Um Breve Estudo sobre Análise de Componentes Esparsos

Everton Z. Nadalin[1], Ricardo Suyama[2], Romis Attux[1]

1 - Departamento de Engenharia de Computação e Automação Industrial (DCA)
2 – Departamento de Microonda e Óptica (DMO)

Faculdade de Engenharia Elétrica e de Computação (FEEC)
Universidade Estadual de Campinas (Unicamp)
Caixa Postal 6101, CEP 13083-970 – Campinas, SP, Brasil

{nadalin, attux}@dca.fee.unicamp.br, rsuyama@dmo.fee.unicamp.br

Abstract – In this work, we present a discussion concerning some fundamental aspects of sparse component analysis (SCA), a method that has been increasingly employed to solve blind source separation (BSS) problems, especially when there are more sources than sensors. In addition to providing an overview on BSS and SCA, we try to point out relevant paths for future research and analyze the idea of compressive sensing, which seems to raise interesting perspectives for practical application.

Keywords – unsupervised signal processing, blind source separation, sparse component analysis, compressive sensing.

1. Introdução

Devido ao barateamento de processadores e à conseqüente revolução que os sistemas embarcados vêm causando, novas técnicas de processamento de sinais têm sido requisitadas. Sempre buscando novos desafios, a teoria de processamento não-supervisionado tem cada vez mais incorporado técnicas e soluções capazes de fazer efetivo uso de informação *a priori* sobre os sinais que se deseja estimar. Nesse contexto, uma técnica que vem ganhando espaço é a análise de componentes esparsos (SCA – *sparse component analysis*), a qual encontra aplicação em áreas como separação cega de fontes (BSS – *blind source separation*) [1] e *compressive sensing* [2].

Intuitivamente, a idéia de esparsidade de um sinal é relativamente simples: um sinal esparso é fundamentalmente aquele em que predominam, no domínio temporal ou em algum outro domínio relevante, valores nulos ou próximos de zero, sendo que poucos componentes possuem a maior parte da energia do sinal. Porém, quando tentamos quantificar essa idéia, surgem diversas interpretações.

Este artigo pretende mostrar algumas formas existentes de quantificação da idéia de esparsidade em sinais esparsos, bem como fazer uma discussão sobre a aplicação de SCA em BSS.

O artigo está dividido da seguinte forma: na seção 2, é feita uma descrição da idéia de SCA; na seção 3, discute-se a aplicação de SCA em BSS; enquanto a seção 4 contém uma discussão final.

2. Quantificando a esparsidade de um sinal

O conceito de esparsidade nos remete sinais que possuem uma grande quantidade de regiões com valores nulos ou quase nulos em algum domínio, ou seja, sinais em que toda a informação está concentrada em uma quantidade pequena de valores que representam aquilo que se analisa. Uma maneira, neste caso, de se quantificar o grau de esparsidade é contar quantas amostras ou quantos coeficientes possuem valores não-nulos.

Este tipo de medição evoca a chamada norma ℓ_0 [3], porém, ela falha na maioria dos casos práticos pelo simples fato de que quase nunca os coeficientes possuem valor nulo, havendo em geral ruído ou um valor residual. Para fugir disto, alguns trabalhos adotaram a norma $\ell_{0\varepsilon}$, que leva em conta um valor ε residual para definir uma espécie de limiar de "nulidade".

Outro tipo de critério de otimização é a chamada norma ℓ_1 [3], definida pela soma dos valores absolutos das amostras do sinal. Uma vantagem desta abordagem é que o uso dessa norma permite, em alguns casos, que se lance mão de metodologias de otimização convexa, as quais, muitas vezes, permitem que certos passos de um procedimento de separação se vinculem a soluções em forma fechada [4].

É importante frisar que, além do uso direto de critérios de otimização que procuram quantificar a esparsidade do sinal, existem outras possibilidades, como as discutidas no trabalho de Hurley et al. [3]. Uma dessas possibilidades, aliás, relaciona-se ao o emprego adicional de uma estatística de ordem superior amplamente usada em análise de componentes independentes (ICA, - *independent component analysis*), a curtose.

Ainda em [3], é realizada uma discussão sobre a qualidade de cada critério no âmbito de um conceito mais amplo de esparsidade. No artigo, seis são os atributos considerados para se medir a esparsidade de um sinal, quatro deles derivados das leis de Dalton relacionadas à distribuição de renda [5] e outros dois descritos em [6]. Estes seis atributos são descritos abaixo.

1 – quanto mais a energia for distribuida entre os coeficientes, menor a esparsidade do sinal.

2 – multiplicar cada coeficiente por uma mesma constante não altera a esparsidade.

3 – o acréscimo de uma mesma constante a cada coeficiente diminui a esparsidade do sinal.

4 – a união de dois sinais idênticos não altera a esparsidade.

5 – conforme um determinado coeficiente vai ficando com cada vez mais energia, a esparsidade do sinal vai aumentando.

6 – a inclusão de coeficientes de valor nulo o sinal aumenta a esparsidade.

Apesar de estes atributos parecerem intuitivamente bem consistentes, Hurley et al. mostram que, das 15 medidas de esparsidade mais comumente utilizadas, somente o índice de Gini [7] obedece a todos os seis. Porém, é importante perceber que nem todos estes critérios precisam ser levados em consideração quando tentamos resolver os problemas de separação de fontes e compressão de sinais, os mais comuns em que se utilizam técnicas de SCA. Para que tal fato fique mais claro, estes dois problemas serão explicados no próximo item.

3. Separação cega de fontes

Nos últimos anos, o estudo do problema de separação cega de fontes (BSS) se consolidou como um dos pilares da teoria de processamento não-supervisionado de sinais. Um bom exemplo de problema de BSS é o comumente chamado de *cocktail party problem* [8], que pode ser exemplificado da seguinte forma: existe uma sala com *n* microfones e *m* pessoas falando, e, tendo acesso apenas aos dados dos microfones, queremos encontrar o sinal produzido por cada pessoa separadamente.

Matematicamente, sendo **s**(n) o vetor dos sinais das fontes de informação (no exemplo acima, esse vetor seria composto pelos sinais de voz de todos os presentes) e **x**(n) o vetor de sinais captados pelos sensores (também chamado de vetor de observações), podemos, supondo que a mistura dos sinais é linear e não envolve espalhamento temporal, escrever

$$\mathbf{x}(n) = \mathbf{A}\mathbf{s}(n)$$

onde **A** é chamada de matriz de mistura do sinal, que, tipicamente, é suposta quadrada e inversível (voltaremos a essa hipótese mais adiante). Sendo assim, o problema acima descrito corresponde a estimar as fontes tendo acesso somente às observações. Para realizar esse processo de separação, adota-se, tipicamente, um sistema separador com estrutura análoga à da mistura, ou seja:

$$\mathbf{y}(n) = \mathbf{W}\mathbf{x}(n)$$

sendo **W** a matriz de separação. Idealmente, o que se busca é obter um vetor **y**(n) que seja igual a **s**(n) a menos de fatores de escala e de uma permutação. Portanto, a matriz de separação ideal seria:

$$\mathbf{W} = \mathbf{P.D.A}^{-1}$$

onde **P** é uma matriz de permutação e **D** uma matriz diagonal.

O modelo acima descrito é, sem dúvida, o mais estudado em BSS, e, sob a hipótese de que as fontes são mutuamente independentes, o problema dele originado vem sendo resolvido com sucesso por meio de diversas técnicas de ICA. No entanto, tal modelo se baseia em algumas hipóteses restritivas que podem não ser válidas em determinadas aplicações práticas. Um exemplo é o fato de que a matriz **A** é considerada quadrada e inversível (ou, eventualmente, de posto completo), o que nem sempre é razoável, pois pode haver, por exemplo, menos sensores que fontes.

4. SCA aplicada à separação de fontes

Em boa parte dos problemas práticos não é possível garantir que haja um número de sensores ao menos igual ao número de fontes porque número de fontes não é conhecido. Neste caso, pode não ser possível aplicar um método para obter um **W** eficiente pelo simples fato de a matriz de mistura não ser inversível: tem-se, portanto, um problema indeterminado. Trabalhos mostram que, em algumas situações em que o sistema é indeterminado [1,9,10,11] e os sinais são esparsos, o fato de que nem sempre todas as fontes estarão ativas ao mesmo tempo traz interessantes perspectivas para a aplicação desta ferramenta em problemas de BSS. Isto se deve ao fato de que, nestas situações, um sistema indeterminado ser localmente determinado, sendo possível a identificação da matriz de mistura e em alguns casos até a separação das fontes. Este tipo de abordagem é comumente chamada de análise de componente esparsos (SCA).

Os primeiros trabalhos utilizando SCA em BSS [1] assumiram que no máximo uma fonte estivesse ativa em cada amostra dos sinais de mistura. Desta forma, não existindo sobreposição de fontes, é possível separá-las de forma integral, mesmo que haja mais fontes do que misturas, conforme mostrado em [9]. Neste caso, métodos como o mascaramento binário [10] já são suficientes para a separação das fontes. Em [9] é dito que, quando é exigido em cada instante que apenas uma fonte esteja ativa, o problema acaba demandando, na verdade, duas restrições. Além da esparsidade das fontes, é necessário também que exista uma ortogonalidade disjunta entre elas na mistura.

Já no artigo [11], Aissa-El-Bay et al. extendem o resultado também para fontes não-disjuntas, em que nem sempre é necessário uma fonte ativa por vez. Outros trabalhos seguem a mesma linha, considerando que o número de fontes ativas seja, na média, unitário [12], ou que, para cada fonte, exista pelo menos um instante na mistura em que somente ela esteja ativa [13]. Porém, essas garantias só se justificam em uma mistura sem ruído.

Alguns trabalhos conseguem extrapolar os resultados para situações em quem em cada instante o número de fontes não ultrapassa o número de misturas [14], porém, neste caso o processo se restringe à estimação da matriz de mistura: as fontes não chegam a ser separadas.

O trabalho [15] mostra que se, além de esparsos, os sinais forem independentes entre si (o que é verdade em boa parte dos sinais práticos), é possível estimar a matriz de mistura.

4.1 Estimação do número de fontes

Um ponto que deve ser levado em consideração é que, a partir do momento em que se trabalha com sistemas indeterminados, tem-se que tomar uma decisão sobre o conhecimento *a priori* ou não do número de fontes.

Praticamente todos os trabalhos citados consideram que o número de fontes é conhecido, porém, na maioria dos casos reais, isto não é verdade, principalmente por dois motivos: não se sabe quantas fontes de ruído existem e não é possível precisar o que o algoritmo vai considerar como sendo ou não fonte.

Os trabalhos [16,17] fazem a estimação do número de fontes a partir de uma estimação automática do número de *clusters* mais adequado para representar os sinais de mistura.

4.2 O que ainda não foi feito

Apesar do avanço que vem ocorrendo nos últimos anos, todas as soluções mostradas dependem de condições específicas das misturas, além da esparsidade dos sinais das fontes. Isso evoca a condição clássica para ICA: quando garantimos que os sinais das fontes são independentes entre si e a matriz de mistura tem posto completo em colunas, é possível realizar uma separação perfeita.

Fica agora a pergunta: sabendo que uma mistura é composta por sinais esparsos, se procurarmos otimizar algum critério que determine as componentes mais esparsas possíveis que compõe a mistura, em quais condições estaremos automaticamente separando as fontes?

Em outras palavras, tomando como base o modelo de mistura de ICA

$$\mathbf{x}(n) = \mathbf{A}\mathbf{s}(n)$$

se encontrarmos um vetor $\mathbf{q}(n)$ a partir de um critério de maximização de esparsidade que obedeça a equação

$$\mathbf{x}(n) = \mathbf{G}\mathbf{q}(n)$$

ele será automaticamente uma estimativa de $\mathbf{s}(n)$?

5. Conclusões

O problema de separação cega de fontes (BSS) é, sem dúvida, um dos pilares da moderna teoria de processamento não-supervisionado. Embora, em certos contextos práticos, tal problema possa ser resolvido de modo satisfatório para meio da análise de componentes independentes (ICA), há casos em que outros tipos de informação *a priori* sobre as fontes são necessários.

Uma possibilidade muito interessante nesse sentido é o uso da análise de componentes esparsos (SCA). Neste trabalho, buscamos apresentar a idéia geral de SCA, bem como alguns aspectos de sua aplicação no problema de BSS que julgamos relevantes e promissores.

Referências

[1] Bofill, P. and Zibulevsky, M., “Blind separation of more sources than mixtures using sparsity of their short-time Fourier transform.” In: *Proceedings of the ICA2000*, pp. 87-92.

[2] Candès, E. J., Wakin, M. B., “An Introdution to Compressive Sampling”, *IEEE Signal Processing Magazine*, vol. 25, pp. 21-30, Março 2008.

[3] Hurley, N., Rickard, S., “Comparing Measures of Sparsity”. In: *IEEE Workshop on Machine Learning for Signal Processing*, Cancún, México, 2008.

[4] Boyd, S., Vandenberghe, L.: Convex Optimization. Cambridge University Press, Nova York, 2004.

[5] Dalton, H., “The measurement of the inequity of incomes”, *Economic Journal*, vol. 30, pp. 348-361, 1920.

[6] Rickard, S., Fallon. M., “The Gini index of speech”. In: *Conference on Information Sciences and Systems*, Princeton, EUA, 2004.

[7] Gini, C., “Measurement of inequity of incomes”, *Economic Journal*, vol. 31, pp. 124-126, 1921.

[8] Hyvärinen, A., Karhunen, J., Oja, E.: Independent Component Analysis. John Wiley & Sons, Nova York, 2001.

[9] Rickard, S., “Sparse sources are separated sources”. In: *Proceedings of the 16th Annual European Signal Processing Conference*, Florence, Italy, 2006.

[10] Yilmaz, O.; Rickard, S., "Blind separation of speech mixtures via time-frequency masking," *Signal Processing, IEEE Transactions on* , vol.52, no.7, pp. 1830-1847, Julho 2004.

[11] Aissa-El-Bey, A.; Linh-Trung, N.; Abed-Meraim, K.; Belouchrani, A.; Grenier, Y., "Underdetermined Blind Separation of Nondisjoint Sources in the Time-Frequency Domain," *Signal Processing, IEEE Transactions on*, vol.55, no.3, pp.897-907, Março 2007.

[12] Georgiev, P.; Theis, F.; Cichocki, A., "Sparse component analysis and blind source separation of underdetermined mixtures," *Neural Networks, IEEE Transactions on*, vol.16, no.4, pp.992-996, Julho 2005.

[13] Kim, S.; Yoo, C.D., "Underdetermined Blind Source Separation Based on Subspace Representation," *Signal Processing, IEEE Transactions on*, vol.57, no.7, pp.2604-2614, Julho 2009.

[14] Movahedi Naini, F., Hosein Mohimani, G., Babaie-Zadeh, M., and Jutten, C. “Estimating the mixing matrix in Sparse Component Analysis (SCA) based on partial k-dimensional subspace clustering.” *Neurocomput.* 71, 10-12 (Jun. 2008), 2330-2343.

[15] Nadalin, E. Z., Suyama, R., and Attux, R. “An ICA-Based Method for Blind Source Separation in Sparse Domains.” In *Proceedings of the 8th international Conference on independent Component Analysis and Signal Separation* (Paraty, Março 15 - 18, 2009).

[16] Arberet, S., Gribonval, R., Bimbot, F.: “A Roubust Method to Count and Locate Audio Sources in a Stereophonic Linear Instantaneous Mixture.” In: In *Proceedings of the 6th international Conference on independent Component Analysis and Signal Separation*. EUA (2006).

[17] Nadalin, E. Z., Suyama, R., Attux, R., “Estimating the number of audio sources in a stereophonic instantaneous mixture.” In: *7o Congresso de Engenharia de Áudio - AES2009*, 2009, São Paulo.