Signal Processing 118 (2016) 153–158

Contents lists available at ScienceDirect

Signal Processing

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

Fast communication

Multiple-measurement vector based implementation for single-measurement vector sparse Bayesian learning with reduced complexity

Jian A. Zhang [a,*], Zhuo Chen [a], Peng Cheng [a], Xiaojing Huang [b]

[a] *The Wireless and Networks Program, CSIRO, Australia*
[b] *University of Technology, Sydney, Australia*

ARTICLE INFO

Article history:
Received 2 March 2015
Received in revised form 15 June 2015
Accepted 28 June 2015
Available online 6 July 2015

Keywords:
Compressive sensing
Sparse Bayesian learning
Direction of arrival estimation

ABSTRACT

Sparse Bayesian learning (SBL) has high computational complexity associated with matrix inversion in each iteration. In this paper, we investigate complexity reduced multiple-measurement vector (MMV) based implementation for single-measurement vector SBL problems. For problems with special structured sensing matrices, we propose two sub-optimal SBL schemes with significantly reduced complexity and slight estimation performance degradation, by exploiting the deterministic correlation in the converted MMV model explicitly. Two application scenarios on channel estimation in multicarrier systems and direction of arrival estimation are presented. Simulation results validate the effectiveness of the schemes.

1. Introduction

Sparse Bayesian learning (SBL) [1,2], also known as Bayesian compressive sensing (BCS) [3], is one of the well-known compressive sensing techniques. It explores a Bayesian model based on a prior knowledge to model signal sparsity, and forms an optimal estimator using the maximum-a-posteriori (MAP) solution. The SBL framework provides an effective proxy to indirectly solve l_0 minimization, leading to a more accurate recovery performance than many l_1 minimization solutions [3], particularly when the modelling of the statistical properties of the unknown non-zero variables is accurate. Unlike l_1 minimization solutions, which typically require that the measurement matrix satisfies the restricted isometry property (RIP) to achieve maximally sparse solution, SBL can still work efficiently when RIP is not satisfied.

In SBL, typically, the posteriori probability density function (pdf) is represented as a multivariate Gaussian function, and its mean and variance are estimated iteratively, using expectation–maximization (EM) algorithms. The estimation process involves matrix inversion and large matrix multiplication in each iteration, and hence SBL generally has much higher computational complexity than many other CS techniques. A fast algorithm for SBL is proposed in [4] and further developed in [3,5] by adding/removing signal basis and avoiding matrix inversion. However, we found such algorithms converge slowly and suffer from significant estimation performance degradation.

In this paper, we investigate complexity-reduced multiple-measurement vector (MMV) based implementation for single-measurement vector (SMV) SBL problems by converting a SMV model to a MMV one. We propose two sub-optimal low-complexity SBL schemes based on multitask BCS [5] and simultaneous SBL [2] algorithms, by jointly combining and processing the outputs in each iteration of these two algorithms to achieve better performance. Unlike existing group sparse Bayesian algorithms [6–8]

* Corresponding author.
E-mail address: Andrew.Zhang@csiro.au (J.A. Zhang).

http://dx.doi.org/10.1016/j.sigpro.2015.06.020
0165-1684/

which learn and exploit the group correlation, our schemes use the deterministic correlation explicitly. For problems with a special sensing matrix, we apply maximum ratio combining (MRC) in each iteration to improve the estimation performance. These solutions reduce the dimension of the matrix to be inverted and can significantly reduce the computational complexity. We first consider the formulation of an MMV model for a general SMV problem in Section 3.1, derive its optimal solution in Section 3.2, and present a general suboptimal solution in Section 3.3. We then focus on a particular signal model where the sensing matrix has a special structure. We present two sub-optimal solutions for this special model in Sections 3.3 and 3.4, and exemplify two applications in Section 4. Simulation results provided in Section 5 demonstrate the effectiveness of the proposed scheme.

2. System model

A general SMV model is given by

$$\mathbf{t} = \mathbf{\Psi}\mathbf{w} + \mathbf{z}, \tag{1}$$

where $\mathbf{t} \in \mathbb{C}^{N'}$ is the measured signal vector, $\mathbf{\Psi} \in \mathbb{C}^{N' \times M}$ is the sensing matrix, $\mathbf{w} = \{w_m\} \in \mathbb{C}^M, m = 1, \ldots, M$, is a sparse signal vector with support K (at most K non-zero values and $K \ll M$), and $\mathbf{z} \in \mathbb{C}^{N'}$, $\mathbf{z} \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I}_{N'}, \mathbf{0})$ is an additive noise vector following a complex Gaussian distribution with zero mean, correlation $\sigma^2 \mathbf{I}_{N'}$ and relation matrix $\mathbf{0}$.

A special model we are particularly interested in is

$$\mathbf{\Psi} = [(\mathbf{\Phi}\mathbf{D}_1)^T, (\mathbf{\Phi}\mathbf{D}_2))^T, \ldots, (\mathbf{\Phi}\mathbf{D}_L)^T]^T, \tag{2}$$

where $\mathbf{\Phi} \in \mathbb{C}^{N \times M}$ is a sub- sensing matrix, $\mathbf{D}_i = \mathrm{diag}\{d_{i,m}\} \in \mathbb{C}^{M \times M}, i = 1, \ldots, L, m = 1, \ldots, M$, is a diagonal matrix. Both $\mathbf{\Phi}$ and $\mathbf{D}_i$ are known. This model represents a large class of applications such as direction of arrival (DoA) estimation [9], channel estimation [10], CFO estimation [11], and linear array synthesis problem [12], where, e.g., the sensing matrix is a partial discrete Fourier Transform (DFT) matrix.

Assume that $w_m, m = 1, \ldots, M$, are independent complex Gaussian random variables[1] with zero mean and variance γ_m. Existing results for real signals, such as simultaneous SBL [2] and multitask compressive sensing [5], can be directly extended to complex signals by changing matrix transpose to conjugate transpose, when

$$\mathrm{E}[\Re(w_m)\Re(w_n)] = \mathrm{E}[\Im(w_m)\Im(w_n)] \quad \text{and}$$
$$\mathrm{E}[\Re(w_m)\Im(w_n)] = -\mathrm{E}[\Re(w_n)\Im(w_m)] \tag{3}$$

for any n, m [13]. The first condition can be satisfied when the real and imaginary parts of every w_m have the same variance and different w_m are independent variables; and the second condition becomes true when the real and imaginary parts of every w_m are statistically independent. Hence these two conditions can be met with reasonable assumptions which have insignificant effect on the performance of SBL, unless the elements in $\mathbf{w}$ show large correlation. In this case, separating the real and imaginary parts of $\mathbf{w}$ and expanding the original SMV problem from dimension $N' \times M$ to $2N' \times M$ will be beneficial, and block SBL [6] can also better capture such correlation. Hereafter, we will assume that the conditions in (3) are satisfied.

To this end, $\mathbf{w}$ follows a multivariate complex Gaussian distribution with mean $\mathbf{0}$, covariance matrix $\mathbf{\Gamma} = \mathrm{diag}\{\gamma_m\}$, $m = 1, \ldots, M$, where γ_m is the variance of w_m, and relation matrix $\mathbf{0}$. Let $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \ldots, \gamma_M)$. The joint prior probability density function (pdf) of $\mathbf{w}$ is then given by

$$p(\mathbf{w}; \boldsymbol{\gamma}) = \prod_{m=1}^{M} \mathcal{CN}(0, \gamma_m, 0). \tag{4}$$

The SBL solution is based on MAP optimization where the posterior $p(\mathbf{w}|\mathbf{t}; \boldsymbol{\gamma})$ can be expressed as a multivariate Gaussian distribution with mean and covariance [2]

$$\mathbf{u}^{(o)} = \mathbf{\Gamma}\mathbf{\Psi}^H \mathbf{\Sigma}_t^{-1} \mathbf{t},$$
$$\mathbf{\Sigma}^{(o)} = \mathbf{\Gamma} - \mathbf{\Gamma}\mathbf{\Psi}^H \mathbf{\Sigma}_t^{-1} \mathbf{\Psi}\mathbf{\Gamma}, \tag{5}$$

where $\mathbf{\Sigma}_t = \sigma^2 \mathbf{I}_{N'} + \mathbf{\Psi}\mathbf{\Gamma}\mathbf{\Psi}^H$.

The MAP solution is given by $\mathbf{u}^{(o)}$. The type-II maximum likelihood approach, such as the EM and fast Mackay algorithms [2], can iteratively compute and update $\boldsymbol{\gamma}$ and $\mathbf{u}^{(o)}$.

The computational complexity of the SMV SBL scheme above is $\mathcal{O}(MN'^2)$, mainly associated with the matrix inversion and matrix multiplication in (5). For the special model (2), in general, no fast algorithm is known for computing the inversion of $\mathbf{\Sigma}_t$ and the whole complexity remains similar, except for some limited occasions of special $\mathbf{D}_i$.[2]

3. MMV formulation and solutions

3.1. MMV formulation of the SMV model

The SMV problem in (1) can be re-formulated as an MMV problem as

$$\mathbf{t}_i = \mathbf{\Phi}_i \mathbf{w} + \mathbf{z}_i, \quad i = 1, \ldots, L, \tag{6}$$

where $\mathbf{t}_i$, $\mathbf{\Phi}_i$ and $\mathbf{z}_i$ contain a fixed number of N rows of $\mathbf{t}$, $\mathbf{\Psi}$ and $\mathbf{z}$, respectively. Note that these rows can be overlapped in different measurement vectors, and all the measurements form a complete measurement of $\mathbf{t}$. Hence $N' \le NL$. It is not necessary to form more than $L = \lceil N'/N \rceil$ MMVs, where $\lceil x \rceil$ denotes the smallest integer no less than x. For example, when $N' = 5$, we can form $\mathbf{t}_1 = [t_1, t_2, t_3]^T$ and $\mathbf{t}_2 = [t_3, t_4, t_5]^T$ for L=2; or $\mathbf{t}_1 = [t_1, t_2]^T$, $\mathbf{t}_2 = [t_3, t_4]^T$ and $\mathbf{t}_3 = [t_4, t_5]^T$ for L=3.

For the special sensing matrix in (2), we note that the above MMV formulation is also applicable, and overlapped rows can be used in different $\mathbf{\Phi}_i = \mathbf{\Phi}\mathbf{D}_i$.

Now the task is to estimate $\mathbf{w}$ over L measurements $\mathbf{T} = \{\mathbf{t}_i\}, i = 1, \ldots, L$.

[1] To encourage sparsity presentation, SBL introduces a sparse prior probability function which should exhibit a sharp peak for w_m=0 and small tails for $w_m \ne 0$, and should also enable simple calculation. A Gaussian probability function has these features and is typically used. Whether the signal actually follows a Gaussian distribution is not that important.

[2] One such example is $\mathbf{D}_i = a_i \mathbf{D}$ where a_i is a scalar coefficient. In this very special case, both the sensing matrices and the signals to be estimated in the MMV model are the same in different measurements, and algorithms operating in the reduced dimension of N, such as the one in [6], can be applied with reduced complexity of $\mathcal{O}(MN^2)$.