Signal Processing 131 (2017) 344–349

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Low complexity sparse Bayesian learning using combined belief propagation and mean field with a stretched factor graph

Chuanzong Zhang [a,b], Zhengdao Yuan [c,d], Zhongyong Wang [a,*], Qinghua Guo [e,f]

[a] School of Information Engineering, Zhengzhou University, Kexue Road 100, Zhengzhou 450001, China
[b] Department of Electronic Systems, Aalborg University, Niels Jernes Vej 12, DK-9220 Aalborg, Denmark
[c] National Digital Switching System Engineering and Technological Research and Development Center, Zhengzhou 450000, China
[d] Zhengzhou Institute of Information Science and Technology, Zhengzhou 450001, China
[e] School of Electrical, Computer and Telecommunications Engineering, University of Wollongong, Wollongong, NSW2522, Australia
[f] School of Electrical, Electronic and Computer Engineering, University of Western Australia, Crawley, WA6009, Australia

ARTICLE INFO

Article history:
Received 25 February 2016
Received in revised form
12 August 2016
Accepted 20 August 2016
Available online 22 August 2016

Keywords:
Sparse Bayesian learning
Message passing
Belief propagation
Mean field

ABSTRACT

This paper concerns message passing based approaches to sparse Bayesian learning (SBL) with a linear model corrupted by additive white Gaussian noise with unknown variance. With the conventional factor graph, mean field (MF) message passing based algorithms have been proposed in the literature. In this work, instead of using the conventional factor graph, we modify the factor graph by adding some extra hard constraints (the graph looks like being 'stretched'), which enables the use of combined belief propagation (BP) and MF message passing. We then propose a low complexity BP-MF SBL algorithm based on which an approximate BP-MF SBL algorithm is also developed to further reduce the complexity. Thanks to the use of BP, 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.

1. Introduction

Recently, compressed sensing [1,2] has received tremendous attention and it has found wide applications in a large variety of engineering areas, e.g. biomagnetic imaging, sparse channel estimation, bandlimited extrapolation and spectral estimation, echo cancellation and image restoration [3]. In compressed sensing, a vector $\boldsymbol{\alpha} \in \mathbb{C}^{L\times 1}$ which exhibits sparsity is estimated based on the measurement vector $\boldsymbol{y} \in \mathbb{C}^{N\times 1}$ with the following model:

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

where $\boldsymbol{\Phi} \in \mathbb{C}^{N\times L}$ is called dictionary matrix and $\boldsymbol{\omega}$ represents an additive white Gaussian noise (AWGN) vector with zero mean and covariance matrix $\lambda^{-1}\boldsymbol{I}$. In this work, we are particularly interested in the case that the variance of the AWGN (or the precision parameter λ) is unknown.

Besides convex [4] and greedy [5] methods, sparse Bayesian learning (SBL) [6–8] is an alternative method of sparse signal estimation, which aims at finding a sparse maximum a posteriori (MAP) estimate $\hat{\boldsymbol{\alpha}} = \operatorname{argmax}_{\boldsymbol{\alpha}} p(\boldsymbol{\alpha}|\boldsymbol{y})$ of the vector $\boldsymbol{\alpha}$ by specifying a priori probability density function (pdf) $p(\boldsymbol{\alpha})$. Instead of working directly with a prior $p(\boldsymbol{\alpha})$, SBL typically employs a two-layer (2-L) hierarchical structure [9] that assumes a conditional prior pdf $p(\boldsymbol{\alpha}|\boldsymbol{\gamma})$ and a hyper-prior pdf $p(\boldsymbol{\gamma})$, so that $p(\boldsymbol{\alpha}) = \int_{\boldsymbol{\gamma}} p(\boldsymbol{\alpha}|\boldsymbol{\gamma})p(\boldsymbol{\gamma})d\boldsymbol{\gamma}$ has a sparsity-inducing nature. Most recently, SBL has been efficiently implemented using belief propagation (BP) [10,11] and approximate message passing [12,13]. However, these methods assume that λ is known, which may not be true in many applications. This work deals with message passing based approaches to SBL with unknown λ.

Mean field (MF) based message passing [14–16], which is also often referred to as variational message passing (VMP), has been widely used for approximate Bayesian inference, especially for exponential distributions. With 2-L or 3-L hierarchical prior structures, Pedersen et al. proposed an MF SBL algorithm (with unknown λ) [17], which was applied to sparse channel estimation in OFDM. As the MF SBL algorithm deals with the sparse signal $\boldsymbol{\alpha}$ in a vector-form, matrix inversion is involved in each iteration and its computational complexity is as high as $O(L^3)$. To address the issue

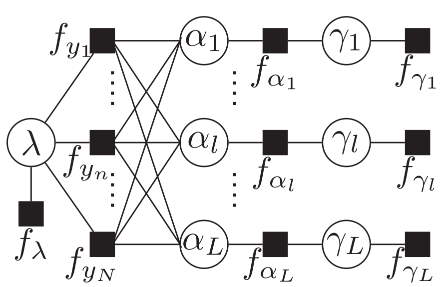
* Corresponding author.
E-mail addresses: ieczzhang@gmail.com (C. Zhang), yuan_zhengdao@foxmail.com (Z. Yuan), iezywang@zzu.edu.cn (Z. Wang), qguo@uow.edu.au (Q. Guo).

http://dx.doi.org/10.1016/j.sigpro.2016.08.027
0165-1684/

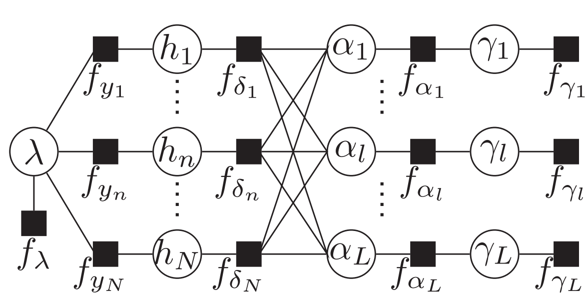
of complexity, a low complexity MF SBL algorithm [18] is then proposed, where the inverse of a large matrix is decomposed into a number of matrix inverses with smaller size. Flexible trade-off between complexity and performance can be achieved by adjusting the size of smaller matrices, which means that the reduction of complexity comes at the cost of performance loss. Apparently, the size of the smaller matrices can be set to be 1, so that the matrix inverses are avoided and we call it scalar-form MF SBL algorithm. Recently, the scalar-form MF SBL algorithm was used for channel gain and delay estimation in [19]. We note that an efficient hyperprior $p(\boldsymbol{\alpha})$ with 2-L structure was proposed in [6], which performs better than the 2-L and 3-L structures in [17].

Different from MF which supposes all the beliefs of variable nodes are independent, BP considers the joint belief of variable nodes neighboring a factor node and makes the most of their correlation. BP, which may achieve exact Bayesian inference, is efficient to deal with discrete probability models and linear Gaussian models. However, BP may have a high complexity, when especially dealing with models involving both discrete and continuous random variables. Recently, a unified message passing framework was proposed in [20] where BP and MF are merged to keep the merits of BP and MF while avoid their drawbacks.

In this work, a low complexity BP-MF SBL algorithm with a 2-L hierarchical prior is proposed. Instead of using the conventional factor graph shown in Fig. 1(a), we modify the factor graph by adding a number of extra hard constraint factors as shown in Fig. 1(b), i.e., the factor graph looks like being 'stretched'. The hard constraint factors seem redundant, which however facilitates the use of BP in the graph, leading to considerable performance improvement. As we assume that the noise variance λ^{-1} is unknown, MF can be used to tackle the exponential factors, while BP is used to handle the hard constraint factors. As we factorize the signal $\boldsymbol{\alpha}$ in a scalar form, the developed BP-MF SBL algorithm avoids matrix inversion and has a low complexity. Inspired by the derivation of the generalized approximate message passing (GAMP) [21], we further simplify the BP message passing by ignoring some minimal terms and develop an approximate BP-MF SBL algorithm.



(a) Conventional factor graph



(b) Stretched factor graph

Fig. 1. Two factor graph representations for the probabilistic model (2).

Numerical examples show that the proposed BP-MF SBL algorithms provide even better mean-square-error (MSE) performance with much lower complexity compared with the vector-form MF SBL algorithm [17], and achieve noticeable MSE performance gain with similar complexity compared with the scalar-form MF SBL algorithm [18,19].

Notation: Boldface lowercase and uppercase letters denote vectors and matrices, respectively. The expectation operator with respect to a pdf $g(x)$ is expressed by $\langle f(x)\rangle_{g(x)} = \int f(x)g(x)dx / \int g(x')dx'$, while $\mathrm{var}[x]_{g(x)} = \left\langle |x|^2 \right\rangle_{g(x)} - |\langle x\rangle_{g(x)}|^2$ stands for the variance. The pdf of a complex Gaussian distribution with mean μ and variance ν is represented by $\mathcal{CN}(x; \mu, \nu)$. The relation $f(x) = cg(x)$ for some positive constant c is written as $f(x) \propto g(x)$.

2. Factor graph model

The joint a posteriori pdf of $\boldsymbol{\alpha}, \boldsymbol{\gamma}$ and λ in (1) with a 2-L hierarchical prior [9] can be factorized as

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

where $f_{y_n}(\boldsymbol{\alpha}, \lambda) \triangleq p(y_n|\boldsymbol{\alpha}, \lambda) = \mathcal{CN}(y_n; \boldsymbol{\Phi}_n\boldsymbol{\alpha}, \lambda^{-1})$, with $\boldsymbol{\Phi}_n$ being the n-th row of matrix $\boldsymbol{\Phi}$, and $f_\lambda(\lambda)$ denotes the prior of noise precision parameter λ. The factor $f_{\alpha_l}(\alpha_l, \gamma_l)$ denotes the conditional pdf $p(\alpha_l|\gamma_l) = \mathcal{CN}(\alpha_l; 0, \gamma_l^{-1})$, which is chosen as a Gaussian prior of α_l and $f_{\gamma_l}(\gamma_l)$ represents a hyperprior $p(\gamma_l) = \mathrm{Ga}(\gamma_l; \epsilon, \eta)$[1] of the hyperparameter γ_l. The factorization in (2) can be visually depicted on the factor graph [22] as shown in Fig. 1(a), which is similar to those in [18,19]. We assume that λ is unknown, and MF can be used to deal with factor nodes $\{f_{y_n}, \forall n \in [1{:}\, N]\}$, which leads to the scalar-form MF SBL algorithm [18]. In [17], the vector-form MF SBL algorithm is derived based on a conventional factor graph, where the vector $\boldsymbol{\alpha}$ is treated as a single variable node.

To facilitate the use of both BP and MF, we modify the factor graph in Fig. 1(a) by adding hard constraint factors $\{f_{\delta_n}(h_n, \boldsymbol{\alpha}) = \delta(h_n - \boldsymbol{\Phi}_n\boldsymbol{\alpha}), \forall n \in [1{:}\, N]\}$ with a new variable vector $\boldsymbol{h} = \boldsymbol{\Phi}\boldsymbol{\alpha}$. Therefore, factor f_{y_n} denotes the likelihood function $p(y_n|h_n, \lambda) = \mathcal{CN}(y_n; h_n, \lambda^{-1})$. The new factor graph, shown in Fig. 1(b), looks like a stretched version of the graph in Fig. 1(a). In the new graph, MF rules with fixed points equations can be used to compute the messages for the exponential factors, while BP rules, often yielding better performance, can be used to deal with the hard constraint factors. The message computations and scheduling are detailed in the following section.

3. BP-MF based SBL

In this section, with the combined BP-MF message update rule [20], we detail the message computations and scheduling on the factor graph shown in Fig. 1(b) to perform sparse signal estimation. All the factors in Fig. 1(b) are represented by set $\mathcal{A}$, and it is divided into two disjoint subsets, a BP subset and an MF subset, which are denoted by $\mathcal{A}_{\mathrm{BP}} = \{f_{\delta_n}, \forall n\}$ and $\mathcal{A}_{\mathrm{MF}} = \mathcal{A} \backslash \mathcal{A}_{\mathrm{BP}}$, respectively.

[1] $\mathrm{Ga}(\cdot; a, b)$ denotes a Gamma pdf with shape parameter a and rate parameter b. Note that, as in [6], we use the Gamma prior for the parameter of precision, rather than for variance [17].