

# Uma Abordagem Geométrica para Separação Cega de Fontes em um Modelo de Mistura baseado na Integral de Choquet

Gláucia Jardim Pissinelli, Leonardo Tomazeli Duarte e Yannick Deville

**Resumo**—A separação cega de fontes (BSS - *Blind Source Separation*) tem como finalidade a extração de sinais a partir de misturas observadas de tais sinais. Os modelos de misturas mais populares de BSS são os lineares. No entanto, há também grande interesse em propostas para abordar a separação de misturas não lineares. Neste trabalho, investigamos a BSS em um modelo não linear baseado na integral de Choquet. Nesta primeira abordagem, focamos no problema de identificação cega do processo de mistura (BMI - *Blind Mixture Identification*). A solução proposta se fundamenta em aspectos geométricos engendrados pelo processo de mistura, e é analisada por meio de um conjunto de experimentos numéricos.

**Palavras-Chave**—Separação Cega de Fontes, Identificação Cega do Processo de Mistura, Integral de Choquet, Misturas Não Lineares.

**Abstract**—Blind source separation (BSS) is intended to extract signals from observed mixtures of such signals. The most popular models of BSS mixtures are linear. However, there is also great interest in proposals to address the separation of nonlinear mixtures. In this paper, we investigate BSS in a nonlinear model based on the Choquet integral. In this first approach, we focus on the blind mixing process identification (BMI) problem. The proposed solution is based on geometric aspects engendered by the mixing process, and is analyzed by a set of numerical experiments.

**Keywords**—Blind Source Separation, Blind Mixture Identification, Choquet Integral, Nonlinear Mixtures.

## I. INTRODUÇÃO

O problema da separação cega de fontes (BSS, do inglês *Blind Source Separation*), foi apresentado inicialmente nos trabalhos de [1], [3] e, desde então, tem sido intensamente estudado pela comunidade de processamento de sinais [5].

Em BSS, o objetivo é recuperar sinais originais submetidos a um processo de mistura sem ter informações prévias sobre as fontes e sobre o processo de mistura em sua totalidade. No modelo mais clássico de BSS, o caso linear e instantâneo, presume-se que as misturas podem ser modeladas como combinações lineares das fontes. Essa formulação pode ser expressa matematicamente como  $\mathbf{X} = \mathbf{A}\mathbf{S}$ , onde  $\mathbf{X} \in \mathbb{R}^{M \times T}$  é uma matriz na qual cada linha corresponde a um sinal misturado. De maneira análoga  $\mathbf{S} \in \mathbb{R}^{N \times T}$  representa os sinais

fontes. As constantes  $M$ ,  $N$  e  $T$  representam os números de misturas, fontes e amostras, respectivamente. Finalmente,  $\mathbf{A} \in \mathbb{R}^{M \times N}$  corresponde à matriz de mistura. No caso em que se deseja unicamente estimar a matriz de mistura, o problema resultante é conhecido como identificação cega do processo de mistura (BMI, do inglês *Blind Mixture Identification*) [14]. Cabe destacar que, para as misturas lineares, a resolução do problema de BMI pode ser utilizada como uma primeira etapa de um método de BSS, pois, feita a identificação do processo de mistura, é possível lançar mão, por exemplo, de uma abordagem de quadrados mínimos para estimação das fontes.

Novas propostas têm surgido para abordar a BSS em misturas não lineares, como por exemplo, os modelos de mistura (*Post-Nonlinear*) [11], compostos por uma seção linear seguida por uma seção não linear. Outros modelos que também lidam com modelos não lineares são discutidos em [9], [14].

No caso de misturas não lineares, um desafio importante refere-se à interpretabilidade dos modelos. De fato, enquanto que, no caso linear, um procedimento simples para analisar a influência de cada fonte em uma dada mistura é verificar os coeficientes da matriz de mistura  $\mathbf{A}$ , tal procedimento, no caso não linear, não é trivial, pois, justamente, a relação não linear entre as observações (misturas) e fontes impede qualquer tipo de inspeção direta sobre as influências das fontes no processo de mistura. Cabe destacar ainda que essa questão da busca por interpretabilidade em modelos não lineares vem sendo um dos principais desafios em análise de dados e aprendizado de máquina [15].

Outro tema central do trabalho é o conceito de integral de Choquet, que tem sido utilizado para resolver problemas de agregação de informações de múltiplas entradas, como em sistemas de recomendação e classificação de dados [6], [8], [7]. A principal característica desse modelo é a possibilidade de incorporar interações não lineares entre diferentes fontes de informação, além de ser um modelo não linear com características inerentes de interpretabilidade.

Neste trabalho, consideramos um modelo de mistura não linear baseado na integral de Choquet e estamos particularmente interessados no problema de BMI, ou seja, na identificação cega dos parâmetros das integrais de Choquet consideradas no processo de mistura. Nossa abordagem, que pode ser entendida como um primeiro passo para a BSS, se fundamenta na geometria das observações.

O artigo está organizado da seguinte forma. Na Seção II, apresentamos um preâmbulo sobre a integral de Choquet. Na

Gláucia Jardim Pissinelli, Faculdade de Tecnologia (FT), Universidade Estadual de Campinas, Limeira-SP, e-mail: g049933@dac.unicamp.br; Leonardo Tomazeli Duarte, Faculdade de Ciências Aplicadas (FCA), Universidade Estadual de Campinas, Limeira-SP, e-mail: leonardo.duarte@fca.unicamp.br; Yannick Deville, Institut de Recherche en Astrophysique et Planetologie (IRAP), Université Paul Sabatier Toulouse 3, Toulouse-FR, e-mail: Yannick.Deville@irap.omp.eu.

Seção III, apresentamos a proposta de mistura baseada na integral de Choquet e a análise do processo de estimação dos parâmetros do modelo de mistura. Na Seção IV, analisamos alguns aspectos da ideia proposta por meio de experimentos numéricos. Finalmente, na Seção V, apresentamos nossas conclusões.

## II. INTEGRAL DE CHOQUET

A integral de Choquet [6] pode ser vista como um operador de agregação não linear e de natureza paramétrica, de modo que os parâmetros desse modelo são representados pelo conceito de capacidade ou medida fuzzy. Seja  $\mathcal{C}$  um conjunto universal finito,  $\mathcal{C} = \{c_1, \dots, c_N\}$ , e  $\mathcal{P}(\mathcal{C})$  o conjunto das partes de  $\mathcal{C}$ , uma medida fuzzy ou capacidade em  $\mathcal{C}$  é uma função  $\mu : \mathcal{P}(\mathcal{C}) \rightarrow [0, 1]$  que satisfaz  $\mu(\emptyset) = 0$ ,  $\mu(\mathcal{C}) = 1$ ,  $\mu(\mathcal{G}) \leq \mu(\mathcal{H})$ , para todo  $\mathcal{G} \subset \mathcal{H} \subset \mathcal{C}$ . Em termos práticos, a capacidade  $\mu(\mathcal{H})$  quantifica a importância da coalizão de cada elemento expressa pelo conjunto  $\mathcal{H}$  [6], [7].

No contexto de agregação de dados, busca-se o mapeamento de um vetor para um escalar. Por exemplo, na área de MCDA (*Multiple Criteria Decision Analysis*), os elementos do vetor representam a avaliação de uma dada alternativa para diferentes critérios [12]. Logo, o objetivo da agregação é extrair um valor único para uma dada alternativa, levando em consideração as avaliações desses diferentes critérios.

Em termos matemáticos, suponha que, considerando a nomenclatura da área de MCDA, o conjunto  $\mathcal{C} = \{c_1, \dots, c_N\}$  represente os critérios a serem agregados, de modo que o vetor  $\mathbf{p} = [p_1, \dots, p_N]$  representa as avaliações de uma dada alternativa em relação aos critérios de  $\mathcal{C}$ . A integral de Choquet de  $\mathbf{p}$  em relação a  $\mu$  é definida por

$$D_{\mathbf{p}}^{\mu} = \sum_{t=1}^N (p_{(t)} - p_{(t-1)}) \mu(T_{(t)}), \quad (1)$$

onde  $p_{(i)}$  representa o  $i$ -ésimo menor elemento de  $\mathbf{p}$ , de modo que  $p_{(1)} \leq \dots \leq p_{(n)}$  e  $p_{(0)} := 0$ , e  $T_{(t)} = \{c_{(t)}, \dots, c_{(N)}\}$ . A integral Choquet respeita a monotonicidade, uma propriedade fundamental no contexto de agregação de dados [6].

Um dos aspectos interessantes da integral de Choquet é que este modelo, ainda que de natureza não linear, admite interpretação de seus parâmetros [8]. De fato, uma ferramenta útil para interpretar a integral de Choquet é fornecida pelo índice de interação generalizada. Este índice é definido para qualquer agrupamento  $\mathcal{H}$  obtido a partir do conjunto dos elementos de  $\mathcal{C}$  e é dado por

$$I(\mathcal{H}) := \left( \sum_{\mathcal{K} \subset \mathcal{C} \setminus \mathcal{H}} \frac{(N - |\mathcal{K}|) - |\mathcal{H}|! |\mathcal{K}|!}{(N - |\mathcal{H}| + 1)!} \right) \times \left( \sum_{\mathcal{B} \subset \mathcal{H}} (-1)^{(|\mathcal{H}| - |\mathcal{B}|)} \mu(\mathcal{K} \cup \mathcal{B}) \right). \quad (2)$$

Os conjuntos  $\mathcal{K}$  e  $\mathcal{B}$  denotam conjuntos auxiliares que atuam como índices para os operadores de soma. Quando calculado para um único elemento de  $\mathcal{C}$ , digamos  $c_i$ , o índice de interação generalizado corresponde ao índice de importância de Shapley, que é denotado por  $I_i = I\{c_i\}$ . O índice de Shapley  $I_i$

quantifica a contribuição média do elemento  $c_i$  em todas as coalizões. De fato, a importância de  $c_i$  deve levar em conta não apenas  $\mu(c_i)$ , mas também todos os valores de  $\mu(\mathcal{H})$ , tais que  $c_i \in \mathcal{H}$ . Os índices de Shapley são sempre positivos e sua soma é igual a 1, ou seja,  $\sum_{i=1}^N I_i = 1$  [7].

A interação entre dois elementos  $c_i$  e  $c_j$  pode ser quantificada calculando (2) para a coalizão  $\mathcal{H} = \{c_i, c_j\}$ . A medida resultante, representada por  $I_{ij} = I(c_i, c_j)$  e denominada índice de interação, permite interpretar o tipo de interação que ocorre entre os elementos. Quando  $I_{ij} > 0$ , essa relação é dita positiva ou complementar, ou seja, a importância conjunta, para fins do processo de agregação, dos elementos  $i$  e  $j$  é maior do que a soma das importâncias individuais. Por outro lado, se  $I_{ij} < 0$ , caracteriza-se uma interação negativa ou substitutiva, pois a importância conjunta dos elementos  $i$  e  $j$  é menor do que a soma das importâncias individuais. Por fim,  $I_{ij} = 0$  significa que os elementos não interagem, portanto sua contribuição para a agregação final é independente. O modelo da integral de Choquet pode ser representado em termos de  $\mu(\mathcal{H})$  ou  $I(\mathcal{H})$  para todas as coalizões de elementos, pois  $I(\mathcal{H})$  é uma transformada inversível de  $\mu(\mathcal{H})$  [6].

Embora seja bastante abrangente, o número de possíveis coalizões na integral de Choquet é de  $2^N - 2$ , o que torna sua utilização bastante difícil. Na prática, é comum utilizar uma restrição no número de coalizões a serem consideradas por meio de classes restritas de capacidades conhecidas como  $k$ -aditivas, que fixa o grau de interação entre os elementos [6], [7].

Neste trabalho, utilizaremos a classe de capacidades 2-aditivas. Assim, as capacidades são completamente determinadas pelos índices de Shapley  $I_i$  e pelos índices de interação  $I_{ij}$  e ainda assim são flexíveis o suficiente para modelar interações não lineares, como as positivas e negativas.

Considerando capacidade 2-aditiva, a integral de Choquet pode ser expressa por [7]

$$D_{\mathbf{p}}^{\mu} = \sum_{I_{ij} > 0} (p_i \wedge p_j) I_{ij} + \sum_{I_{ij} < 0} (p_i \vee p_j) |I_{ij}| + \sum_{i=1}^N p_i \left( I_i - \frac{1}{2} \sum_{j \neq i} |I_{ij}| \right), \quad (3)$$

onde  $\wedge$  e  $\vee$  são os operadores de mínimo e máximo, respectivamente. A representação expressa (3) evidencia o papel dos índices de interação  $I_{ij}$  como termos não lineares do modelo de agregação, enquanto os índices de Shapley estão associados à parte linear do modelo.

## III. PROPOSTA DE MÉTODO DE BSS PARA UM MODELO BASEADO NA INTEGRAL DE CHOQUET

### A. Formulação Geral do Problema de BSS

No caso mais geral de misturas não lineares, as observações podem ser expressas por [5]:

$$\mathbf{x}(t) = \mathbf{F}(\mathbf{s}(t)), \quad (4)$$

onde  $\mathbf{F} : \mathbb{R}^N \rightarrow \mathbb{R}^M$  é a função de mistura desconhecida,  $N$  e  $M$  são o número de fontes e de sinais de observação respectivamente e  $\mathbf{s}(t) = [s_1(t), s_2(t), \dots, s_N(t)]^T$  e  $\mathbf{x}(t) =$

$[x_1(t), x_2(t), \dots, x_M(t)]^T$  ( $T$  denota a transposição do vetor) representam as fontes e as misturas, respectivamente.

### B. Modelo de mistura baseado na integral de Choquet

No caso da integral de Choquet para o caso 2-aditivo, é possível reescrever os operadores  $\wedge$  e  $\vee$  em termos de módulos  $|\cdot|$  [6]. Fazendo tal substituição, e considerando que os elementos a serem agregados são as fontes e que o resultado da agregação é uma mistura, podemos estabelecer um modelo de mistura no qual a  $k$ -ésima mistura é dada por:

$$x_k(t) = \sum_{i=1}^N I_i^{(k)} s_i(t) - \sum_{1 \leq i < j \leq N} \frac{I_{i,j}^{(k)}}{2} |s_i(t) - s_j(t)|, \quad (5)$$

onde, para uma dada mistura  $k$ ,  $I_i^{(k)}$  representa o índice de Shapley, e  $I_{ij}^{(k)}$  representa o índice de interação entre as fontes  $i$  e  $j$ .

Restringimos inicialmente nosso estudo ao caso em que  $N = M = 2$ , de modo que as misturas podem ser descritas como

$$\begin{aligned} x_1(t) &= I_1^{(1)} s_1(t) + I_2^{(1)} s_2(t) - \frac{I_{1,2}^{(1)}}{2} |s_1(t) - s_2(t)| \\ x_2(t) &= I_1^{(2)} s_1(t) + I_2^{(2)} s_2(t) - \frac{I_{1,2}^{(2)}}{2} |s_1(t) - s_2(t)|. \end{aligned} \quad (6)$$

Nossa proposta se apoia em uma fundamentação geométrica, motivada pelo que o modelo expresso em (6) tem uma estrutura linear por partes.

### C. Proposta de uma Abordagem Geométrica

A partir dos sinais misturas, representados por  $x_1(t)$  e  $x_2(t)$ , analisamos inicialmente a geometria do problema no gráfico de dispersão das misturas. Neste sentido, uma observação em cada instante  $t$  pode ser representada como um ponto  $(x_1, x_2)$  na Figura 1. O processo de mistura, dado pela Equação (6), define uma transformação, apresentada na Figura 1, de forma que um ponto  $(s_1(t), s_2(t))$  é levado a  $(x_1(t), x_2(t))$ . Podemos utilizar as equações da mistura para definir as equações paramétricas das retas ( $x_2 = \alpha_m x_1 + \beta_m$ ), para  $m = 1, 2, 3$  e 4, que definem o polígono que envolve todas as amostras no gráfico de dispersão das misturas.

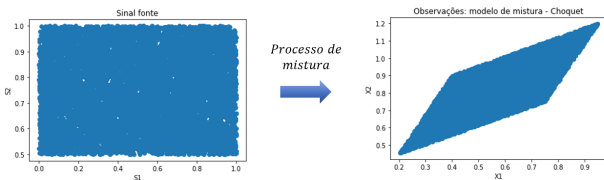


Fig. 1. Geometria das distribuições conjuntas de fontes (esquerda) e misturas (direita).

De modo a expressar o modelo de mistura (6) considerando a notação usual da área de BSS, podemos reescrevê-lo da seguinte maneira:

$$\begin{aligned} x_1 &= a_{11}s_1 + a_{12}s_2 - a_{13}|s_1 - s_2| \\ x_2 &= a_{21}s_1 + a_{22}s_2 - a_{23}|s_1 - s_2|, \end{aligned} \quad (7)$$

onde  $a_{ij}$  são os coeficientes do modelo de mistura e a dependência temporal ( $t$ ) é omitida, uma vez que o sistema de mistura não tem memória e a dependência temporal das amostras não será utilizada. Cabe destacar que trabalharemos com as fontes no suporte  $[0,1]$ .

O próximo passo é estimar os coeficientes angulares ( $\alpha_1$  e  $\alpha_2$ ) e os lineares ( $\beta_1$  e  $\beta_2$ ), conforme apresentado na Figura 2; as retas nesta figura são denotadas por  $A'$ ,  $B'$ ,  $C'$  e  $D'$ . No presente trabalho, assumiremos uma estimação ideal de tais coeficientes. Na prática, é possível estimar tais coeficientes com uma estratégia similar àquela apresentada em [16] ou por meio de transformadas da área de processamento de imagens, como a transformada de Hough.

No presente trabalho, consideraremos que  $\alpha_1 \geq 1 \geq \alpha_3$ , assim como, nos restringiremos a situações nas quais o gráfico de dispersão da Figura 2 não apresenta sobreposição. Cabe lembrar que em BSS é natural operar com uma ambiguidade de permutação na recuperação das fontes [5].

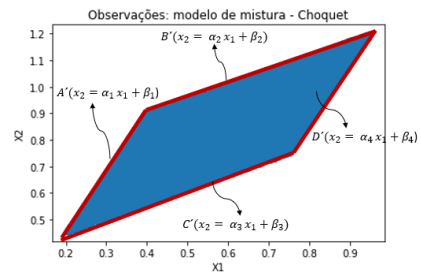


Fig. 2. Geometria das distribuições das misturas.

Logo, se considerarmos  $s_2 \geq s_1$  nas Equações (7), é possível expressar os coeficientes angulares ( $\alpha_1$  e  $\alpha_2$ ) e os lineares ( $\beta_1$  e  $\beta_2$ ) de  $A'$  e  $B'$ , respectivamente, da seguinte forma:

$$\begin{aligned} x_1 &= a_{11}s_1 + a_{12}s_2 - a_{13}(-s_1 + s_2) \\ x_2 &= a_{21}s_1 + a_{22}s_2 - a_{23}(-s_1 + s_2). \end{aligned} \quad (8)$$

Assim, quando  $s_1 = 0$ , o que caracteriza o lugar geométrico da reta  $A'$ , temos que:

$$\alpha_1 = \frac{a_{22} - a_{23}}{a_{12} - a_{13}} \quad (9)$$

$$\beta_1 = 0. \quad (10)$$

Da mesma forma, quando  $s_2 = 1$ , temos que:

$$\alpha_2 = \frac{a_{21} + a_{23}}{a_{11} + a_{13}} \quad (11)$$

$$\beta_2 = \frac{a_{11}a_{22} - a_{11}a_{23} - a_{12}a_{21} - a_{12}a_{23} + a_{13}a_{21} + a_{13}a_{22}}{a_{11} + a_{13}}. \quad (12)$$

Analogamente, se considerarmos  $s_2 \leq s_1$  e reescrevermos as Equações (7), podemos expressar os coeficientes angulares ( $\alpha_3$  e  $\alpha_4$ ) e os lineares ( $\beta_3$  e  $\beta_4$ ) de  $C'$  e  $D'$  (ver Figura 2).

Assim, quando  $s_2 = 0$ , temos que:

$$\alpha_3 = \frac{a_{21} - a_{23}}{a_{11} - a_{13}} \quad (13)$$

$$\beta_3 = 0. \quad (14)$$

Se  $s_1 = 1$ , temos

$$\alpha_4 = \frac{a_{22} + a_{23}}{a_{12} + a_{13}} \quad (15)$$

$$\beta_4 = \frac{(-a_{11}a_{22} - a_{11}a_{23} + a_{12}a_{21} - a_{12}a_{23} + a_{13}a_{21} + a_{13}a_{22})}{(a_{12} + a_{13})}. \quad (16)$$

O cálculo de  $\alpha_1, \alpha_2, \alpha_3, \alpha_4, \beta_2$  e  $\beta_4$ , nos permite estabelecer um sistema não linear de seis equações com relação a  $a_{11}, a_{12}, a_{13}, a_{21}, a_{22}$  e  $a_{23}$ , representado pelas Equações (17), (18), (19), (20), (21) e (22). Vale ressaltar que  $\beta_1 = 0$  e  $\beta_3 = 0$  e, por este motivo não aparecem no sistema de equações.

Dado que visamos estimar os coeficientes do modelo de mistura, o próximo passo consiste em reescrevê-los em função dos coeficientes angulares e lineares  $\alpha_m$  e  $\beta_m$ . Para tanto, é necessário resolver o sistema de equações a seguir:

$$\alpha_1 a_{12} - \alpha_1 a_{13} - a_{22} + a_{23} = 0 \quad (17)$$

$$\alpha_2 a_{11} + \alpha_2 a_{13} - a_{21} - a_{23} = 0 \quad (18)$$

$$\alpha_3 a_{11} - \alpha_3 a_{13} - a_{21} + a_{23} = 0 \quad (19)$$

$$\alpha_4 a_{12} + \alpha_4 a_{13} - a_{22} - a_{23} = 0 \quad (20)$$

$$\beta_2 a_{11} + \beta_2 a_{13} - a_{11} a_{22} + a_{11} a_{23} + a_{12} a_{21} + a_{12} a_{23} - a_{13} a_{21} - a_{13} a_{22} = 0 \quad (21)$$

$$\beta_4 a_{12} + \beta_4 a_{13} + a_{11} a_{22} + a_{11} a_{23} - a_{12} a_{21} + a_{12} a_{23} - a_{13} a_{21} - a_{13} a_{22} = 0. \quad (22)$$

Esse processo se inicia resolvendo as Equações (17), (18), (19) e (20) em termos de  $a_{11}$  e  $a_{21}$ , assim obtemos:

$$a_{12} = \frac{2a_{21}(\alpha_1 - \alpha_2 - \alpha_3 + \alpha_4)}{(\alpha_2 - \alpha_3)(\alpha_1 - \alpha_4)} - \frac{a_{11}(-\alpha_1\alpha_2 - \alpha_1\alpha_3 + 4\alpha_2\alpha_3 - \alpha_2\alpha_4 - \alpha_3\alpha_4)}{(-\alpha_2 + \alpha_3)(\alpha_1 - \alpha_4)} \quad (23)$$

$$a_{22} = -\frac{a_{21}(\alpha_1\alpha_2 + \alpha_1\alpha_3 - 4\alpha_1\alpha_4 + \alpha_2\alpha_4 + \alpha_3\alpha_4)}{(\alpha_2 - \alpha_3)(\alpha_1 - \alpha_4)} + \frac{2a_{11}(-\alpha_1\alpha_2\alpha_3 + \alpha_1\alpha_2\alpha_4 + \alpha_1\alpha_3\alpha_4 - \alpha_2\alpha_3\alpha_4)}{(-\alpha_2 + \alpha_3)(\alpha_1 - \alpha_4)} \quad (24)$$

$$a_{13} = \frac{2a_{21}}{\alpha_2 - \alpha_3} - \frac{a_{11}(\alpha_2 + \alpha_3)}{\alpha_2 - \alpha_3} \quad (25)$$

$$a_{23} = -\frac{a_{21}(-\alpha_2 - \alpha_3)}{\alpha_2 - \alpha_3} + \frac{2a_{11}\alpha_2\alpha_3}{-\alpha_2 + \alpha_3}. \quad (26)$$

Substituindo os termos de  $\alpha_m$  que multiplicam  $a_{11}$  e  $a_{21}$ , nas Equações (23), (24), (25) e (26), por  $f_1, f_2, g_1, g_2, h_1, h_2, l_1, l_2$ , respectivamente, obtemos:  $a_{11} = a_{11}$ ,  $a_{12} = f_1 a_{11} + f_2 a_{21}$ ,  $a_{21} = a_{21}$ ,  $a_{22} = g_1 a_{11} + g_2 a_{21}$ ,  $a_{13} = h_1 a_{11} + h_2 a_{21}$  e  $a_{23} = l_1 a_{11} + l_2 a_{21}$ , e aplicando na Equação (21), temos que:  $\beta_2 + \beta_2 h_1 = A_1$ ;  $\beta_2 h_2 = B_1$ ;  $-g_1 + l_1 + f_1 l_1 - g_1 h_1 = C_1$ ;  $-g_2 + l_2 + f_1 + f_1 l_2 + f_2 l_1 - h_1 - g_2 h_1 - g_1 h_2 = D_1$  e  $f_2 + f_2 l_2 - h_2 - g_2 h_2 = E_1$ . Logo, é possível reescrevê-las em termos de uma equação quadrática completa:

$$A_1 a_{11} + B_1 a_{21} + C_1 a_{11}^2 + D_1 a_{11} a_{21} + E_1 a_{21}^2 = 0. \quad (27)$$

A mesma operação poderá ser repetida para Equação (22), e obtemos

$$A_2 a_{11} + B_2 a_{21} + C_2 a_{11}^2 + D_2 a_{11} a_{21} + E_2 a_{21}^2 = 0. \quad (28)$$

Se considerarmos,  $a_{21} = a_{11}k$ , obtemos dois resultados possíveis para as Equações (27) e (28). Se  $a_{11} = 0$ , ambas as equações se tornam 0. Portanto,  $a_{11} = 0$  e  $a_{21} = 0$  é uma solução trivial para o problema. Se  $a_{11} \neq 0$ , temos:

$$\frac{A_1 + B_1 k}{C_1 + D_1 k + E_1 k^2} = \frac{A_2 + B_2 k}{C_2 + D_2 k + E_2 k^2}. \quad (29)$$

Portanto, podemos expressar essa relação na forma de uma equação cúbica:

$$T k^3 + S k^2 + Z k + N = 0, \quad (30)$$

onde  $T = (B_1 E_2 - B_2 E_1)$ ,  $S = (B_1 D_2 + A_1 E_2 - B_2 D_1 - A_2 E_1)$ ,  $Z = (B_1 C_2 + A_1 D_2 - B_2 C_1 - A_2 D_1)$  e  $N = (A_1 C_2 - A_2 C_1)$ . Para obter um coeficiente unitário para o termo cúbico, dividimos a Equação (30) por  $T$ , e considerando  $a = \frac{S}{T}$ ,  $b = \frac{Z}{T}$  e  $c = \frac{N}{T}$  podemos reescrever a Equação (30) da seguinte maneira:

$$k^3 + a k^2 + b k + c = 0. \quad (31)$$

A fórmula cúbica de Cardano é uma solução geral para equações cúbicas. Podemos transformar a equação cúbica em uma fórmula reduzida, eliminando o termo quadrático, através da substituição de  $k = y - a/3$  [13]. Dessa forma, temos:

$$y^3 + p y + q = 0, \quad (32)$$

sendo que,  $p = \frac{3b - a^2}{3}$  e  $q = c + \frac{2a^3}{27} - \frac{ab}{3}$ . Assim, é possível calcular o discriminante da seguinte forma:  $D = \left(\frac{p}{3}\right)^3 + \left(\frac{q}{2}\right)^2$ .

A fórmula de Cardano requer o cálculo do discriminante da equação cúbica. A partir desse valor, é possível calcular as três raízes cúbicas,  $\gamma_1, \gamma_2$  e  $\gamma_3$ , e determinar o tipo da raiz cúbica [13]:

- 1)  $D > 0$ , então a equação cúbica tem uma raiz real  $\gamma_1$  e duas raízes conjugadas complexas  $\gamma_2$  e  $\gamma_3$ ;
- 2)  $D = 0$ , então a equação tem três soluções reais, incluindo uma solução dupla;
- 3)  $D < 0$ , então a equação tem três raízes reais distintas  $\gamma_1, \gamma_2$  e  $\gamma_3$ .

Todos os passos apresentados visam, portanto, estimar os coeficientes do modelo de mistura a partir dos aspectos geométricos dos dados, representados pelos coeficientes angulares e lineares  $\alpha_m$  e  $\beta_m$ . Ou seja, no uso efetivo da metodologia proposta tem-se uma etapa que estima os coeficientes de mistura, tal estimativa pode ser denotada  $\hat{a}_{ij}$ .

#### IV. EXPERIMENTOS NUMÉRICOS

Nesta seção, apresentamos alguns resultados numéricos para estimação dos coeficientes do modelo de mistura a partir da geometria das observações (demonstrada na Seção III-C). Cabe ressaltar que para estimar os coeficientes da mistura, seguindo a sequência definida na seção anterior, pressupomos que inicialmente os valores dos coeficientes lineares  $\alpha_m$  e angulares  $\beta_m$  são conhecidos.

No processo de estimação dos coeficientes de mistura a partir de  $\alpha_m$  e  $\beta_m$ , um ponto fundamental é compreender se

o sistema de equações proposto resulta em soluções única ou múltiplas. Para conduzir essa análise, podemos inspecionar o discriminante de Cardano. A depender do valor desse discriminante, é possível determinar o tipo de raízes sem a necessidade de calcular explicitamente cada raiz individualmente. O experimento realizado, consiste na recuperação dos valores de  $a_{ij}$  a partir de  $\alpha_m$  e  $\beta_m$  considerando o equacionamento da seção III-C.

Foram realizadas 5000 iterações com as seguintes características: os valores dos coeficientes de mistura  $a_{ij}$  são gerados aleatoriamente (distribuição uniforme entre  $[0, 1]$ ) e as condições de monotonicidade são respeitadas.

Os resultados mostram que é possível estimar os coeficientes  $a_{ij}$  em 97,74% das iterações. Nas demais execuções, os coeficientes de mistura não foram estimados corretamente por completo; é importante ressaltar que apenas os valores de  $a_{11}$  e  $a_{22}$ , não foram recuperados por completo.

Em relação à análise do discriminante, é possível observar que em 96,84% das iterações, o valor do discriminante é negativo ( $D < 0$ ), indicando que a equação cúbica de Cardano (30) possui três raízes reais distintas. Neste caso, os coeficientes de mistura foram estimados considerando a melhor entre as três soluções  $\gamma_1$ ,  $\gamma_2$  e  $\gamma_3$ , que correspondem às raízes obtidas com a equação de Cardano. O discriminante é igual a zero ( $D = 0$ ) em 3,16% das iterações, indicando que há três soluções reais, incluindo uma solução dupla. E em nenhuma iteração o discriminante foi positivo ( $D > 0$ ). Note que tal procedimento não é cego, pois exige o conhecimento dos verdadeiros valores dos coeficientes de mistura. No entanto, procedemos desta maneira apenas para verificar se, dentre as três soluções obtidas por Cardano, uma delas se aproxima dos valores verdadeiros dos coeficientes de mistura.

## V. CONCLUSÕES

Neste trabalho, propomos uma abordagem geométrica para separação cega de fontes em um modelo baseado na integral de Choquet. A abordagem geométrica utilizada para estimar a geometria das observações do problema mostrou-se eficaz na maioria dos casos analisados, permitindo uma estimativa satisfatória dos coeficientes de mistura. Além disso, a análise do discriminante exigirá mais investigações, sendo necessário compreender como selecionar as raízes de modo não supervisionado.

Além disso, exploraremos situações onde há recobrimento no espaço de misturas, ou seja, casos nos quais não há mais relação de bijetividade entre as fontes e as misturas. Finalmente, abordaremos o problema de estimação dos coeficientes angulares  $\alpha_m$  e lineares  $\beta_m$ .

## AGRADECIMENTOS

Agradecemos o CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) e ao programa mobilidade internacional de estudantes de pós graduação - Santander/DERI-Unicamp pelo apoio financeiro para a realização desta pesquisa.

## REFERÊNCIAS

- [1] J. Héroult, B. Ans, "Circuits neuronaux à synapses modifiables: décodage de messages composites par apprentissage non supervisé". C.R. Acad. Sci., 299 (1984), pp. 525-528
- [2] J. Héroult, C. Jutten, B. Ans, "Détection de grandeurs primitives dans un message composite par une architecture de calcul neuromimétique en apprentissage non supervisé", in: Actes du Xème colloque GRETSI, Nice, France, May 1985, pp. 1017–1022.
- [3] B. Ans, J. Héroult, C. Jutten, "Adaptive neural architectures: Detection of primitives", in: Proceedings of COGNITIVA'85, Paris, France, June 1985, pp. 593–597.
- [4] Pierre Comon, Independent component analysis, "A new concept?", Signal Processing", vol. 36, Issue 3, 1994, pp. 287-314.
- [5] P. Comon and C. Jutten, editors. "Handbook of blind source separation: independent component analysis and applications". Academic Press, 2010.
- [6] M. Grabisch, "A graphical interpretation of the Choquet integral", in IEEE Transactions on Fuzzy Systems, vol. 8, no. 5, pp. 627-631, Oct. 2000.
- [7] L. T. Duarte, "A novel multicriteria decision aiding method based on unsupervised aggregation via the Choquet integral", in IEEE Transactions on Engineering Management, vol. 65, no. 2, pp. 293-302, May 2018.
- [8] M. Grabisch, "K-order additive discrete fuzzy measures and their representation," Fuzzy Sets Syst., vol. 92, no. 2, pp. 167–189, 1997.
- [9] Y. Deville and L. T. Duarte, "An overview of blind source separation methods for linear-quadratic and post-nonlinear mixtures", in Latent Variable Analysis and Signal Separation. Springer, 2015, pp. 155–167.
- [10] C.G. Puntonet, A. Prieto, C. Jutten, M. Rodríguez-Alvarez, J. Ortega, "Separation of sources: A geometry-based procedure for reconstruction of n-valued signals", Signal Processing, vol. 46, Issue 3, 1995, pp. 267-284.
- [11] A. Taleb and C. Jutten, "Source separation in post-nonlinear mixtures", IEEE Transactions on Signal Processing, 1999, vol. 47, pp. 2807-2820.
- [12] S. Greco, "Multiple criteria decision analysis: state of the art surveys". International Series in Operations Research & Management Science, 2006, Springer New York.
- [13] J. W. Harris, and H. Stocker, "Handbook of Mathematics and Computational Science", 1998, Springer New York.
- [14] Y. Deville, L. T. Duarte, S. Hosseini, "Nonlinear blind source separation and blind mixture identification. Methods for bilinear, linear-quadratic and polynomial mixtures", SpringerBriefs in Electrical and Computer Engineering, Springer Nature, 2021.
- [15] D. V. Carvalho, E. M. Pereira, J. S. Cardoso, "Machine learning interpretability: A survey on methods and metrics". Electronics, 2019, 8(8), 832.
- [16] A. Mansour, N. Ohnishi, C.G. Puntonet, "Blind multiuser separation of instantaneous mixture algorithm based on geometrical concepts", Signal Processing, vol. 82, issue 8, 2002.