

# 基于拉伸因子图的低复杂度贝叶斯稀疏信号算法研究

卞孝丽

(郑州财税金融职业学院 郑州 450048)

**摘要** 建立加性高斯白噪声的线性数学模型,针对此模型对基于稀疏贝叶斯学习的消息传递算法进行研究。对传统的因子图通过添加额外的硬约束节点得到改进的因子图,然后在改进的因子图中利用联合 BP-MF 规则,提出低复杂度的 BP-MF SBL 算法。为了进一步降低复杂度,在 BP-MF SBL 的基础上提出近似 BP-MF SBL 算法。仿真结果表明与向量形式的 MF 算法相比,所提方法复杂度低,且性能有所提升;与标量形式的 MF 算法相比,在复杂度相似的情况下,所提方法的性能更好。

**关键词** 加性高斯白噪声,稀疏贝叶斯学习,拉伸因子图,低复杂度,BP-MF SBL 算法

**中图分类号** TP301 **文献标识码** A

## Low Complexity Bayesian Sparse Signal Algorithm Based on Stretched Factor Graph

BIAN Xiao-li

(Zhengzhou Vocational College of Finance and Taxation, Zhengzhou 450048, China)

**Abstract** The linear mathematical model of additive Gauss white noise was established, and the message passing algorithm based on Sparse Bayesian learning was studied in this model. In this work, we modified the factor graph by adding some extra hard constraints which enables the use of combined belief propagation (BP) and MF message passing. This paper proposed a low complexity BP-MF SBL algorithm, based on which an approximate BP-MF SBL algorithm was also developed to further reduce the complexity. The BP-MF SBL algorithms show their merits compared with state-of-the-art MF SBL algorithms. They deliver even better performance with much lower complexity compared with the vector-form MF SBL algorithm and they significantly outperform the scalar-form MF SBL algorithm with similar complexity.

**Keywords** Additive Gauss white noise, Sparse Bayesian learning, Stretched factor graph, Low complexity, BP-MF SBL algorithm

## 1 引言

近年来,压缩感知<sup>[1]</sup>备受社会关注,并且已被应用于各个工程领域,如磁场成像、稀疏信道估计、频谱估计、图像重建<sup>[2]</sup>。在压缩感知中,通常根据下列数学模型中的观测向量  $\mathbf{y} \in \mathbb{C}^{N \times 1}$  估计稀疏向量  $\mathbf{a} \in \mathbb{C}^{L \times 1}$ 。

$$\mathbf{y} = \Phi \mathbf{a} + \boldsymbol{\omega} \quad (1)$$

在此线性模型中,  $\Phi \in \mathbb{C}^{N \times L}$  被称为字典矩阵;  $\boldsymbol{\omega}$  表示均值为 0、协方差矩阵为  $\lambda^{-1} \mathbf{I}$  的高斯白噪声向量,其中  $\lambda$  未知。

除了贪婪算法<sup>[3]</sup>和凸优化<sup>[4]</sup>算法,稀疏贝叶斯学习(SBL)<sup>[5-7]</sup>是另一种可供选择的稀疏信号估计方法,通过给定待估计向量  $\mathbf{a}$  的先验概率  $p(\mathbf{a})$ ,得到  $\mathbf{a}$  的最大后验概率(MAP)估计  $\hat{\mathbf{a}} = \arg \max_{\mathbf{a}} p(\mathbf{a} | \mathbf{y})$ 。为了使  $p(\mathbf{a})$  具有稀疏特性,稀疏贝叶斯学习(SBL)通常利用 3-L 贝叶斯分层模型<sup>[8]</sup>  $p(\mathbf{a}) = \int_{\gamma} p(\mathbf{a} | \gamma) p(\gamma) d\gamma$  来代替  $p(\mathbf{a})$  求  $\mathbf{a}$  的最大后验概率估计。近年来, SBL 能够有效地利用置信传播(BP)<sup>[9-10]</sup>近似消息传递(AMP)<sup>[11-12]</sup>,然而, BP 和 AMP 这两种方法都假设  $\lambda$  是已知的,这使其受限于实际应用。本文讨论了  $\lambda$  未知的情况下的 SBL 消息传递算法。平均场(MF)消息传递算法<sup>[13-15]</sup>,又称为变分消息传递(VMP),被广泛用于近似贝叶

斯推理,尤其是用于具有指数型分布函数的 2-L 及 3-L 分层模型结构。Pedersen 提出了 MF SBL 算法( $\lambda$  未知)<sup>[16]</sup>,并将其用于 OFDM 系统稀疏信道估计中。由于 MF SBL 算法在每次迭代估计向量  $\mathbf{a}$  时都涉及矩阵求逆的问题,导致其具有很高的计算复杂度。为了解决这个问题,文献<sup>[13]</sup>提出一种低复杂度 MF SBL 算法,通过将大矩阵转换为小矩阵来降低矩阵运算复杂度,但这是以牺牲算法性能为代价的。很明显,如果转换为小矩阵的维度等于 1,那么将不会出现矩阵求逆,我们将这种算法称为标量形式的 MF SBL 算法。近年来,标量形式的 MF SBL 算法被用于信道增益和时延估计<sup>[17]</sup>。文献<sup>[5]</sup>提出先验  $p(\mathbf{a})$  的分层模型,文献<sup>[16]</sup>具体分析了 2-L 和 3-L 分层先验模型的特征。

在 MF 算法中,与函数相连的各个变量节点的置信(belief)都是互相独立的。而对于 BP 算法而言,与函数节点相邻的所有变量节点的置信是联合非独立的。这使得 BP 算法能够有效地处理离散概率模型和线性高斯模型。然而, BP 算法在处理既有离散和连续变量的线性模型时,会有很高的复杂度。针对此问题,文献<sup>[18]</sup>提出一种归一化消息传递算法框架,并提出了联合 BP-MF 算法,避免了 BP 算法和 MF 算法的缺点。

本文提出基于 2-L 分层先验模型的低复杂度 BP-MF

SBL算法。此算法是针对改进因子图(通过在传统的因子图上添加一系列硬约束因子而获得)提出的,看似改变“多余”,实则提升了算法性能。由于噪声方差 $\lambda$ 未知, MF SBL算法可以用于处理指数型函数节点, BP SBL算法可以用于处理硬约束因子。由于待估计的向量 $\alpha$ 分解为标量形式,基于改进因子图的BP-MF SBL算法避免了矩阵求逆,大大降低了算法复杂度。受文献[19]中广义近似消息传递(GAMP)的启发,通过GAMP近似BP可以得到复杂度更低的近似BP-MF SBL算法。仿真结果显示,所提出的BP-MF SBL算法在复杂度方面远小于向量形式的MF SBL算法[16],在MSE性能方面要优于向量形式的MF SBL算法。在复杂度基本相等的情况下, MSE性能接近于标量形式的MF SBL算法[17,20]。

函数 $g(x)$ 的期望可以表示为 $\langle f(x) \rangle_{g(x)} = \int f(x)g(x) dx / \int g(x')dx'$ , 其方差为 $\text{var}[x]_{g(x)} = \langle |x|^2 \rangle_{g(x)} - |\langle x \rangle_{g(x)}|^2$ 。 $\mathcal{CN}(x; \mu, \nu)$ 表示均值为 $\mu$ 、方差为 $\nu$ 的复高斯分布。 $f(x) = cg(x)$ 表示为 $f(x) \propto g(x)$ , 其中 $c$ 为常数。

## 2 图模型

$\alpha, \gamma, \lambda$ 的联合后验PDF的2-L分层先验模型[8]的因子分解为:

$$p(\alpha, \gamma, \lambda | \mathbf{y}) \propto f_\lambda(\lambda) \prod_n f_{y_n}(\alpha, \lambda) \prod_l f_{\alpha_l}(\alpha_l, \gamma_l) f_{\gamma_l}(\gamma_l) \quad (2)$$

其中,  $f_{y_n}(\alpha, \lambda) \triangleq p(y_n | \alpha, \lambda) = \mathcal{CN}(y_n; \Phi_n \alpha, \lambda^{-1})$ ,  $\Phi_n$ 表示矩阵 $\Phi$ 的第 $n$ 行向量。 $f_\lambda(\lambda)$ 表示噪声精准系数 $\lambda$ 的先验。函数 $f_{\alpha_l}(\alpha_l, \gamma_l)$ 表示条件概率 $p(\alpha_l | \gamma_l) = \mathcal{CN}(\alpha_l; 0, \gamma_l^{-1})$ , 其服从高斯分布。 $f_{\gamma_l}(\gamma_l)$ 表示超先验分布, 超先验参数 $\gamma_l$ 服从伽马分布 $p(\gamma_l) = \text{Ga}(\gamma_l; \varepsilon, \eta)$ 。

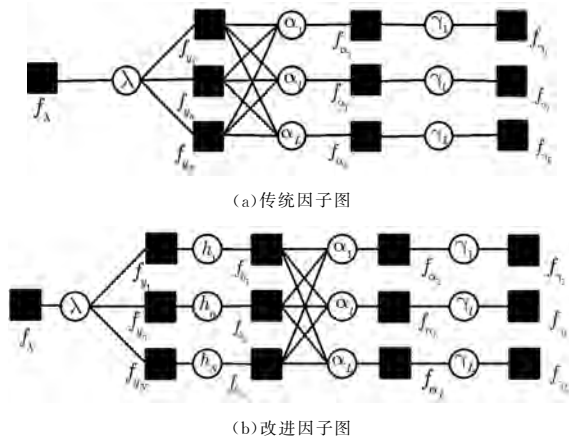


图1

式(2)的因子图[21]如图1(a)所示,其与文献[17,20]中的图模型一致。假设 $\lambda$ 未知,并且MF可以被用于处理函数节点 $\{f_{y_n}, \forall n \in [1:N]\}$ 。这与文献[20]中标量形式的MF SBL算法一致。文献[16]基于传统的因子图推导得出了向量形式的MF SBL算法,被估计变量 $\alpha$ 被视为单一的变量节点。

为了能够使用MF和BP消息传递规则,对图1(a)进行改进,通过增加硬约束函数节点 $\{f_{\delta_n}(h_n, \mathbf{a}) = \delta(h_n - \Phi_n \mathbf{a}), \forall n \in [1:N]\}$ 和变量节点 $\mathbf{h} = \Phi \alpha$ ,同时改变函数 $f_{y_n}$ 的分布 $p(y_n | h_n, \lambda) = \mathcal{CN}(y_n; h_n, \lambda^{-1})$ ,得到改进后的因子图模型,如图1(b)所示。在改进的因子图中,对固定的指数型函数节点部分使用MF消息传递规则更新消息;而在硬约束部分,使用

BP消息规则更新消息,以确保获得更好的性能。下面给出消息更新的流程及步骤。

## 3 BP-MF SBL

本节根据联合BP-MF消息更新规则[18],在图1(b)中给出详细的稀疏信号估计消息更新及迭代推导过程。图1(b)中所有变量集合为 $A$ ,其被分为两个不相关的子集:BP子集 $\mathcal{A}_{BP} = \{f_{\delta_n}, \forall n\}$ 和MF子集 $\mathcal{A}_{MF} = A \setminus \mathcal{A}_{BP}$ 。

### 3.1 消息计算

消息更新计算可分为前向(从左向右)和后向(从右向左)两部分。在更新前向消息时用到的相关的后向消息由上一次迭代获得。

#### 3.1.1 前向消息计算

假设噪声精准系数 $\lambda$ 的置信 $b(\lambda)$ 已知,利用MF规则计算从观测变量 $f_{y_n} \in \mathcal{A}_{MF}$ 到 $h_n$ 的消息 $m_{f_{y_n} \rightarrow h_n}(h_n)$ 。

$$m_{f_{y_n} \rightarrow h_n}(h_n) = \exp\{\langle \log f_{y_n}(h_n, \lambda) \rangle_{b(\lambda)}\} \propto \mathcal{CN}(h_n; y_n, \lambda^{-1}) \quad (3)$$

$$\lambda = \langle \lambda \rangle_{b(\lambda)}$$

其中,利用BP规则计算消息 $m_{f_{\delta_n} \rightarrow \alpha_l}(\alpha_l)$ ,  $n_{h_n \rightarrow f_{y_n}}(h_n) = m_{f_{y_n} \rightarrow h_n}(h_n)$ 以及 $\{n_{\alpha_l \rightarrow \gamma_l}(\alpha_l'), \forall l' \neq l\}$ ,得到:

$$m_{f_{\delta_n} \rightarrow \alpha_l}(\alpha_l) = \langle f_{\delta_n}(h_n, \mathbf{a}) \rangle_{n_{h_n \rightarrow f_{y_n}}(h_n) \prod_{l' \neq l} n_{\alpha_l \rightarrow f_{\delta_n}}(\alpha_l')} \propto \mathcal{CN}(\alpha_l; \hat{\alpha}_{n \rightarrow l}, \nu_{\alpha_{n \rightarrow l}}) \quad (4)$$

其中:

$$\hat{\alpha}_{n \rightarrow l} \triangleq \frac{y_n - \hat{p}_n + \Phi_{nl} \alpha_{l \rightarrow n}}{\Phi_{nl}} \quad (5)$$

$$\nu_{\alpha_{n \rightarrow l}} \triangleq \frac{\lambda^{-1} + \nu_{p_n} - |\Phi_{nl}|^2 \nu_{\alpha_{l \rightarrow n}}}{|\Phi_{nl}|^2} \quad (6)$$

$$\hat{p}_n \triangleq \sum_l \Phi_{nl} \alpha_{l \rightarrow n} \quad (7)$$

$$\nu_{p_n} \triangleq \sum_l |\Phi_{nl}|^2 \nu_{\alpha_{l \rightarrow n}} \quad (8)$$

为描述方便,所有高斯消息的乘积形式由下式表示:

$$q_l(\alpha_l) = \prod_n m_{f_{\delta_n} \rightarrow \alpha_l}(\alpha_l) \propto \mathcal{CN}(\alpha_l; q_l, \nu_{q_l}) \quad (9)$$

其中:

$$\nu_{q_l} \triangleq \left( \sum_n \frac{1}{\nu_{\alpha_{n \rightarrow l}}} \right)^{-1} \quad (10)$$

$$q_l \triangleq \nu_{q_l} \left( \sum_n \frac{\hat{\alpha}_{n \rightarrow l}}{\nu_{\alpha_{n \rightarrow l}}} \right) \quad (11)$$

给定消息 $m_{f_{y_n} \rightarrow \alpha_l}(\alpha_l) \propto \mathcal{CN}(\alpha_l; 0, \gamma_l^{-1})$ ,变量 $\alpha_l$ 的置信 $b(\alpha_l)$ 可表示为:

$$b(\alpha_l) \propto q_l(\alpha_l) m_{f_{y_n} \rightarrow \alpha_l}(\alpha_l) \propto \mathcal{CN}(\alpha_l; \hat{\alpha}_l, \nu_{\alpha_l}) \quad (12)$$

其中:

$$\hat{\alpha}_l \triangleq \frac{q_l}{1 + \nu_{q_l} \gamma_l} \quad (13)$$

$$\nu_{\alpha_l} \triangleq \left( \frac{1}{\nu_{q_l}} + \gamma_l \right)^{-1} \quad (14)$$

由于 $f_{\delta_n}$ 被归为MF子集,利用MF消息更新规则计算消息 $m_{f_{\delta_n} \rightarrow \gamma_l}(\gamma_l)$ :

$$m_{f_{\delta_n} \rightarrow \gamma_l}(\gamma_l) = \exp\{\langle \log f_{\delta_n}(\alpha_l, \gamma_l) \rangle_{b(\alpha_l)}\} \propto \gamma_l \exp\{-\gamma_l (|\hat{\alpha}_l|^2 + \nu_{\alpha_l})\} \quad (15)$$

超先验参数 $\gamma_l$ 的置信 $b(\gamma_l)$ 可表示为 $b(\gamma_l) \propto m_{f_{\delta_n} \rightarrow \gamma_l}(\gamma_l)$

$$f_{\gamma_l}(\gamma_l) \propto \gamma_l^{\epsilon+1} \exp\{-\gamma_l(\eta + |\hat{\alpha}_l|^2) + \nu_{\alpha_l}\}.$$

### 3.1.2 后向消息计算

首先利用 MF 准则计算从  $f_{\alpha_l}$  到  $\alpha_l$  的消息  $m_{f_{\alpha_l} \rightarrow \alpha_l}(\alpha_l)$ :

$$m_{f_{\alpha_l} \rightarrow \alpha_l}(\alpha_l) = \exp\{\langle \log f_{\alpha_l}(\alpha_l, \gamma_l) \rangle_{b(\gamma_l)}\} \propto \mathcal{CN}(\alpha_l; 0, \hat{\gamma}_l^{-1}) \quad (16)$$

其中:

$$\hat{\gamma}_l = \langle \gamma_l \rangle_{b(\gamma_l)} = \frac{\epsilon + 1}{\eta + |\hat{\alpha}_l|^2 + \nu_{\alpha_l}} \quad (17)$$

由于函数节点  $f_{\delta_n} \in \mathcal{A}_{BP}$ , 将利用 BP 规则更新消息  $n_{\alpha_l \rightarrow f_{\delta_n}}(\alpha_l)$ :

$$n_{\alpha_l \rightarrow f_{\delta_n}}(\alpha_l) = \frac{b(\alpha_l)}{m_{f_{\delta_n} \rightarrow \alpha_l}(\alpha_l)} \propto \mathcal{CN}(\alpha_l; \hat{\alpha}_{l \rightarrow n}, \nu_{\alpha_l \rightarrow n}) \quad (18)$$

其中:

$$\nu_{\alpha_l \rightarrow n} \triangleq \left( \frac{1}{\nu_{\alpha_l}} - \frac{1}{\nu_{\alpha_l \rightarrow l}} \right)^{-1} \quad (19)$$

$$\hat{\alpha}_{l \rightarrow n} \triangleq \nu_{\alpha_l \rightarrow n} \left( \frac{\hat{\alpha}_l}{\nu_{\alpha_l}} - \frac{\alpha_{n \rightarrow l}}{\nu_{\alpha_n \rightarrow l}} \right) \quad (20)$$

然后利用 BP 规则更新消息  $m_{f_{\delta_n} \rightarrow h_n}(h_n)$ , 得到:

$$m_{f_{\delta_n} \rightarrow h_n}(h_n) = \langle f_{\delta_n}(h_n, \alpha) \rangle_{\prod_l n_{\alpha_l \rightarrow f_{\delta_n}}(\alpha_l)} \propto \mathcal{CN}(h_n; \hat{p}_n, \nu_{p_n}) \quad (21)$$

通过  $b(h_n) \propto m_{f_{\delta_n} \rightarrow h_n}(h_n) n_{h_n \rightarrow f_{\delta_n}}(h_n) \propto \mathcal{CN}(h_n; \hat{h}_n, \nu_{h_n})$  计算置信  $b(h_n)$ 。其中:

$$\nu_{h_n} \triangleq \left( \hat{\lambda} + \frac{1}{\nu_{p_n}} \right)^{-1} \quad (22)$$

$$\hat{h}_n \triangleq \nu_{h_n} \left( y_n \hat{\lambda} + \frac{\hat{p}_n}{\nu_{p_n}} \right) \quad (23)$$

利用 MF 准则计算消息  $m_{f_{\delta_n} \rightarrow \lambda}(\lambda) \propto \lambda \exp\{-\langle |y_n - h_n|^2 \rangle_{b(h_n)}\}$ , 而  $f(\lambda) = \frac{1}{\lambda}$ , 置信  $b(\lambda)$  的更新为:

$$b(\lambda) \propto m_{f_{\delta_n} \rightarrow \lambda}(\lambda) f_{\delta_n}(\lambda) \propto \lambda^{N-1} \exp\{-\lambda \sum_n \langle |y_n - h_n|^2 \rangle_{b(h_n)}\} \quad (24)$$

计算参数  $\lambda$ :

$$\hat{\lambda} = \langle \lambda \rangle_{b(\lambda)} = \frac{N}{\sum_n \langle |y_n - h_n|^2 \rangle_{b(h_n)}} \quad (25)$$

### 3.1.3 标记

本文提出的 BP-MF SBL 算法与标量形式的 MF SBL 算法的最大区别是: BP-MF SBL 是基于改进因子图计算网络消息的部分, 具体来说是  $m_{f_{\delta_n} \rightarrow \alpha_l}(\alpha_l)$  消息的计算。针对图 1(b), 有:

$$\begin{aligned} m_{f_{\delta_n} \rightarrow \alpha_l}(\alpha_l) &= \int f_{\delta_n}(h_n, \alpha) n_{h_n \rightarrow f_{\delta_n}}(h_n) \prod_{l' \neq l} n_{\alpha_{l'} \rightarrow f_{\delta_n}}(\alpha_{l'}) d\alpha_l dh_n \\ &= \frac{\int f_{\delta_n}(h_n, \alpha) n_{h_n \rightarrow f_{\delta_n}}(h_n) \prod_{l' \neq l} n_{\alpha_{l'} \rightarrow f_{\delta_n}}(\alpha_{l'}) \prod_{l' \neq l} d\alpha_{l'} dh_n}{n_{\alpha_l \rightarrow f_{\delta_n}}(\alpha_l)} \\ &= \frac{\int b(h_n, \alpha) \prod_{l' \neq l} d\alpha_{l'} dh_n}{n_{\alpha_l \rightarrow f_{\delta_n}}(\alpha_l)} \end{aligned} \quad (26)$$

其中,  $b(h_n, \alpha)$  表示与函数节点  $f_{\delta_n}$  相邻的变量节点  $\{\alpha_1, \dots, \alpha_L, h_n\}$  的联合置信。MF SBL 则是基于传统因子图(见图 1(a))计算  $m_{f_{\delta_n} \rightarrow \alpha_l}(\alpha_l)$ 。变量节点  $\{\alpha_1, \dots, \alpha_L\}$  的置信  $b(\alpha_l)$  相互独立。

## 3.2 BP-MF SBL 算法流程

在图 1(b) 中由于网络的存在, 其消息更新流程有很多种。本文选择一种简单的算法流程, 按照顺序更新前向后向消息, 且同时更新  $N$  个和  $L$  个变量的信息。算法 1 给出了 BP-MF SBL 算法的流程。

### 算法 1 BP-MF SBL 算法

1. 初始化  $\hat{\alpha}_l, \nu_{p_n}, \hat{p}_n, \hat{\lambda}, \hat{\alpha}_{l \rightarrow n}, \nu_{\alpha_{l \rightarrow n}}$ 。
2. 通过式(5)和式(6)更新  $\hat{\alpha}_{l \rightarrow n}$  和  $\nu_{\alpha_{l \rightarrow n}}$ 。
3. 通过式(10)和式(11)更新  $\hat{q}_l$  和  $\nu_{\alpha_l}$ 。
4. 通过式(13)和式(14)更新  $\hat{\alpha}_l$  和  $\nu_{\alpha_l}$ 。
5. 利用式(17)更新  $\hat{\gamma}_l$ 。
6. 利用式(13)和式(14)再次更新  $\hat{\alpha}_l$  和  $\nu_{\alpha_l}$ 。
7. 利用式(19)和式(20)更新  $\nu_{\alpha_{l \rightarrow n}}$  和  $\hat{\alpha}_{l \rightarrow n}$ 。
8. 利用式(7)和式(8)更新  $\hat{p}_n$  和  $\nu_{p_n}$ 。
9. 利用式(22)和式(23)更新  $\hat{h}_n$  和  $\nu_{h_n}$ 。
10. 利用式(25)更新  $\hat{\lambda}$ 。

## 4 近似 BP-MF SBL 算法

从图 1 可以看出, 有  $NL$  条边连接变量节点  $\alpha_l$  和函数节点  $f_{\delta_n}$ 。因此, 在每一次迭代过程中, 在计算前向和后向消息时, 需要计算  $2NL$  次。为了得到复杂度更低的算法, 用 GAMP 算法近似 BP 部分高斯消息的均值和方差, 得到近似 BP-MF SBL 算法。

### 4.1 近似消息

将式(14)代入式(19)得:

$$\nu_{\alpha_{l \rightarrow n}} = (1/\nu_{\alpha_l} + \hat{\gamma}_l - 1/\nu_{\alpha_{l \rightarrow l}})^{-1} \approx \nu_{\alpha_l} \quad (27)$$

当  $N$  足够大时, 由式(10)知:  $\frac{1}{\nu_{\alpha_l}} \gg \frac{1}{\nu_{\alpha_{l \rightarrow n}}}$ , 同理将式(5)和式(6)代入式(20), 得:

$$\begin{aligned} \hat{\alpha}_{l \rightarrow n} &= \nu_{\alpha_{l \rightarrow n}} \left( \frac{\hat{\alpha}_l}{\nu_{\alpha_l}} - \frac{\Phi_{nl}^*(y_n - \hat{p}_n + \Phi_{nl} \alpha_{l \rightarrow n}^{-1})}{\hat{\lambda}^{-1} + \nu_{p_n} - |\Phi_{nl}|^2 \nu_{\alpha_{l \rightarrow n}}^{-1}} \right) \\ &\approx \hat{\alpha}_l - \nu_{\alpha_l} \frac{y_n - \hat{p}_n}{\hat{\lambda}^{-1} + \nu_{p_n}} \Phi_{nl}^* = \alpha_l - \nu_{\alpha_l} s_n \Phi_{nl}^* \end{aligned} \quad (28)$$

其中:

$$s_n \triangleq \frac{y_n - \hat{p}_n}{\hat{\lambda}^{-1} + \nu_{p_n}} \quad (29)$$

上述公式成立的条件是变量  $\alpha$  的数量足够多, 以保证可以从式(8)得到:  $\nu_{p_n} \gg |\Phi_{nl}|^2 \nu_{\alpha_{l \rightarrow n}}$  和  $\hat{p}_n \gg \Phi_{nl} \alpha_{l \rightarrow n}$ 。

分别将式(27)和式(28)代入式(8)和式(7), 得到近似的均值方差:

$$\nu_{p_n} \approx \sum_l |\Phi_{nl}|^2 \nu_{\alpha_l} \quad (30)$$

$$\hat{p}_n \approx \sum_l \Phi_{nl} (\alpha_l - \nu_{\alpha_l} s_n \Phi_{nl}^*) \approx \sum_l \Phi_{nl} \alpha_l - s_n \nu_{p_n} \quad (31)$$

将式(5)和式(6)代入式(10)和式(11), 设置  $L$  足够大, 如下所示:

$$\nu_{\alpha_l} = \left( \sum_n \frac{|\Phi_{nl}|^2}{\hat{\lambda}^{-1} + \nu_{p_n} - |\Phi_{nl}|^2 \nu_{\alpha_{l \rightarrow n}}} \right)^{-1} \approx \left( \sum_n \frac{|\Phi_{nl}|^2}{\hat{\lambda}^{-1} + \nu_{p_n}} \right)^{-1} \quad (32)$$

$$\begin{aligned} \hat{q}_l &= \nu_{\alpha_l} \left( \sum_n \frac{\Phi_{nl}^* (y_n - \hat{p}_n + \Phi_{nl} \alpha_{l \rightarrow n}^{-1})}{\hat{\lambda}^{-1} + \nu_{p_n} - |\Phi_{nl}|^2 \nu_{\alpha_{l \rightarrow n}}} \right) \\ &\approx \nu_{\alpha_l} \sum_n s_n \Phi_{nl}^* + \nu_{\alpha_l} \sum_n \frac{|\Phi_{nl}|^2}{\hat{\lambda}^{-1} + \nu_{p_n}} \alpha_{l \rightarrow n} \\ &\approx \nu_{\alpha_l} \sum_n s_n \Phi_{nl}^* + \hat{\alpha}_l \end{aligned} \quad (33)$$

### 4.2 近似 BP-MF SBL 算法流程

与 BP-MF SBL 算法流程相似,在近似 BP-MF SBL 算法中,利用上述近似公式代替 BP-MF SBL 算法中对应的部分。利用式(32)和式(33)代替式(10)和式(11)估计参数  $\nu_{q_i}$  和  $\hat{q}_i$ , 利用式(30)和式(31)代替式(8)和式(7)计算  $\nu_{p_n}$  和  $\hat{p}_n$ 。此外,在 BP-MF SBL 算法的第 3 行和第 8 行中,不需要计算  $\hat{\alpha}_{n \rightarrow l}$ ,  $\nu_{\alpha_i \rightarrow n}$  和  $\nu_{\alpha_i \rightarrow l}$ , 而插入中间变量  $s_n$ 。算法 2 给出近似 BP-MF SBL 算法的流程。

#### 算法 2 近似 BP-MF SBL 算法

1. 初始化  $\nu_{p_n}, \hat{\gamma}_1, s_n, \hat{\alpha}_1, \hat{\lambda}$ 。
2. 利用式(32)和式(33)更新  $\nu_{q_i}$  和  $\hat{q}_i$ 。
3. 利用式(13)和式(14)更新  $\hat{\alpha}_1$  和  $\nu_{\alpha_i}$ 。
4. 利用式(17)更新  $\hat{\gamma}_1$ 。
5. 利用式(13)和式(14)再次更新  $\hat{\alpha}_1$  和  $\nu_{\alpha_i}$ 。
6. 利用式(30)和式(31)更新  $\hat{p}_n$  和  $\nu_{p_n}$ 。
7. 利用式(29)更新  $s_n$ 。
8. 利用式(22)和式(23)更新  $\hat{h}_n$  和  $\nu_{h_n}$ 。
9. 更新  $\hat{\lambda}$ 。

## 5 仿真结果说明

### 5.1 性能分析

本节利用蒙特卡洛仿真评价所提出的算法性能,考虑数学模型式(1)并设置  $M \times N (M=100, N=256)$  阶字典矩阵  $\Phi$  中的元素独立同分布,并服从均值为 0、方差为 1 的复高斯随机变量。假设长度为  $N$  的待估计向量  $\alpha$  有  $K$  个非零值,并且均匀分布在  $\alpha$  中。此外,这些非零值独立同分布,并且也服从均值为 0、方差为 1 的复高斯分布。仿真迭代 100 次,且每次迭代的字典矩阵  $\Phi$ 、待估计向量  $\alpha$  和噪声  $\omega$  随机独立产生。

本文将所提算法与当前较好的算法的 MSE 性能进行对比。BP-MF 和 A-BP-MF 分别表示 BP-MF SBL 算法和近似 BP-MF SBL 算法。MF-向量形式和 MF-标量形式分别表示向量形式的 MF SBL 算法<sup>[16]</sup>和标量形式的 MF SBL 算法<sup>[17]</sup>。以上算法都按照文献[5]中的 2-L 分层模型及参数进行设置。

图 2 给出了不同算法的 MSE 在不同信噪比下的变化曲线。所有算法的迭代次数均为 20。待估计向量  $\alpha$  中的非零元素个数  $K=20$ 。BP-MF, A-BP-MF 算法和 MF-向量形式的算法的 MSE 略好于 MF-标量形式。与 MF-向量形式的算法相比, BP-MF 和 A-BP-MF 算法的 MSE 性能优很多。

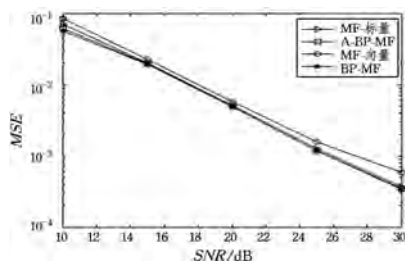


图 2 不同算法的 MSE 比较图

图 3 是在 SNR = 14 dB, K = 20 的情况下各种算法的 MSE 曲线随迭代关系的变化图。从图 3 可以看出,由于采用串行消息更新流程, MF-标量形式算法的收敛速度最快。本

文所提算法的收敛速度比 MF-标量形式及 MF-向量形式-3L 的收敛速度慢,但是其 MSE 性能优于这两种算法。

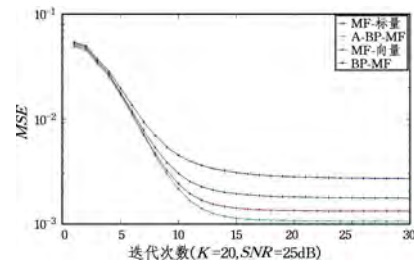


图 3 迭代次数与 MSE 关系图

图 4 给出了 BP-MF(经过拉伸加约束点)与 BP-MF(不加约束点)在不同信噪比下的 MSE 变化曲线。从图 4 中可以看出,随着信噪比的增加, BP-MF(经过拉伸加约束点)的算法性能优于 BP-MF(不加约束点)的算法性能。这是因为基于传统因子图的置信  $b(\alpha_l)$  相互独立,而基于改进因子图的  $b(h_n, \alpha)$  为联合置信。

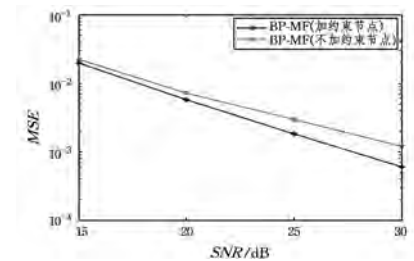


图 4 两种算法的性能对比图

### 5.2 计算复杂度

在所提算法中,由于对参数  $\gamma_l$  和  $\lambda$  的更新算法都是一致的,因此仅分析  $h$  和  $\alpha$  的消息计算复杂度。由于避免了矩阵求逆,在一次迭代过程中 MF-向量形式的复杂度为  $O(L^3)$ , MF-标量形式的则为  $O(NL)$ ,由于提出的 BP-MF 和 A-BP-MF 算法是基于标量形式的因子图的,因此其复杂度与 MF-标量形式相似。具体说来, BP-MF 算法需要计算  $O(NL)$  消息和  $O(NL)$  来存储消息的参数(均值和方差),而 MF-标量形式和 A-BP-MF 只需要更新和存储  $O(N+L)$  消息。然而在更新置信  $b(\alpha_l)$  时, MF-标量形式比 BP-MF 的时间复杂度更大。

### 5.3 算法仿真时间

表 1 列出了 5 种算法的仿真时间。从表中可以看出, MF-向量形式的仿真时间最长,因为其算法复杂度最高为  $O(L^3)$ ; BP-MF(不加约束节点)的仿真时间次之,因为其是基于传统因子图提出的,其在计算变量节点  $\alpha_l$  与因子节点  $f_y$  的  $O(NL)$  消息后,又利用 MF-向量形式更新了  $O(N)$  个  $\hat{\lambda}$  消息,复杂度为  $O(N^2L)$ 。BP-MF 由于要计算  $O(NL)$  消息和  $O(NL)$  来存储消息的参数(均值和方差),仿真时间低于 BP-MF(不加约束节点),多于 A-BP-MF 和 MF-标量形式。MF-标量形式和 A-BP-MF 只需要更新和存储  $O(N+L)$  消息。然而在更新置信  $b(\alpha_l)$  时, MF-标量形式比 A-BP-MF 花费的时间更多,需要  $M \times L$  个消息相乘和  $M \times L$  个消息相加。

表 1 各算法仿真时间的比较

(单位:s)

|      | BP-MF    | A-BP-MF | MF-<br>向量形式 | MF-<br>标量形式 | BP-MF<br>(不加约束节点) |
|------|----------|---------|-------------|-------------|-------------------|
| 仿真时间 | 145.5231 | 91.5478 | 165.6217    | 133.5231    | 155.4432          |

**结束语** 本文讨论了基于消息传递的稀疏贝叶斯学习(SBL),提出了基于拉伸因子图的低复杂度 BP-MF SBL 算法,此拉伸因子图是通过传统因子图添加额外的硬约束因子而得到的。仿真结果显示,在 MSE 性能和复杂度方面,BP-MF SBL 算法的性能优于当前较好的算法 MF SBL。

### 参 考 文 献

- [1] DONOHO D L. Compressed sensing[J]. IEEE Trans. Inform. Theory, 2016, 52(4): 1289-1306.
- [2] WIPF D P, RAO B D. Sparse Bayesian learning for basis selection[J]. IEEE Trans. Signal Processing, 2014, 52(8): 2153-2164.
- [3] CHEN S, DONOHO D L, SAUNDERS M A. Atomic decomposition by basic pursuit[J]. SIAM Journal on Scientific Computing, 2008(1): 33-61.
- [4] TROPP J A. Greed is good: algorithmic results for sparse approximation[J]. IEEE Trans. Information Theory, 2013: 2231-2242.
- [5] TIPPING M E. Sparse Bayesian learning and the relevance vector machine[J]. Journal of Machine Learning Research, 2011, 1: 211-244.
- [6] TIPPING M E, FAUL A C. Fast marginal likelihood maximisation for sparse Bayesian models[C] // Proc. 9th International Workshop on Artificial Intelligence and Statistics. 2003: 3-6.
- [7] SHUTIN D, BUCHGRABER T, KULKARNI S R, et al. Fast variational sparse Bayesian learning with automatic relevance determination for superimposed signals[J]. IEEE Trans. Signal Processing, 2011, 59(12): 6257-6261.
- [8] PEDERSEN N L, MANCHÓN C N, BADIU M A, et al. Sparse estimation using Bayesian hierarchical prior modeling for real and complex linear models[J]. Signal Processing, 2015, 115: 94-109.
- [9] TAN X, LI J. Computationally efficient sparse Bayesian learning via belief propagation[C] // 2009 Conference Record of the Forty-Third Asilomar Conference on Signals, Systems and Computers, 2009: 1566-1570.
- [10] BARON D, SARVOTHAM S, BARANIUK R. Bayesian compressive sensing via belief propagation[J]. IEEE Trans. Signal Processing, 2010, 58(1): 269-280.
- [11] SOM S, SCHNITER P. Compressive imaging using approximate message passing and a Markov-tree prior[J]. IEEE Trans. Signal Processing, 2012, 60(7): 3439-3448.
- [12] AL-SHOUKAIRI M, RAO B. Sparse Bayesian learning using approximate message passing[C] // 2014 48th Asilomar Conference on Signals, Systems and Computers. 2014: 1957-1961.
- [13] XING E P, JORDAN M I, RUSSELL S A. A generalized mean field algorithm for variational inference in exponential families[C] // Proceedings of the Nineteenth Conference on Uncertainty in Artificial Intelligence (UAI'03). San Francisco, CA, USA, 2013: 583-591.
- [14] BISHOP C M, WINN J. Structured variational distributions in VIBES[C] // Proceedings of Artificial Intelligence and Statistics. 2003: 3-6.
- [15] DAUWELS J. On variational message passing on factor graphs[C] // Proc. IEEE International Symposium on Information Theory (ISIT 2007). 2007: 2546-2550.
- [16] PEDERSEN N L, MANCHÓN C N, SHUTIN D, et al. Application of Bayesian hierarchical prior modeling to sparse channel estimation[C] // 2007 IEEE International Conference on Communications (ICC 2012). 2012: 3487-3492.
- [17] HANSEN T L, JØRGENSEN P B, BADIU M, et al. Joint sparse channel estimation and decoding: Continuous and discrete domain sparsity[J]. arXiv: 1507. 02954.
- [18] RIEGLER E, KIRKELUND G E, MANCHÓN C N, et al. Merging belief propagation and the mean field approximation: A free energy approach[J]. IEEE Trans. Inform. Theory, 2013, 59(1): 588-602.
- [19] RANGAN S. Generalized approximate message passing for estimation with random linear mixing[C] // Proc. IEEE Int. Symp. on Inform. Theory (ISIT 2011). 2011: 2168-2172.
- [20] PEDERSEN N L, MANCHÓN C N, FLEURY B H. Low complexity sparse Bayesian learning for channel estimation using generalized mean field[C] // 20th European Wireless Conference. 2014: 838-843.
- [21] KSCHISCHANG F, FREY B, LOELIGER H A. Factor graphs and the sum-product algorithm[J]. IEEE Trans. Inform. Theory, 2001, 47(2): 498-519.
- (上接第 109 页)
- [2] FAN Z X, CHEN S Y, ZHA L, et al. A Text Clustering Approach of Chinese News Based on Neural Network Language Model[J]. International Journal of Parallel Programming, 2016, 44(1): 198-206.
- [3] CAO Q M, GUO Q, WANG Y L, et al. Text clustering using VSM with feature clusters[J]. Neural Computing & Applications, 2015, 26(4): 995-1003.
- [4] GUO L T, LI Y, MU D J, et al. A LDA model based topic detection method[J]. Journal of Northwestern Polytechnical University, 2016, 34(4): 98-102.
- [5] 王刚, 钟国祥. 一种基于本体相似度计算的文本聚类研究[J]. 计算机科学, 2010, 37(9): 222-224.
- [6] HU X H, ZHANG X, et al. Exploiting Wikipedia as External Knowledge for Document Clustering[C] // ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Paris, France, 2009: 389-396.
- [7] HUANG A, MILNE D, FRANK E, et al. Clustering Documents using a Wikipedia Based Concept Representation [M] // Advanced in Knowledge Discovery and Data Mining. Springer Berlin Heidelberg, 2009: 628-636.
- [8] HOFMANN T. Probabilistic latent semantic indexing[C] // 22nd International ACM SIGIR Conference on Research and Development in Information Retrieval. Berkeley, CA, USA, 1999: 50-57.
- [9] BLEI D, NG A, JORDAN M. Latent dirichlet allocation[J]. Journal of Machine Learning Research, 2003(3): 993-1022.
- [10] 徐戈, 黄厚峰. 自然语言处理中主题模型的发展[J]. 计算机学报, 2011, 34(8): 1423-1437.
- [11] 曹娟, 张勇东. 一种基于密度的自适应最优 LDA 模型选择方法[J]. 计算机学报, 2008, 31(10): 1780-1788.
- [12] TEH Y, JORDAN M, BEAL M, et al. Hierarchical dirieght processes[J]. Journal of the American Statistical Association, 2007, 101(476): 1566-1581.
- [13] 张超, 陈利, 李琼. 一种 PST\_LDA 中文文本相似度计算方法[J]. 计算机应用研究, 2016, 33(2): 375-377.
- [14] 黄承慧, 印鉴, 侯昉. 一种结合词项语义信息和 TF-IDF 方法的文本相似度度量方法[J]. 计算机学报, 2011, 34(5): 856-864.