An Analysis for Density Evolution on Noisy Sparse Recovery via Belief Propagation

Jaewook Kang, Heung-No Lee, Kiseon Kim
Gwangju Institute of Science and Technology (GIST)
{jwkkang,heungno,kskim}@gist.ac.kr

Abstract

In this paper, an analysis for density evolution of a BP-based sparse recovery algorithm is provided. We model the signal message used in the BP algorithm using a family of spike-and-slab densities characterized by three parameters. By applying Binomial expansion and Gaussian approximation, we derive an analytical update rule of the parameters over the BP-iteration.

Ⅰ. Introduction

Recently, belief propagation (BP)-based sparse recovery algorithms have received attention [1],[2]. In conjunction with sparse measurement matrices, the BP-approach enables fast measurements generation as well as low complexity signal recovery.

Generally, BP iteratively approximates the marginal posterior corresponding to each signal element by exchanging probability messages over a bipartite graph associated with the measurement matrix. Hence, the marginal posterior seems to evolve from the prior density and converge to a certain density during the iteration. We call such behavior *density evolution* by borrowing the term from the coding theory [3].

Our aim is to analyze the behavior of a BP-based recovery algorithm proposed in [2] called BHT-BP, using the idea of the density evolution. BHT-BP detects the support set of the target signal based on the marginal posterior obtained from BP. Therefore, obtaining the analytical expression of the marginal posterior is highly significant work as the first step of the behavior analysis.

In this paper, we model the signal message in BP using a family of spike-and-slab densities, which is fully characterized by three parameters: μ,σ,q. Then, we derive an analytical update rule of the parameters and demonstrate that the parameters are converged correctly as the iteration proceeded.

Ⅱ. Signal Model

Let $\mathbf{x}\in\mathbb{R}^N$ denote a random K-sparse vector with a state vector $\mathbf{s(x)}$ indicating the support set of $\mathbf{x}$; therefore $\|\mathbf{s(x)}\|_0=K$. Each element $s_i\in\mathbf{s(x)}$ is defined as

$$s_i=\begin{cases}1, & \text{if } x_i\neq 0\\ 0, & \text{else}\end{cases}\ \text{forall}\, i\in\{1,...,N\}. \quad (1)$$

We limit our discussion to the random vector $\mathbf{X}$ whose elements are i.i.d. random variables. Then, the decoder observes a measurement vector $\mathbf{z}\in\mathbb{R}^M$, given as $\mathbf{z}=\mathbf{\Phi x}_0+\mathbf{n}$, where $\mathbf{x}_0\in\mathbb{R}^N$ is a deterministic realization of $\mathbf{x}$; $\mathbf{n}\sim\mathcal{N}(0,\sigma_n^2\mathbf{I}_M)$ is an additive Gaussian noise vector. For the measurement matrix $\mathbf{\Phi}$, we employ sparse-binary matrices $\mathbf{\Phi}\in\{0,1\}^{M\times N}$ with rank $(\mathbf{\Phi})\leq M$ and $M\leq N$. In addition, we specify the sparsity of $\mathbf{\Phi}$ by adjusting the column weight to L such that $\|\phi_{jth\,col}\|_2^2=L$.

Ⅲ. Analysis of Density Evolution of Marginal Posterior

It is known that partially recovered support owing to noise effect is a subset of the true support in terms of l_1-based convex solvers [4]. Such a fact is also observed under the use of BHT-BP, and this indicates that the misdetection happened by noise occurs when $s_i(\hat{x}_i)=0$ given $s_i(x_i)=1$ in BHT-BP. Therefore, the analysis of the marginal posteriors can be confined to the case of nonzero signal elements.

Let $x_i\in\mathbf{x}$ denote a nonzero element. With respect to $\mathbf{\Phi}$, x_i has L neighbored measurements $\{z_{j_1},...,z_{j_L}\}$, and each measurement has $R-1$ neighbored signal elements $\{x_{k_1},...,x_{k_{R-1}}\}$ except x_i, where clearly $R\sim Binomial\left(N,\frac{L}{M}\right)$ given the fixed L. In addition, let α denote the number of the connection to the nonzero elements except x_i from z_{j_1}. The corresponding graph representation is shown in Fig. 1. For convenience, we suppose that the first α elements correspond to the nonzeros and the remaining $R-1-\alpha$ correspond to zeros, that is,

$$\begin{aligned}&x_{k_n}=x_{min}\ \forall n\in\{1,...,\alpha\}\\ &x_{k_n}=0\quad \forall n\in\{\alpha+1,...,R-1\},\end{aligned} \quad (2)$$

where we limit all nonzero values to the minimum signal value $x_{\min}$ to consider the worst case. Let $\mathbf{a}_{k\to j_1}$ denote a signal message from $\{x_{k_1},...,x_{k_{R-1}}\}$ to z_{j_1} in BP. Given the support knowledge, we model the signal message $\mathbf{a}_{k\to j}$ via a family of spike-and-slab densities, respectively:

$$\begin{aligned}&\mathbf{a}_{k\to j_1}(x\,|\,s_k=1)\equiv q_{nz}\mathcal{N}(x;\mu_{nz},\sigma_a^2)+(1-q_{nz})\delta_0\\ &\mathbf{a}_{k\to j_1}(x\,|\,s_k=0)\equiv q_z\mathcal{N}(x;\mu_z,\sigma_a^2)+(1-q_z)\delta_0,\end{aligned} \quad (3)$$

where $q_{nz},q_z\in(0,1]$ and δ_x denotes the delta function having the peak at x. We assume that $\mathbf{a}_{k\to j_1}(x\,|\,s_k=1)$ and $\mathbf{a}_{k\to j_1}(x\,|\,s_k=0)$ are identically

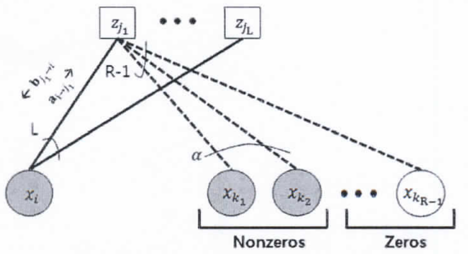

Fig. 1 Graph representation for the analysis

distributed, respectively since the signal value is x_{min} or zero, the noise elements are i.i.d., and the column weight is fixed to L for all signal elements.

From (13) in [2] and using (3), the measurement messages $\mathbf{b}_{j_1\to i}$ is expressed as

$$\mathbf{b}^l_{j_1\to i}(x) = \mathcal{N}(x; z_{j_1}, \sigma_n^2) \otimes \left(\mathbf{a}^l(-x \mid s_k = 1)\right)^{\otimes\alpha} \otimes \left(\mathbf{a}^l(-x \mid s_k = 0)\right)^{\otimes(R-1-\alpha)}. \quad (4)$$

Since the convolution operation ⊗ satisfies the distributive law and the commutative law, $\mathbf{a}^l(-x \mid s_k = 1)^{\otimes\alpha}$ and $\mathbf{a}^l(-x \mid s_k = 0)^{\otimes\alpha}$ in (4) can be represented in the form of Binomial expansion, that is,

$$\mathbf{a}^l(-x \mid s_k = 1)^{\otimes\alpha} = \sum_{m=0}^{\alpha} \binom{\alpha}{m} (1-q_{nz})^m q_{nz}^{(\alpha-m)} \mathcal{N}(-x, \mu_{nz}, \sigma_a^2)^{\otimes(\alpha-m)} \delta_x^{\otimes\alpha}, \quad (5)$$
$$\mathbf{a}^l(-x \mid s_k = 0)^{\otimes R-1-\alpha} = \sum_{m=0}^{R-1-\alpha} \binom{R-1-\alpha}{m} (1-q_z)^m q_z^{(R-1-\alpha)-m} \mathcal{N}(-x, \mu_z, \sigma_a^2)^{\otimes((R-1-\alpha)-m)} \delta_x^{\otimes\alpha}.$$

Under the assumption that the recovery is converged, we postulate that $q_{nz} \to 1$ and $q_z \to 0$ as $l \to \infty$. Then, the first order approximation to (5) is tightly fit, that is,

$$\mathbf{a}^l(-x \mid s_k = 1)^{\otimes\alpha} \approx (q_{nz})^\alpha \times \mathcal{N}(-x; \alpha\mu_{nz}, \alpha\sigma_a^2),$$
$$\mathbf{a}^l(-x \mid s_k = 0)^{\otimes(R-1-\alpha)} \approx (1-q_z)^{(R-1-\alpha)} \delta_x, \quad (6)$$

where we use the fact that *sum of Gaussian is Gaussian*. By substituting (6) to (4), we obtain a Gaussian approximation of $\mathbf{b}^l_{j_1\to i}$ such that

$$\mathbf{b}^l_{j_1\to i}(x) \approx c_1 \times \mathcal{N}(x; z_{j_1} - \alpha\mu_{nz}, \alpha\sigma_a^2 + \sigma_n^2), \quad (7)$$

where $c_1 = (q_{nz})^\alpha (1-q_z)^{(R-1-\alpha)}$.

Then, using (12) in [2] and (7), the signal message for the next iteration is approximately calculated as

$$\mathbf{a}^{l+1}_{i\to j_1}(x \mid s_i = 1) \approx \eta\left[f_x(x) \times \prod_{n=2}^{L} \mathcal{N}(x; z_{j_n} - \alpha\mu_{nz}, \alpha\sigma_a^2 + \sigma_n^2) \right], \quad (8)$$

where we can cancel the constant c_1 for each $\mathbf{b}^l_{j_1\to i}$ with the normalization function $\eta[\cdot]$, and $f_x(x)$ is the prior density given in (6) of [2]. It is well known that the product of Gaussian density functions will result in a scaled Gaussian density function, that is, $\mathcal{N}(x; \mu_1, \sigma_1^2) \times \mathcal{N}(x; \mu_2, \sigma_2^2) \propto \mathcal{N}(x; b, B)$, where $B = \frac{\sigma_1^2\sigma_2^2}{\sigma_1^2+\sigma_2^2}$, $b = \frac{\mu_1\sigma_2^2 + \mu_2\sigma_1^2}{\sigma_1^2+\sigma_2^2}$. Since $x_i = x_{\min}\ \forall i \in \{1,...,N\} : s_i = 1$, the value of z_{j_n} is given as $z_{j_n} = (\alpha+1)x_{\min} + n_{j_n}\ \forall n \in \{1,...,L\}$. Using these facts, the expression in (8) is rewritten as

$$\mathbf{a}^{l+1}_{i\to j_1}(x) \approx \eta\left[q \times c_3 \mathcal{N}\left(x; \tfrac{\mu_b\sigma_x^2}{\sigma_x^2+\sigma_b^2}, \tfrac{\sigma_x^2\sigma_b^2}{\sigma_x^2\sigma_b^2}\right) + \tfrac{1-q}{\sqrt{2\pi\sigma_b^2}} \exp\left[-\tfrac{\mu_b^2}{2\sigma_b^2}\right] \delta_x \right], \quad (9)$$

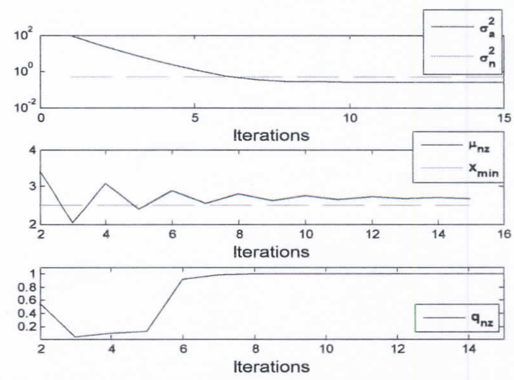

Fig. 2 Behavior of density evolution in the BP-iteration when $q = 0.1$, $\alpha = 0.7$, $x_{\min} = 2.5$, $\sigma_n^2 = 0.48$

where $\mu_b = \alpha(x_{min} - \mu_{nz}) + x_{min} + \frac{1}{L-1}\sum_{n=2}^{L} n_{j_n}$ and $\sigma_b = \sqrt{\frac{\alpha\sigma_a^2 + \sigma_n^2}{L-1}}$, and c_3 is given as $c_3 = \exp\left[\frac{\mu_b^2\sigma_b^2}{2(\sigma_x^2+\sigma_b^2)}\right] / \sqrt{2\pi(\sigma_x^2+\sigma_b^2)}$.

From (9), we can track the parameters set $(\mu_{nz}, \sigma_a, q_{nz})$ of the signal message $\mathbf{a}_{i\to j_1}$, using the following update rules:

$$(\sigma_a^2)^{l+1} = \frac{\sigma_x^2\sigma_b^2}{\sigma_x^2+\sigma_b^2} = \frac{\alpha(\sigma_a^2)^l\sigma_x^2 + \sigma_n^2\sigma_x^2}{\alpha(\sigma_a^2)^l + (L-1)\sigma_x^2 + \sigma_n^2},$$
$$(\mu_{nz})^{l+1} = \frac{\mu_b\sigma_x^2}{\sigma_x^2+\sigma_b^2} = \frac{\alpha(x_{min}\sigma_x^2 - (\mu_{nz})^l\sigma_x^2) + x_{\min}\sigma_x^2 + \frac{1}{L-1}\left(\sum_{n=2}^{L} n_{j_n}\right)\sigma_x^2}{\sigma_x^2 + \frac{\alpha(\sigma_a^2)^l + \sigma_n^2}{L-1}}, \quad (10)$$
$$(q_{nz})^{l+1} = \frac{c_3 q}{c_3 q + \frac{1-q}{\sqrt{2\pi\sigma_b^2}}\exp\left[-\frac{\mu_b^2}{2\sigma_b^2}\right]}.$$

IV. Discussion

In this paper, we analyzed the density evolution behavior in the BP-iteration of BHT-BP. Fig.2 shows that each parameter describing the signal message is converged correctly, according to (10). By applying results in (7),(10) to (14) in [2], we will finally obtain the marginal posterior of a signal element x_i. The remaining step is to consider about α in (10), which a random variable is related to R. We will expand the result of this paper including α in future.

ACKNOWLEDGMENT

This work was supported by National Research Foundation of Korea grant (Haek-Sim Research Program, NO. 2011-0027682, Do-Yak Research Program, NO. 2012-0005656)

References

[1] D. Baron *et al.*, Bayesian compressive sensing via belief propagation," *IEEE Trans. Signal Process.*, vol. 58, no. 1, pp. 269-280, Jan. 2010.

[2] Jaewook Kang *et al.*, Bayesian hypothesis test for sparse support recovery using belief propagation," accepted at *IEEE Statistical Signal Processing Workshop (SSP)*, Ann Arbor, USA, Aug. 2012.

[3] T. Richardson *et al.*, " The capacity of low-density parity check codes under message-passing decoding," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 599-618, Feb. 2001.

[4] D. L. Donoho *et al.*, " Stable recovery of sparse overcomplete representations in the presence of noise," *IEEE Trans. Inf. Theory*, vol. 52, no. 1, pp. 6-18, Jan. 2006.