

稀疏贝叶斯学习(Sparse Bayesian Learning)

张智林 (Zhilin Zhang)

z4zhang@ucsd.edu

Department of Electrical and Computer Engineering, University of California, San Diego,
La Jolla, CA 92093-0407, USA

1 引言

稀疏贝叶斯学习 (Sparse Bayesian Learning, SBL) 最初作为一种机器学习算法由 Tipping 于 2001 年前后提出 [Tipping2001], 随后被引入到稀疏信号恢复/压缩感知领域 [Wipf2004, Ji2008]。Wipf 和 Rao 等人对 SBL 进行了深入的理论研究。与广泛使用的基于 L1 惩罚项的算法 (比如 Lasso, Basis Pursuit) 相比 (以下简称 L1 算法), SBL 具有一系列显著的优势: (1) 在无噪情况下, 除非满足一些严格的条件 [Donoho2003], L1 算法的全局最小点 (global minimum) 并不是真正最稀疏的解 [Wipf2004]。因此, 在一些应用中, 当真实的解是最稀疏的解, 采用 SBL 是更好的选择。(2) 当感知矩阵 (sensing matrix) 的列与列相关性很强时, L1 算法的性能会变得非常差。事实上不光是 L1 算法, 绝大多数已知的压缩感知算法 (比如 Approximate Message Passing 算法, Matching Pursuit 算法) 在这种情况下性能都会变得很差。相比之下, SBL 算法仍旧具有良好的性能 [Wipf_NIPS2011]。因此, 在雷达追踪, 波达方向估计, 脑源定位, 特征提取, 功率谱估计等一些列领域, SBL 都具备显著的优势。(3) 业已证明, SBL 算法等价于一种迭代加权 L1 最小化算法 (iterative reweighted L1 minimization), 而 L1 算法仅仅只是其第一步 [Wipf2010]。Candes 等人指出, 迭代加权 L1 最小化算法更易获得真正最稀疏解 [Candes2008]。从这个角度也就不难理解 SBL 的优越性。(4) 在很多实际问题中, 所期望的稀疏解常常有一些结构, 而利用这些结构可以获得更好的性能 [ModelCS]。作为一种贝叶斯算法, SBL 算法对利用这些解的结构信息提供了更多的灵活性。这种灵活性最主要来自于 SBL 采用参数化的高斯分布为解的先验分布。最近 Zhang 和 Rao 提出了块稀疏贝叶斯学习框架 (Block Sparse Bayesian Learning, BSBL) [Zhang_IEEE2011, Zhang_TSP2012]。该框架提供了一种利用解的空间结构 (spatial structure) 和时序结构 (temporal structure) 的解决方案。由其框架得到的算法在多任务学习 (multi-task learning) [Wan2012], 生理信号的无线传输和远程监控 [Zhang_TBME2012a, Zhang_TBME2012b], 脑源定位和脑-机接口 [Zhang_PIEEE2012] 等许多领域获得了极大的成功。

下面将首先介绍基本的 SBL 框架, 然后对 BSBL 框架及其算法进行详细介绍, 并在最后给出一些代表性的实验结果。

2 稀疏贝叶斯学习

压缩感知的基本模型可描述为:

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{v} \quad (1)$$

其中 \mathbf{A} 为 $N \times M$ 的感知矩阵, \mathbf{y} 为 $N \times 1$ 维压缩信号, \mathbf{x} 为 M 维待求的解向量, \mathbf{v} 为未知的噪声向量。为求解 \mathbf{x} , SBL 假设 \mathbf{x} 中的每个元素都服从一个参数化的均值为 0 方差为 γ_i 的高斯分布 [Wipf2004]:

$$p(x_i; \gamma_i) = N(0, \gamma_i), \quad i = 1, \dots, M \quad (2)$$

其中 x_i 表示 \mathbf{x} 中的第 i 个元素, γ_i 是未知的参数, 将会由算法自动估计出来。这样一种先验分布常被称为 automatic relevance 先验分布, 最初出现于人工神经网络领域[ARD1996]。在算法运行中, 绝大部分的 γ_i 将会变成 0 (无噪情况下) 或者趋于 0 (有噪情况下)。SBL 通常会采用一个阈值将趋近于 0 的 γ_i 置为 0 (该阈值的大小通常和信噪比有关)。当 $\gamma_i = 0$ 时, 相应的 x_i 则为 0。因此, γ_i 与解的稀疏程度密切相关, 也从而决定了 γ_i 的学习规则是 SBL 算法中最核心的部分。在 SBL 框架中, 噪声 \mathbf{v} 通常假设为高斯白噪声向量, 即 $p(\mathbf{v}; \lambda) = N(\mathbf{0}, \lambda \mathbf{I})$, 其中 λ 为噪声方差。根据以上的假设, 利用贝叶斯规则很容易获得后验分布, 其也为一高斯分布。当所有的未知参数(即 $\{\gamma_i\}_{i=1}^M, \lambda$) 都被估计出来后, \mathbf{x} 的最大后验估计 (Maximum A Posterior) 由这个高斯分布的均值给出。而这些未知参数可以由第二类最大似然估计 (Type II Maximum Likelihood) 获得[Tipping2001, MacKay1992]。

在以上的 SBL 框架中, 我们把 γ_i 作为一未知的确定性参数, 而没有把它视为一随机变量从而进一步假设它的先验分布。事实上, 这等同于假设 γ_i 的先验分布是一个 non-informative prior。Wipf 和 Rao 已从理论上证明, 这种 SBL 框架可以获得真正的解 (即最稀疏的解) [Wipf2004], 而若对 γ_i 赋予一个非 non-informative prior, 有可能导致算法的不稳定或者解的不正确[Wipf_PhDThesis]。另外也需注意到, Tipping 提出的 SBL 算法[Tipping2001] 是假定 x_i 的 precision (即方差的倒数) 具有一参数化的 Gamma prior, 而这些参数最终被固定为某些特殊的值使得 x_i 具有一 improper prior, 即 $p(x_i) \propto 1/|x_i|$ 。这种 prior 类似于 Laplace prior, 起着促进稀疏解的作用。通过比较 Wipf 和 Rao 的 SBL 算法和 Tipping 的算法, 我们不难发现, 前者的 SBL 算法恰好是后者的 SBL 算法取该 improper prior 的形式。从这个角度也不难理解为什么前者的 SBL 算法可以获得稀疏解。

除了 Tipping, Wipf 等人的 SBL 算法外, 还有其它一些 SBL 算法赋予 x_i 的 precision 其它的分布, 或者假设 x_i 的先验分布为一 Laplace prior[BCS_laplace]。这些算法多数情况下无法证明其全局解是真正稀疏解 (即最稀疏解), 或者本身稳定性存在问题, 不能保证良好的收敛性能。值得注意的是, 赋予 x_i 不同的先验分布并不能导致相应的 SBL 算法在实际应用中具有明显的优势。这是因为大多数实际问题都和理想的感知压缩模型相去甚远, 比如感知矩阵 (sensing matrix) 的列与列之间具有强相关性, 噪声很强, \mathbf{x} 并不是非常稀疏等等。在这些情况下, 不少参数的估计将会有较大的误差, 从而导致最终的解具有较大的误差。最明显的是, 绝大多数 SBL 算法对噪声方差 λ 的估计都不有效, 尤其是当感知矩阵的列与列之间具有强相关性且噪声很大的时候。而对该方差估计的准确性对 x 的估计的准确性影响非常大。Zhang 和 Rao 最近给出了噪声方差的另外一个学习规则 [Zhang_IEEE2011]。试验表明该学习规则可以获得更加鲁棒的效果。

事实上要想在实际中获得更好的结果, 充分利用解的自身结构信息是更加有效的策略。接下来我们将介绍利用解的空间结构信息和时序结构信息的 SBL 算法。特别的, 我们将介绍如何利用解的各个元素

之间的相关性来提升算法的性能。

3 利用解的结构信息的稀疏贝叶斯学习

3.1 解的空间信息和块稀疏贝叶斯学习

解的空间信息是指在模型(1)中解向量 \mathbf{x} 具有某些结构。最常见的结构是块结构 (block structure), 或称为组群结构 (group structure) [groupLasso, ModelCS, Eldar2010BSS], 即

$$\mathbf{x} = [\underbrace{x_1, \dots, x_{d_1}}_{\mathbf{x}_1^T}, \dots, \underbrace{x_{d_{g-1}+1}, \dots, x_{d_g}}_{\mathbf{x}_g^T}]^T \quad (3)$$

基于这个块划分的基本压缩感知模型 (即公式(1)(3)) 称为块稀疏模型 (Block Sparse Model)。在这个模型中, 解向量 \mathbf{x} 可以划分为 g 个块结构 (每个块结构包含的元素有多有少), 而 \mathbf{x} 的非零的元素则聚集在少数几个块内。基于这个模型, 目前已经有了不少算法, 比如 Group Lasso [groupLasso], Block-OMP [Eldar2010BSS], Block-CoSaMP [ModelCS]等等。遗憾的是, 很少有算法考虑每个块内的元素之间的相关性 (幅值的相关性)。为方便, 下面我们称该相关性为块内相关性(Intra-Block Correlation)。

块内相关性之所以还没有引起重视, 是因为在大多数情况下目前已有的算法并没有显示出其性能受到该相关性的影响。块内相关性对算法性能的影响直到最近才被Zhang和Rao通过提出块稀疏贝叶斯学习 (Block Sparse Bayesian Learning, BSBL) 而发现[Zhang_TSP2012], 并被成功的运用到非稀疏生理信号的无线传输 [Zhang_TBME2012a, Zhang_TBME2012b]。

在BSBL中, 每一个块 \mathbf{x}_i 被假设为满足一多元高斯分布:

$$p(\mathbf{x}_i) = N(\mathbf{0}, \gamma_i \mathbf{B}_i) \quad (3)$$

其中 \mathbf{B}_i 为一未知的正定矩阵, 用于对该块内的元素之间的相关结构进行建模, 而 γ_i 为一未知的参数,

用于决定该块是否为 $\mathbf{0}$ 。类似于基本的 SBL 框架, 当 $\gamma_i = 0$, 相应的块 $\mathbf{x}_i = \mathbf{0}$ 。这样的 prior 可以认为是一种结构化的 Automatic Relevance Prior。由于 automatic relevance determination(ARD)机制, 在算法学习过程中大多数 γ_i 最终为 0 或者趋近于 0, 从而促成了了解的块稀疏性(Block Sparsity)。同样, 假设噪声服从 $p(\mathbf{v}; \lambda) = N(\mathbf{0}, \lambda \mathbf{I})$ 。这样我们可以利用贝叶斯规则得到 \mathbf{x} 的后验分布。利用第二类最大似然估计可以估计出各种参数, 从而最终得到 \mathbf{x} 的最大后验估计值。

Zhang 和 Rao 证明 [Zhang_IEEE2011], 在无噪情况下 BSBL 的全局解即是真正最稀疏解; 而无论 \mathbf{B}_i 的值是多少都不影响这一结论。事实上, \mathbf{B}_i 的值仅仅只影响算法的局部解的性质, 即算法收敛到局部解的概率。这一结论带来了极大的好处, 那就是我们可以灵活采用一些策略来规范化 (regularize) \mathbf{B}_i 的估计从而克服 overfitting, 而无须担忧是否会影响到算法的全局解的性质。在 [Zhang_IEEE2011, Zhang_TSP2012, Zhang_ICASSP2012] 中多种规范化策略被提出来。比如当每个块包含有同样数目的元素时, 我们可以平均所有的 \mathbf{B}_i 得到一矩阵 \mathbf{B} , 把它作为每个 \mathbf{B}_i 的最终估计值 [Zhang_TSP2012]。我们还可以进一步对 \mathbf{B} 做规范化, 使 $\mathbf{B} \leftarrow \mathbf{B} + \eta \mathbf{I}$, 其中 η 为一正的常数 [Zhang_IEEE2011]。试验表明,

存在若干中规范化策略，均可获得类似的性能。

从算法层次上来说，BSBL 揭示了在标准的压缩感知试验条件下一个有意思的现象[Zhang_TSP2012, RaoZhangJin2012]：当算法忽略块内相关性时（即所有的矩阵 \mathbf{B}_i 都被强制为单位矩阵），无论块内相关性是大是小，算法的性能并不发生显著的变化；当算法利用块内相关性时（即运行矩阵 \mathbf{B}_i 的学习规则），算法的性能随块内相关性的增大而提高。考虑到目前绝大多数算法都没有利用块间相关性，我们不难得到一个启示：通过改进已有的算法（比如 Group Lasso）使其可以利用块间相关性，我们可以进一步提升该算法的性能。

事实上，在 [Zhang_TSP2012] 中，Zhang 和 Rao 揭示了 BSBL 与 Group Lasso 等许多算法的联系。Zhang 和 Rao 证明，BSBL 的代价函数（cost function）等价于下面一迭代加权算法：

$$\mathbf{x}^{(k+1)} = \arg \min_{\mathbf{x}} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda \sum_{i=1}^g w_i^{(k)} \sqrt{\mathbf{x}_i^T \mathbf{B}_i^{-1} \mathbf{x}_i} \quad (4)$$

其中 k 为迭代次数， w_i 为一加权因子，其值取决于上一次迭代的结果。显然，当所有的 \mathbf{B}_i 为单位矩阵，且所有的加权 w_i 均相同且恒为一常数时，该算法即为 Group Lasso 的最常见形式。这一联系提供了如何

改进已有的基于 $\sum_{i=1}^g \|\mathbf{x}_i\|_p^q$ 惩罚项的算法。感兴趣的读者可以参考 [Zhang_ICML2011] 了解如何改进类

似 Group Lasso 的算法（即 $p = 2, q = 1$ ），或者参考 [Zhang_ICASSP2011] 了解如何改进迭代加权 L2 最小化算法（ $p = 2, q = 2$ ）。另外一方面，（4）也提供了如何提高 BSBL 算法的速度的一种方案。试

验表明，在大多数情况下只需要迭代 3 – 5 次便可收敛，而每一次迭代也即运行一次 Group Lasso。考虑到目前 Group Lasso 的运算速度不断提升，采用这种迭代方式的 BSBL 算法的速度也将同步提升。

注意到在公式（3）中块的划分是已知的。在一些情况下这种划分是无法获知的。对于这种情况，文献[Zhang_ICASSP2012, Zhang_TSP2012]提出了一种扩展 BSBL 框架。根据这种框架，未知块划分的情况被转化为一种简单的 BSBL 模型，而在这种 BSBL 模型中每个块的长度成为一种规范化参数 (regularization parameter)。试验表明这种扩展 BSBL 框架在噪声情况下尤为有效。

值得一提的是，即使块结构未知，BSBL 框架在有些情况下也仍然适用。在 [Zhang_TBME2012a, Zhang_TBME2012b] 中，BSBL 被用于恢复非稀疏的但是具有相关结构的生理信号，成功的解决了压缩传感应用在生理信号的低能耗无线传输的瓶颈问题。在试验部分我们将给出一个代表性的例子。

3.2 解的时序信息

在一些应用中（比如波达方向估计，源定位，雷达探测），在相继的时刻（假设为时刻 t_1, t_2, \dots, t_L ）我们可以建立一系列基本的压缩传感模型： $\mathbf{y}(t_1) = \mathbf{A}\mathbf{x}(t_1) + \mathbf{v}(t_1)$ ， $\mathbf{y}(t_2) = \mathbf{A}\mathbf{x}(t_2) + \mathbf{v}(t_2)$ ， \dots ， $\mathbf{y}(t_L) = \mathbf{A}\mathbf{x}(t_L) + \mathbf{v}(t_L)$ 。在这些模型里，感知矩阵 \mathbf{A} 均一样，每个解向量 $\mathbf{x}(t_n)$ ($n = 1, \dots, L$) 的非零元素的位置也一样（但幅值可能不一样）。我们可以把这 L 个模型

合并到一起，写成矩阵的形式：

$$\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{V}$$

其中 $\mathbf{Y} = [\mathbf{y}(t_1), \dots, \mathbf{y}(t_L)]$ ， $\mathbf{X} = [\mathbf{x}(t_1), \dots, \mathbf{x}(t_L)]$ 。根据前面的假设， \mathbf{X} 仅有少数几行是非零行，而绝大多数行都为零行。这样一种模型称为多观测向量模型（Multiple Measurement Vector Model, MMV）[Cotter2005]。相比于仅使用一个观测向量的模型（即这 L 个基本压缩传感模型中的任意一个），MMV 模型能够更加准确的估计出非零行的位置（或者说每个解向量 $\mathbf{x}(t_n)$ 的非零元素的位置），从而更准确的获得最终的解 [RaoZhangJin2012, Cotter2005, Eldar2010, Tang2011]。

目前已有不少针对该模型的 MMV 算法，包括基于这种 MMV 模型的 SBL 算法 [Wipf2007]。遗憾的是，绝大多数算法都忽略了 \mathbf{X} 中每一个非零行内的元素之间的相关性。试验显示 [Zhang_IEEE2011]，如果算法中不考虑这些行内的相关性，算法的性能将会极大的降低。为利用这些行内的相关性，Zhang 和 Rao 注意到块稀疏模型与 MMV 模型的联系，把 BSBL 框架用于 MMV 模型，而行内的相关结构则由 BSBL 框架中的矩阵 \mathbf{B}_i 来建模，从而得到了利用时序结构的 SBL 算法 [Zhang_IEEE2011]。这种利用时序结构的 SBL 算法相比于其它 MMV 算法具有明显的优势。

需要指出的是，利用时序结构信息的 MMV 算法不但可以用于信号处理领域，还可以用于机器学习领域，作为一种多任务学习（multi-task learning）算法用于特征提取。在 [Wan2012] 中，利用时序结构的 SBL 算法与常见的多任务学习算法的联系被揭示出来；前者可以视为一种具有自适应核（adaptive kernels）的多任务学习算法，可以自动捕捉数据的内在结构，因而较传统的多任务学习算法以及依靠人工设定核的多任务学习算法更加优越。

3.3 解的空-时结构信息和时变信息

BSBL 框架提供了一种思路来灵活的合并解的空间结构信息和时序结构信息。此外，利用 BSBL 框架还可以有效利用解的时变结构信息。限于篇幅我们不在这里做介绍。对这两个部分感兴趣的读者可以参考综述 [RaoZhangJin2012]。

4 试验仿真

4.1 非稀疏的生理信号的压缩传感及其在医学远程监控中的应用

经由无线体域网（Wireless Body-Area Network, WBAN）的生理信号的远程监控是目前医疗通讯领域的一个主要发展方向。最近开始有学者提出把压缩传感技术用于这一领域 [Mamaghanian2011, Dixon_BME2012, Eduardo2010]。压缩传感技术用于这一领域有很明显的优势：（1）当采用元素仅为0或者1的稀疏矩阵（binary sparse matrix）为传感矩阵时，压缩传感可以比传统的小波压缩技术更加减少无线体域网的能量的损耗 [Mamaghanian2011]。而减少能量损耗是无线体域网研究的一个核心问题 [Milenkovic2006, Calhoun2012]。（2）从压缩质量上来看，压缩传感和小波压缩有类似的压缩率和恢复质量。

但是压缩传感技术在这一领域取得的成果仅仅局限在及少数几种特别稀疏的生理信号（比如非常干净无噪的成人心电图信号），或者局限在少数几种特殊用途上。压缩传感技术还不能用于更多的生理信号（比如胎儿心电图信号，脑电图信号），以及还不能用于更为广泛的用途。这主要是因为：

（1）绝大多数的生理信号是非稀疏的信号，即使被投影到变换域上（比如小波域）仍旧不够稀疏 [Zhang_TBME2012a]。尽管这些生理信号可以认为是一种可压缩（compressive）的信号（即：虽然信号的所有元素都非零，但是只有少数一些元素显著非零，而绝大部分元素接近为零），但是压缩传感算法只能保证恢复信号中的显著的非零元素，而接近为零的元素无法准确恢复。但是对于医学远程监控

来说,生理信号中一些接近为零的元素常常是重要的诊断指标。比如在胎儿心电图,感兴趣的胎儿心跳信号常常非常微弱,接近于零,甚至很难被肉眼察觉。

(2)生理信号被传输到远程终端之后,通常还需要经过一系列的信号处理和模式识别。比如当多通道脑电信号被传输到远程终端之后,通常还需要进行独立分量分析 [Makeig1996],从脑电信号中提取感兴趣的独立分量,然后再进行模式识别和医疗诊断。这些应用要求压缩传感算法能够完整的把生理信号压缩并恢复出来,而不损失任何细微的波形特征,否则无法进行后继的信号处理和医疗诊断 [Zhang_TBME2012b]。

以上两个因素要求压缩传感算法能够以极高质量恢复非稀疏信号或者可压缩信号。而就我们所知,除了BSBL算法外,目前还没有任何一个压缩传感算法做到这一点。接下来我们将给出BSBL算法解决这一瓶颈问题的几个代表性例子。这些例子取自 [Zhang_TBME2012a, Zhang_TBME2012b]。

首先给出用BSBL算法实现胎儿心电信号的压缩传感。作为示例,我们选用BSBL框架中的BSBL-BO算法 [Zhang_TSP2012]。

图1(a)给出一段采集到的原始胎儿心电图信号(数据来自: <http://www.esat.kuleuven.ac.be/sista/daisy>)。其中两个最明显的峰为母亲心电信号的QRS成分。另外有两个小峰分别出现在第80个采样点和第200个采样点附近,他们分别对应胎儿心电信号的QRS成分,也是我们真正感兴趣的成分。令这段原始信号为 \mathbf{x} 。我们任意生成一个125x250维的binary稀疏传感矩阵,记为 \mathbf{A} 。 \mathbf{A} 中每一列仅有15个元素为1,其余元素都为0。由 \mathbf{x} 和 \mathbf{A} 便得到了压缩的信号 $\mathbf{y} = \mathbf{Ax}$ 。 \mathbf{y} 经由无线体域网传到智能手机并

经过互联网最终传输到远程医疗终端。在该终端,BSBL-BO算法由压缩的信号 \mathbf{y} 和共享的传感矩阵 \mathbf{A} 恢复原始的胎儿心电信号 \mathbf{x} 。图1(b)是恢复的结果。显然,胎儿心电信号的QRS成分被很好的恢复出来。图1(c)给出当BSBL-BO不利用块内相关性的结果。此时,胎儿心电信号的QRS成分不能完整的恢复出来,这说明了利用块内相关性的重要性。

在这个例子中,BSBL-BO算法假设 \mathbf{x} 可以划分为10个块,每个块都包含25个元素。显然这种块划分并不符合 \mathbf{x} 的真实块结构(见图1(a))。但是事实上BSBL算法并不需要这种块划分和信号的真实块划分一致;这种假设的块划分可以视为一种规范化(regularization)。尽管假设的不同块划分会产生性能的一些微小变化,但是算法的性能对块划分并不敏感 [Zhang_TBME2012a, Zhang_TSP2012, Zhang_ICASSP2012]。这一现象其实也可以借助贝叶斯估计理论从理论上解释。

作为比较,我们挑选了六种代表性的压缩传感算法来恢复该原始信号。它们分别是CoSaMP [Needell2009], Elastic Net [Zou2005], SL0 [Mohimani2009], Block-OMP [Eldar2010BSS], StructOMP [Huang2009], 和BM-MAP-OMP [Faktor2012]。其中前三种不利用任何结构信息,后三种利用了块结构信息。具体的参数设置见 [Zhang_TBME2012a]。图2给出了这些算法恢复的结果。显然,所有的算法都不能成功恢复出胎儿心跳信号的QRS峰。

这个例子清晰的说明了BSBL算法相比于其它算法的显著优势,同时也说明了利用块内相关性的好处。其实BSBL算法的优势远不止这些。我们前面已说过,生理信号在传输到远程终端后,通常还要经过一系列的信号处理和模式识别。一旦恢复的信号中所引入的畸变稍大,则会导致后继的信号处理产生极大的误差。为了研究BSBL算法是否真正有助于远程监控和医疗诊断,在 [Zhang_TBME2012a, Zhang_TBME2012b] 中分别对胎儿心电信号和脑电信号进行了一系列试验。试验中,BSBL-BO首先恢复出多通道的生理信号(胎儿心电图,脑电信号等),然后在对恢复的多通道信号进行高级信号处理(比如独立分量分析),得到最终信号。然后,对原始采集到的多通道生理信号进行同样的高级信号处理,得到另外一组最终信号。最后比较两次获得的最终信号。结果显示这两次获得的最终信号几乎一模一样。这个结果说明了BSBL-BO算法在恢复非稀疏的生理信号中,可以保证极高的恢复质量,对远程监控和医疗诊断具有实际的价值。我们预期在不远的将来BSBL-BO或者其它类似的BSBL算法将在这一领域获得广泛的应用。

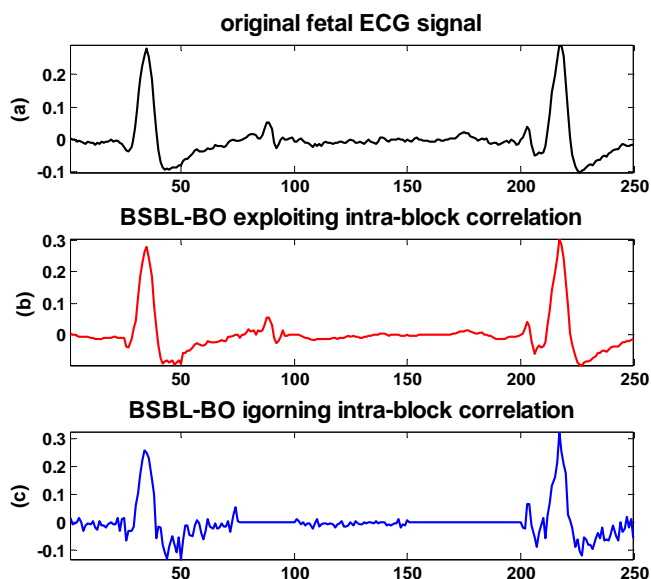


图1.原始的胎儿心电图和用BSBL-BO算法恢复的胎儿心电图。其中(b)是BSBL-BO利用块内相关性的结果。(c)是不利用块内相关性的结果。

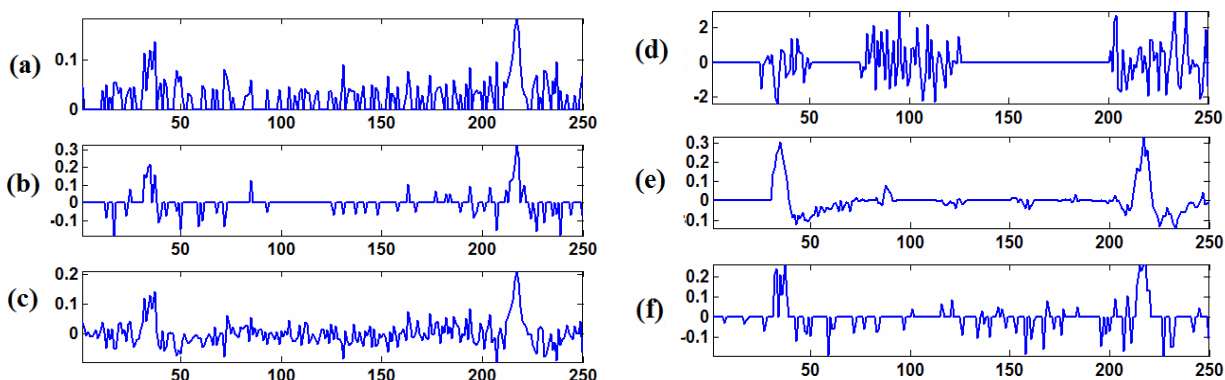


图2.六种代表性的算法恢复的结果: (a) CoSaMP, (b) Elastic Net, (c) SL0, (d) Block-OMP, (e) StructOMP, (f) BM-MAP-OMP。

参考文献

- [Tipping2001] M.E. Tipping. Sparse Bayesian learning and the relevance vector machine [J]. Journal of Machine Learning Research, 2001, 1: 211-244
- [Wipf2004] D.P. Wipf, B.D. Rao. Sparse Bayesian learning for basis selection [J]. IEEE Trans. on Signal Processing, 2004, 52(8): 2153-2164
- [Ji2008] S. Ji, Y. Xue, L. Carin. Bayesian compressive sensing [J]. IEEE Trans. on Signal Processing, 2008, 56(6): 2346-2356
- [Donoho2003] D.L. Donoho, M. Elad. Optimally sparse representation in general (nonorthogonal) dictionaries via L1 minimization [J]. PNAS, 2003, 100(5): 2197-2202

- [Wipf_NIPS2011] D.P. Wipf. Sparse estimation with structured dictionaries [C]. Advances in Neural Information Processing Systems 24, 2011: 2016-2024
- [Wipf2010] D. Wipf, S. Nagarajan. Iterative reweighted L1 and L2 methods for finding sparse solutions [J]. IEEE Journal of Selected Topics in Signal Processing, 2010, 4(2): 317-329
- [Candes2008] E.J. Candes, M.B. Wakin, S.P. Boyd. Enhancing sparsity by reweighted L1 minimization [J]. Journal of Fourier Analysis and Applications, 2008, 14: 877-905
- [ModelCS] R.G. Baraniuk, V. Cevher, M.F. Duarte, C. Hegde. Model-based compressive sensing [J]. IEEE Trans. on Information Theory, 2010, 56(4): 1982-2001
- [Zhang_IEEE2011] Z. Zhang, B.D. Rao. Sparse signal recovery with temporally correlated source vectors using sparse Bayesian learning [J]. IEEE Journal of Selected Topics in Signal Processing, 2011, 5(5): 912-926
- [Zhang_TSP2012] Z. Zhang, B.D. Rao. Extension of SBL algorithms for the recovery of block sparse signals with intra-block correlation [J]. submitted to IEEE Trans. on Signal Processing, 2012. Available: arXiv:1201.0862
- [Wan2012] J. Wan, Z. Zhang, J. Yan, T. Li, B.D. Rao, et al. Sparse Bayesian multi-task learning for predicting cognitive outcomes from neuroimaging measures in Alzheimers disease [C]. CVPR 2012
- [Zhang_TBME2012a] Z. Zhang, T.-P. Jung, S. Makeig, B.D. Rao. Low energy wireless body-area networks for fetal ECG telemonitoring via the framework of block sparse Bayesian learning [J]. submitted to IEEE Trans. on Biomedical Engineering, 2012. Available: arXiv:1205.1287v1
- [Zhang_TBME2012b] Z. Zhang, T.-P. Jung, S. Makeig, B.D. Rao. Compressed sensing of non-sparse physiological signals for telemonitoring via wireless body-area networks with low energy consumption and inexpensive hardware [J]. submitted to IEEE Trans. on Biomedical Engineering, 2012
- [Zhang_PIEEE2012] S. Makeig, C. Kothe, T. Mullen, N. Bigdely-Shamlo, Z. Zhang, K. Kreutz-Delgado. Evolving signal processing for brain-computer interfaces [J]. Proceedings of the IEEE, 2012, 100(Special Centennial Issue): 1567-1584
- [ARD1996] R.M. Neal. Bayesian learning for neural networks [M]. Springer, 1996
- [MacKay1992] D.J.C. MacKay. The evidence framework applied to classification networks [J]. Neural Computation, 1992, 4(5): 720-736
- [Wipf_PhDThesis] D.P. Wipf. Bayesian methods for finding sparse representations. Ph.D. Thesis. University of California, San Diego, 2006
- [BCSlaplace] S.D. Babacan, R. Molina, A.K. Katsaggelos. Bayesian compressive sensing using Laplace priors [J]. IEEE Trans. on Image Processing, 2010, 19(1): 53-63
- [groupLasso] M. Yuan, Y. Lin. Model selection and estimation in regression with grouped variables [J]. Journal of Royal Statistical Society, 2006, 68: 49-67
- [Eldar2010BSS] Y.C. Eldar, P. Kuppinger, H. Bolcskei. Block-sparse signals: uncertainty relations and efficient recovery [J]. IEEE Trans. on Signal Processing, 2010, 58(6): 3042-3054
- [Zhang_ICASSP2012] Z. Zhang, B.D. Rao. Recovery of block sparse signals using the framework of block sparse Bayesian learning [C]. Proc. of the 37th International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2012), 2012
- [RaoZhangJin2012] B.D. Rao, Z. Zhang, Y. Jin. Sparse signal recovery in the presence of intra-vector and inter-vector correlation [C]. International Conference on Signal Processing and Communications, 2012
- [Zhang_ICML2011] Z. Zhang, B.D. Rao. Exploiting correlation in sparse signal recovery problems: multiple measurement vectors, block sparsity, and time-varying sparsity [C]. ICML 2011 Workshop on Structured Sparsity: Learning and Inference, 2011. Available: <http://arxiv.org/pdf/1105.0725>
- [Zhang_ICASSP2011] Z. Zhang, B.D. Rao. Iterative reweighted algorithms for sparse signal recovery with

temporally correlated source vectors [C]. Proc. of the 36th International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2011), 2011

[Cotter2005] S.F. Cotter, B.D. Rao, K. Engan, K. Kreutz-Delgado. Sparse solutions to linear inverse problems with multiple measurement vectors [J]. IEEE Trans. on Signal Processing, 2005, 53(7): 2477-2488

[Eldar2010] Y.C. Eldar, H. Rauhut. Average case analysis of multichannel sparse recovery using convex relaxation [J]. IEEE Trans. on Information Theory, 2010, 56(1): 505-519

[Tang2010] G. Tang, A. Nehorai. Performance analysis for sparse support recovery [J]. IEEE Trans. on Information Theory, 2010, 56(3): 1383-1399

[Wipf2007] D.P. Wipf, B.D. Rao. An empirical Bayesian strategy for solving the simultaneous sparse approximation problem [J]. IEEE Trans. on Signal Processing, 2007, 55(7): 3704-3716

[Mamaghanian2011] H. Mamaghanian, N. Khaled, D. Atienza, and P. Vandergheynst. Compressed sensing for real-time energy-efficient ECG compression on wireless body sensor nodes [J]. IEEE Trans. on Biomedical Engineering, 2011, 58(9): 2456-2466

[Dixon_BME2012] A.M.R. Dixon, E.G. Allstot, D. Gangopadhyay, D.J. Allstot. Compressed sensing system considerations for ECG and EMG wireless biosensors [J]. IEEE Trans. on Biomedical Circuits and Systems, 2012, 6(2): 156-166

[Eduardo2010] C. Eduardo, P. O. Adrian, S. Pedro, et al. Implementation of compressed sensing in telecardiology sensor networks [J]. International Journal of Telemedicine and Applications, 2010 (2010)

[Milenkovic2006] A. Milenkovic, C. Otto, E. Jovanov. Wireless sensor networks for personal health monitoring: Issues and an implementation [J]. Computer communications, 2006, 29(13-14): 2521-2533

[Calhoun2012] B. Calhoun, J. Lach, J. Stankovic, et al. Body sensor networks: a holistic approach from silicon to users [J]. Proceedings of the IEEE, 2012, 100(1): 91-106

[Makeig1996] S. Makeig, A.J. Bell, T.P. Jung, T.J. Sejnowski, et al. Independent component analysis of electroencephalographic data [C]. Advances in neural information processing systems, 1996: 145-151

[Needell2009] D. Needell, J.A. Tropp. CoSaMP: iterative signal recovery from incomplete and inaccurate samples [J]. Applied and Computational Harmonic Analysis, 2009, 26(3): 301-321

[Zou2005] H. Zou, T. Hastie. Regularization and variable selection via the elastic net [J]. Journal of the Royal Statistical Society: Series B, 2005, 67(2): 301-320

[Mohimani2009] H. Mohimani, M. Babaie-Zadeh, C. Jutten. A fast approach for overcomplete sparse decomposition based on smoothed L0 norm [J]. IEEE Trans. on Signal Processing, 2009, 57(1): 289-301

[Huang2009] J. Huang, T. Zhang, D. Metaxas. Learning with structured sparsity [C]. Proc. of the 26th Annual International Conference on Machine Learning, 2009:417-424

[Faktor2012] T. Faktor, Y. Eldar, M. Elad. Exploiting statistical dependencies in sparse representations for signal recovery [J]. IEEE Trans. on Signal Processing, 2012