^2_{\text{noise}}\}$ 이다.

Proposition 1: 모든 s 에 대하여, $\text{rank}(\mathbf{F}_{s,\mathcal{J}}) = K$, $M > K$, $\delta > 0$, 그리고 $\mathcal{J}$ 가 잘못된 support 집합이라

하자. 만약 $S \to \infty$ 이면, $\Pr\{\mathrm{E}(\mathbf{r},\mathcal{J},\delta)\}$ 선형적으로 0에 수렴하고, 수렴속도는 $p(\mathcal{T}^*,\mathcal{J},\lambda_{\min})$ 이다.

Proposition 2: 모든 s 에 대하여, $\mathrm{rank}(\mathbf{F}_{s,\mathcal{J}})=K$, $M>K$, $\delta>0$, 그리고 $\mathcal{I}$ 가 support 집합이라 하자. 만약 $S \to \infty$ 이면, $\Pr\{\mathrm{E}(\mathbf{y},\mathcal{I},\delta)^c\}$ 선형적으로 0에 수렴하고, 수렴속도는 $p(\mathcal{T}^*,\mathcal{I})$ 이다.

위 Proposition 1과 2는 S 를 증가시켰을 때, $\Pr\{\mathrm{E}(\mathbf{r},\mathcal{J},\delta)\}$ 와 $\Pr\{\mathrm{E}(\mathbf{y},\mathcal{I},\delta)^c\}$ 가 선형적으로 0으로 수렴한다. 즉 S 가 커지면, JT 디코더의 검출 실패 사건이 일어날 가능성이 줄어드는 것이다. 이는 다음 정리 1이다.

정리 1: 모든 s 에 대하여, $\mathrm{rank}(\mathbf{F}_{s,\mathcal{J}})=K$, $M>K$, 그리고 $\delta>0$ 이다. 만약 $S \to \infty$ 이면, JT 디코더의 검출 실패 사건이 일어날 확률 $\Pr\{\mathrm{E}(D_{failure})\}$ 은 선형적으로 0으로 수렴하고, 수렴 속도는 $\max\left(p(\mathcal{T}^*,\mathcal{I}),p(\mathcal{T}^*,\mathcal{J}^*,\lambda_{\min})\right)$ 이다.

Proposition 3: 모든 $\mathbf{A}_i$ 들이 서로 상호 독립적이라 하자. 만약 $x^*_{\min} \to \infty$ 이면, $\Pr\{\mathrm{E}(\mathbf{y},\mathcal{I},\delta)^c\} \to 0$ 이다.

$x^*_{\min}$ 은 일반적인 경우에는, $\lambda_{\min}$ 과 $x^*_{\min}$ 의 관계를 명확하게 표현하지 않았다. 하지만, 모든 $\mathbf{A}_i$ 들이 서로 상호 독립적이면, $\lambda_{\min}$ 은 행렬 $\mathbf{R}$ 의 대각선 성분중 가장 작은 값 $\min_s\left[\sigma^2_{\text{noise}}+\sum_{u\in\mathcal{I}\setminus\mathcal{J}} x_s(u)^2\right]$ 이다. Proposition 3와 Lemma 3를 조합시키면 다음 아래 정리 2가 나온다.

정리 2: 모든 s 에 대하여, $\mathrm{rank}(\mathbf{F}_{s,\mathcal{J}})=K$, $M>K$, $\delta>0$, 그리고 $\Pr\{\mathrm{E}(\mathbf{y},\mathcal{I},\delta)^c\} \le \alpha$ 라 하자. 만약 $\lambda_{\min} \to \infty$ 이면, $\Pr\{\mathrm{E}(D_{failure})\} \to \alpha$ 이다. 특별히, 모든 $\mathbf{A}_i$ 들이 서로 상호 독립적이고, $x^*_{\min} \to \infty$ 이면, $\Pr\{\mathrm{E}(D_{failure})\} \to \alpha$ 이다.

정리 2의 결과는, 앞에서 언급한, Wainwright[4], Gastpar[6]가 보고한 결과들과 비슷한 형태이다.

V. 결론

이 논문에서, 우리는 multiple underdetermined system에 대해서 공부하였다. Multiple underdetermined system을 분석하기 위해서, 가상의 JT 디코더를 정의하였다. JT 디코더는 모든 $\mathbf{A}_i$ 와 $\mathbf{y}_i$ 를 알고 있을 때, $SM\times 1$ 벡터 $\mathbf{y}$ 와 $\delta-$jointly typical 해지는 집합 $\mathcal{J}$ 를 찾는다. 우리는 JT 디코더가 support 집합 $\mathcal{I}$ 를 올바르게 찾지 못할 확률의 upper bound를 제시하고, S 가 커짐에 따라 제시된 upper bound가 0으로 수렴하는 것을 보였다. 또한 $x^*_{\min} \to \infty$ 일때, upper bound가 특정 값으로 수렴하는 것을 증명하였다.

다음 연구주제로, 우리는 모든 $\mathbf{A}_i$ 들이 서로 상호독립적인 경우와, 그렇지 않은 경우의 JT 디코더의 성능에 대해 연구하고자 한다.

Acknowledgement

이 논문은 2012년도 정부 (교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (중견연구자-도약연구사업, N0. 2012-0005656)

참고문헌

[1] D. Donoho, “ Compressive Sensing,” *IEEE Trans. Inform. Theory*, 2006.

[2] E. Candes, J. Romberg, and T. Tao, “ Stable Signal Recovery from Incomplete and Inaccurate Measurements,” *Comm. On Pure and Applied Math.*, 2006.

[3] M. Wainwright, “ Sharp thresholds for high-dimensional and noisy recovery of sparsity,” *Allerton*, 2006

[4] M. Wainwright, “ Information-theoretic bounds on sparsity recovery in the high dimensional and noisy setting,” *ISIT*, 2007

[5] A. K. Fletcher, S. Rangan, and V. K. Goyal, “ Necessary and sufficient conditions on sparsity pattern recovery,” *IEEE Trans. Inform. Theory*, 2009.

[6] Galen Reeves and Michael Gastpar, “ Sampling Bounds for Sparse Support Recovery in the Presence of Noise,” *ISIT*, 2008.

[7] Mehmet Akcakaya and Vahid Tarokh, “ Shannon-Theoretic Limits on Noisy Compressive Sampling,” *IEEE Trans. Inform. Theory.*

[8] Sangjun Park and Heung-No Lee, “ Per-Sensor Measurements required for a Noisy Distributed Compressive Sensor System,” In preparation.