

O teorema de Darmois-Skitovich e a separação cega de fontes

Flávio R. M. Pavan e Maria D. Miranda

Escola Politécnica, Universidade de São Paulo, São Paulo, SP, Brasil

E-mails: frmp@lcs.poli.usp.br, maria@lcs.poli.usp.br

Abstract - *The Darmois-Skitovich theorem (1953) had a fundamental role in the development of independent component analysis (ICA) techniques, usually employed in the context of blind source separation (BSS). Although implicitly used in BSS, little emphasis is given in general literature to this theorem's derivation, which is not evident, and to its interpretation. The goal of this paper is to revisit and interpret, in a more intuitive manner, the Darmois-Skitovich theorem in this context.*

Palavras-chave: *separação cega de fontes, análise de componentes independentes, variáveis aleatórias independentes.*

1 Introdução

Em diversos contextos práticos, como, por exemplo, em Processamento de Sinais Biomédicos [1], Processamento de Áudio [2], Comunicações [3], dentre outros, surge a necessidade de recuperar individualmente sinais desconhecidos, combinados de forma também desconhecida [4, 5]. Durante muitos anos, a solução desse problema, atualmente consagrado como separação cega de fontes (BSS, do inglês, *blind source separation*), foi dada como impossível [6].

Um esquema simplificado do problema de BSS é mostrado na Figura 1. Nesse esquema, N fontes indicadas pelo vetor $\mathbf{s}(n)$ são misturadas, resultando em L misturas indicadas pelo vetor $\mathbf{x}(n)$. O objetivo de BSS é obter o sistema separador \mathbf{W} que resulte em boas estimativas $\mathbf{y}(n)$ de $K \leq N$ fontes.

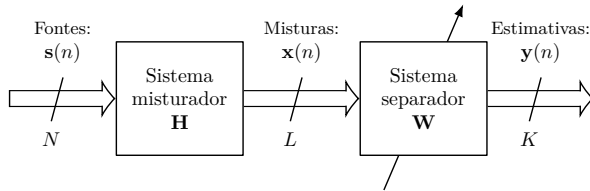


Figura 1: Esquema simplificado do problema de BSS: separação de K fontes a partir da observação de L misturas.

Estudos na área de BSS começaram no início dos anos 1980, considerando apenas sistemas misturadores lineares e instantâneos [6]. Em 1985, B. Ans, J. Héroult e C. Jutten mostraram a possibilidade de resolver o problema de BSS através do uso de estruturas não lineares no sistema separador [7]. Durante essa época, pouco ainda se sabia sobre as limitações teóricas das técnicas de separação existentes e por que de fato elas funcionavam [6]. Apenas no início dos anos 90, P. Comon introduziu um conjunto de técnicas conhecidas coletivamente como análise de componentes independentes (ICA, do inglês, *independent component analysis*) [8, 9]. Essas técnicas exploram, de uma forma ou de outra, dois conceitos importantes: *independência estatística* e *não gaussianidade*.

Em [9], estabeleceram-se condições de *separabilidade das fontes* no contexto de ICA. Tratam-se de condições teóricas sobre o sistema misturador e sobre as distribuições das fontes para garantir uma separação adequada. Além disso, mostrou-se que, mesmo que o sistema separador garanta componentes independentes de $\mathbf{y}(n)$, isso não implica

a separação das fontes. É preciso adicionar a restrição de não gaussianidade das fontes para que a recuperação de independência garanta a separação. Essa relação muito pouco evidente entre independência estatística e não gaussianidade foi realçada por Comon com base no *teorema de Darmois-Skitovich*, publicado em 1953 nos trabalhos independentes de G. Darmois [10] e V. Skitovich [11], na área de Estatística conhecida como Análise Fatorial.

A questão da separabilidade foi fundamental para desenvolver e compreender os limites de funcionamento da ICA. Cabe notar, além disso, que [9] evidenciou a possibilidade de tirar proveito das estatísticas dos sinais para separar fontes e provou-se um marco na área de BSS, inspirando diversos outros trabalhos; veja, por exemplo, [4]–[6], [12] e [13].

O teorema de Darmois-Skitovich teve, portanto, um papel importante no desenvolvimento das técnicas de ICA; daí a sua relevância no contexto de separação cega de fontes. Entretanto, apesar de ser tacitamente empregado, pouco enfoque é dado a sua dedução, que não é evidente, e também a sua interpretação no contexto de BSS. Tendo isso em vista, o objetivo deste artigo é revisar e interpretar, com o auxílio de alguns exemplos, o teorema de Darmois-Skitovich.

O artigo está estruturado como segue. Na Seção 2, os teoremas preliminares necessários na demonstração simplificada do teorema de Darmois-Skitovich são introduzidos e comentados com exemplos. Na Seção 3, o teorema de Darmois-Skitovich é demonstrado inspirando-se em [14] e [15]. Além disso, o teorema decorrente provado por Comon, que garante separabilidade de fontes no contexto de ICA, é apresentado nessa seção. Na Seção 4, as condições de separabilidade são interpretadas em um exemplo de BSS para misturas instantâneas e lineares. Na Seção 5, são apresentadas as conclusões do artigo.

2 Teoremas preliminares

Teorema 1 (Cramér, 1936). *Sejam $\{X_i\}$, com $i = 1, 2, \dots, N$, variáveis aleatórias independentes. Se*

$$Y = \sum_{i=1}^N a_i X_i \quad (1)$$

é gaussiana, então todas as variáveis aleatórias X_i para as quais $a_i \neq 0$ também são gaussianas.

Esse teorema foi inicialmente conjecturado por P. Lévy e depois provado para $N = 2$ em [16], e estendido para qualquer $N \geq 2$ em [14]. Embora seja um fato conhecido que a combinação linear de variáveis aleatórias independentes gaussianas também é gaussiana [17], a recíproca não é evidente.

Pode-se contrastar o Teorema 1 com o teorema central do limite (TCL), largamente utilizado em Estatística. Em linhas gerais, nota-se que se $\{X_i\}$ com $i = 1, 2, \dots, N$ são variáveis aleatórias independentes e identicamente distribuídas, à medida que N tende ao infinito, a distribuição de Y em (1) tende, em um sentido probabilístico, a uma distribuição gaussiana [17]. Com base nesse fato, seria possível alegar que, como o TCL não restringe as variáveis

aleatórias somadas a serem gaussianas, o Teorema 1 estaria errado. Porém, não é esse o caso: enquanto o TCL considera a soma de uma *infinitude* de variáveis aleatórias não necessariamente gaussianas, o Teorema 1 considera a composição de uma variável aleatória gaussiana a partir da soma de uma quantidade *finita* de variáveis aleatórias independentes.

Teorema 2 (Marcinkiewicz-Dugué, 1951). *Seja X uma variável aleatória cuja função característica é dada por [17]*

$$\Phi_X(\omega) = \mathbb{E} [e^{j\omega X}], \quad (2)$$

em que $\mathbb{E}[\cdot]$ é o operador valor esperado. As únicas variáveis aleatórias que possuem função característica da forma $\Phi_X(\omega) = e^{p(\omega)}$, em que $p(\omega)$ é um polinômio, são as variáveis aleatórias constantes e gaussianas. Em outras palavras, se $e^{p(\omega)}$ é uma função característica de uma variável aleatória qualquer, então o grau de $p(\omega)$ é obrigatoriamente menor ou igual a 2.

Alguns casos particulares desse teorema foram observados na tese de M. G. Kuznetz em 1937. A sua demonstração geral foi feita por M. Marcinkiewicz em 1940. Além disso, D. Dugué indicou o teorema em 1939 e apresentou uma prova mais curta em 1951 [18]. Cabe notar que a dedução do teorema é complicada: utiliza conceitos de análise complexa e funções meromórficas, ou seja, funções analíticas exceto em alguns pontos isolados. A dedução é feita estabelecendo-se condições sobre $p(\omega)$ para que $e^{p(\omega)}$ resulte uma função característica. Nota-se que essas condições apenas são satisfeitas para $p(\omega)$ com grau zero, um ou dois. No exemplo a seguir, verifica-se o que pode ser dito sobre a variável aleatória X em cada um desses casos.

Exemplo 1. Se a função densidade de probabilidade de X existir e valer f_X , então aplicando a definição de valor esperado em (2), sua função característica pode ser escrita como

$$\Phi_X(\omega) = \int_{-\infty}^{+\infty} e^{j\omega x} f_X(x) dx. \quad (3)$$

Nota-se que a função característica de X é a transformada de Fourier (com sinal de frequência trocado) de f_X . Com isso, pode-se verificar os seguintes casos de $\Phi_X(\omega) = e^{p(\omega)}$, com $p(\omega)$ polinomial:

- grau $[p(\omega)] = 0$: como $\Phi_X(0) = 1$, então $p(\omega) = 0$. Portanto, $\Phi_X(\omega) = 1$ e X será uma variável aleatória com função densidade de probabilidade $f_X(x) = \delta(x)$, em que $\delta(\cdot)$ é o delta de Dirac. Assim, X é constante e vale 0.
- grau $[p(\omega)] = 1$: como $f_X(x)$ é real, $\Phi_X(-\omega) = \Phi_X^*(\omega)$, em que $(\cdot)^*$ denota o complexo conjugado. Portanto, o coeficiente de ω em $p(\omega)$ é imaginário puro. Tomando $p(\omega) = j\omega x_0$, tem-se $f_X(x) = \delta(x - x_0)$. Nesse caso, X é constante e vale x_0 .
- grau $[p(\omega)] = 2$: tomando-se X gaussiana com média μ e variância σ^2 , tem-se

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

e, notando-se que a transformada de Fourier da gaussiana também é uma gaussiana, resulta, a partir da aplicação do par transformado [17],

$$\Phi_X(\omega) = e^{j\omega\mu - \omega^2 \frac{\sigma^2}{2}}.$$

Conclui-se, pela biunivocidade do par transformado, que $p(\omega)$ tem grau dois se, e somente se, X é gaussiana.

Lema 1 (Darmois, 1953). *Sejam f_1, f_2, \dots, f_N e g_1, g_2 funções contínuas em um aberto V . Se existir uma decomposição*

$$T(x, y) = \sum_{\ell=1}^N f_\ell(a_\ell x + b_\ell y) = g_1(x) + g_2(y), \forall x, y \in V, \quad (4)$$

em que $a_1, a_2, \dots, a_N, b_1, b_2, \dots, b_N$ são constantes não nulas tais que

$$a_i b_j - a_j b_i \neq 0, \quad \forall i \neq j, \quad (5)$$

então as funções $f_i(x)$ são polinômios de grau menor ou igual a N .

A prova desse lema é feita em [10] para $N = 2$ a partir do uso sucessivo de diferenças finitas. Com esse lema, é possível garantir que as funções para as quais a decomposição da Equação (4) ocorre são necessariamente polinômios. Além disso, os graus desses polinômios devem ser menores que o número N de funções originalmente somadas. A seguir, apresentam-se alguns casos particulares para $N = 2$ e $f_1, f_2 : \mathbb{R} \rightarrow \mathbb{R}$.

Exemplo 2. Sejam $f_1(x) = e^x$ e $f_2(x) = \sqrt{|x|}$. As funções f_1 e f_2 não são polinômios e, conforme o lema, a função $T(x, y)$ não admite uma decomposição como em (4).

Exemplo 3. Sejam $f_1(x) = (x+1)^2$ e $f_2(x) = 2x$. Nota-se que uma decomposição seguindo (4) é possível somente se $b_1 = 0$, e esse caso não é contemplado pelo lema. Trata-se de um exemplo em que a recíproca não vale.

Exemplo 4. Sejam $f_1(x) = x^2$ e $f_2(x) = -x^2$. Nesse caso, tem-se a decomposição de $T(x, y)$ segundo (4) se $a_1 b_1 = a_2 b_2$, válida até mesmo para coeficientes não nulos e para $a_1 b_2 \neq a_2 b_1$. De fato, o lema nos garante que f_1 e f_2 são polinômios de grau no máximo $N = 2$; o que, nesse caso, é verdade.

3 O teorema de Darmois-Skitovich

Teorema 3 (Darmois, 1953; Skitovich, 1953). *Sejam S_1, S_2, \dots, S_N variáveis aleatórias independentes. Suponha que*

$$\begin{aligned} Y_1 &= a_1 S_1 + a_2 S_2 + \dots + a_N S_N \\ Y_2 &= b_1 S_1 + b_2 S_2 + \dots + b_N S_N \end{aligned} \quad (6)$$

e que Y_1 e Y_2 são independentes. Se para um valor qualquer de i , com $1 \leq i \leq N$, vale $a_i b_i \neq 0$, então S_i é gaussiana.

Demonstração. A função característica conjunta de Y_1 e Y_2 é igual a

$$\Phi_{Y_1 Y_2}(\omega_1, \omega_2) = \mathbb{E} [\exp(j(\omega_1 Y_1 + \omega_2 Y_2))]. \quad (7)$$

Notando-se que $Y_1 = \sum_{i=1}^N a_i S_i$ e $Y_2 = \sum_{i=1}^N b_i S_i$ resulta

$$\Phi_{Y_1 Y_2}(\omega_1, \omega_2) = \mathbb{E} \left[\exp \left(j \sum_{i=1}^N (\omega_1 a_i + \omega_2 b_i) S_i \right) \right]. \quad (8)$$

Devido à independência dos S_i , com $1 \leq i \leq N$, a seguinte fatoração é possível:

$$\Phi_{Y_1 Y_2}(\omega_1, \omega_2) = \prod_{i=1}^N \Phi_{S_i}(\omega_1 a_i + \omega_2 b_i). \quad (9)$$

Como Y_1 e Y_2 são independentes, vale

$$\Phi_{Y_1 Y_2}(\omega_1, \omega_2) = \Phi_{Y_1}(\omega_1) \Phi_{Y_2}(\omega_2), \quad (10)$$

e a partir de (9), tem-se

$$\Phi_{Y_1}(\omega_1)\Phi_{Y_2}(\omega_2) = \prod_{i=1}^N \Phi_{S_i}(\omega_1 a_i + \omega_2 b_i). \quad (11)$$

Aplicando-se o logaritmo de ambos os lados dessa equação, e denotando $\Psi(\cdot) = \log(\Phi(\cdot))$, resulta

$$\Psi_{Y_1}(\omega_1) + \Psi_{Y_2}(\omega_2) = \sum_{i=1}^N \Psi_{S_i}(\omega_1 a_i + \omega_2 b_i). \quad (12)$$

Quando $a_i b_i = 0$, pode-se agrupar Ψ_{S_i} a algum dos dois termos do lado esquerdo de (12). Esse procedimento deve ser repetido para todos os casos possíveis. Em seguida, supondo que $a_i b_j - a_j b_i \neq 0$ para todo $i \neq j$, aplica-se o Lema 1 na equação resultante. Com isso, observa-se que para os casos em que $a_i b_i \neq 0$, Ψ_{S_i} deve ser um polinômio. Utilizando-se o Teorema 2 e excluindo-se os casos determinísticos, conclui-se que S_i é gaussiana se $a_i b_i \neq 0$.

Para o caso em que $a_i b_j - a_j b_i = 0$ para algum $i \neq j$, pode ser feita uma combinação prévia de variáveis aleatórias. Por exemplo, pode-se definir uma variável aleatória S_{ij} tal que

$$\begin{aligned} \frac{S_{ij}}{b_i} &= a_i S_i + a_j S_j \\ \frac{S_{ij}}{a_i} &= b_i S_i + b_j S_j \end{aligned} \quad (13)$$

se $a_i b_i \neq 0$. Após serem feitas todas as combinações possíveis de variáveis aleatórias em (6), os coeficientes resultantes $\{\bar{a}_i, \bar{b}_i\}$ deverão satisfazer $\bar{a}_i \bar{b}_j - \bar{a}_j \bar{b}_i \neq 0$ para todo $i \neq j$. Prossegue-se realizando os passos que vão da Equação (7) até (12) e aplica-se o Lema 1 como feito anteriormente. Se a partir do lema obter-se que S_{ij} é gaussiana, então aplicando-se o Teorema 1 pode-se garantir que S_i e S_j também são gaussianas. \square

A fim de se compreender o teorema de Darmois-Skitovich no contexto de separação de fontes, remete-se ao esquema da Figura 1 com $K = 2$. Interpretando-se as variáveis $\{S_i\}$ como fontes e $\{Y_i\}$ como estimativas das fontes, nota-se que os coeficientes a_i e b_i variam conforme se ajusta o sistema separador \mathbf{W} . Escolhendo-se \mathbf{W} de modo a se garantir $\{Y_i\}$ independentes entre si, o Teorema 3 estabelece que as fontes serão separadas se não forem gaussianas. Se para a i -ésima fonte vale $a_i b_i \neq 0$, o que significa que ela não está sendo separada, então ela é gaussiana. Em outras palavras, escolhendo-se um sistema separador \mathbf{W} de modo a se garantir componentes independentes de $\{Y_i\}$, garante-se separação das fontes a menos que uma ou mais sejam gaussianas.

Tomando por base o teorema de Darmois-Skitovich, Comon introduziu a ideia de *separação por extração de componentes independentes*, além de generalizar o Teorema 3 para $K > 2$ e estabelecer as condições de separabilidade das fontes. Esses resultados estão resumidos no teorema a seguir.

Teorema 4 (Comon, 1992). *Seja $\mathbf{y} = \mathbf{W}\mathbf{H}\mathbf{s}$, em que $\mathbf{W}\mathbf{H}$ é uma matriz ortogonal $N \times N$ e \mathbf{s} é um vetor de fontes com elementos (ou componentes) estatisticamente independentes, e no máximo um de seus elementos é gaussiano. Então, \mathbf{y} é um vetor aleatório com componentes mutuamente independentes se, e somente se,*

$$\mathbf{W}\mathbf{H} = \mathbf{\Lambda}\mathbf{P}, \quad (14)$$

em que $\mathbf{\Lambda}$ é uma matriz diagonal e \mathbf{P} é uma matriz de permutação.

A partir desse teorema, foi possível justificar o funcionamento das técnicas de ICA, bem como estabelecer seus limites teóricos de funcionamento: ao recuperar a independência estatística nas saídas do sistema separador, recuperam-se as fontes a menos de ambiguidades de escala e permutações, se no máximo uma fonte for gaussiana [9, 5].

A seguir, são ilustradas as implicações práticas dos Teoremas 3 e 4 em um exemplo simples de branqueamento de mistura instantânea e linear.

4 Interpretação no contexto de BSS

Considere o esquema mostrado na Figura 1, com $N = L = 2$. As misturas são obtidas segundo

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{H}\mathbf{s} = \mathbf{H} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix}, \quad (15)$$

em que

$$\mathbf{H} = [\mathbf{h}_1 \quad \mathbf{h}_2] = \begin{bmatrix} 0,1302 & 0,9683 \\ -0,5191 & -0,3545 \end{bmatrix}. \quad (16)$$

Considera-se que as misturas \mathbf{x} passam por um processo posterior de branqueamento [5, 19], resultando no vetor de misturas branqueadas $\tilde{\mathbf{x}}$.

Na Figura 2, são mostrados os diagramas de dispersão para as fontes s_i , misturas x_i e misturas após branqueamento \tilde{x}_i , com $i = 1, 2$ e 1000 realizações independentes de cada variável aleatória. São considerados três pares de fontes independentes com distribuições diferentes, representados nas colunas (a), (b) e (c). Na coluna (a), considera-se duas fontes com distribuição uniforme; na coluna (b), fontes com distribuições distintas, sendo que uma delas possui distribuição bimodal; e na coluna (c), considera-se que as fontes são conjuntamente gaussianas.

As abscissas e ordenadas dos diagramas de dispersão das fontes, mostradas na Figura 2, são também representadas após as transformações lineares que resultam na mistura e no branqueamento da mistura. Para os diagramas de dispersão das misturas, esse eixos apontam na direção das colunas da matriz de misturas \mathbf{H} , ou seja, \mathbf{h}_1 e \mathbf{h}_2 .

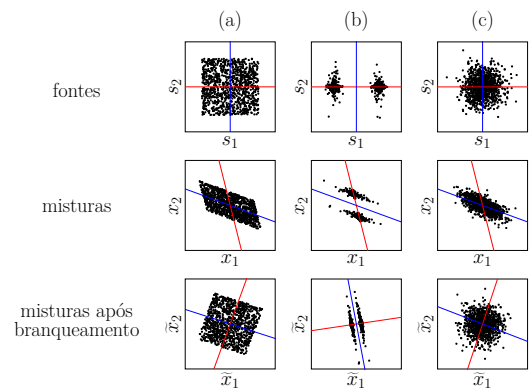


Figura 2: Diagramas de dispersão para as fontes independentes s_i , suas misturas x_i e o efeito do seu branqueamento \tilde{x}_i , com $i = 1, 2$. (a) Fontes com distribuição uniforme. (b) Fontes com distribuições distintas. (c) Fontes conjuntamente gaussianas.

Nota-se, como conhecido, que a técnica de branqueamento é capaz de recuperar a “estrutura” do suporte da distribuição das fontes a menos de uma rotação em todos os casos (a), (b) e (c) da Figura 2. O passo seguinte de separação de fontes, que envolve, por exemplo, técnicas de ICA,

é responsável por descobrir qual é a rotação que, aplicada nas misturas após branqueamento, resultará nas estimativas mais adequadas das fontes independentes.

Percebe-se que há mais de uma rotação possível a ser feita nas misturas após o branqueamento para recuperar a estrutura do diagrama de dispersão das fontes independentes. A depender da rotação escolhida, as fontes poderão ser recuperadas com uma permutação em relação às fontes originais, ou seja, a fonte 2 no lugar da fonte 1, e a fonte 1 no lugar da fonte 2. Trata-se de uma ambiguidade de permutação na solução do problema cego, como previsto pelo Teorema 4 da Seção 3.

Observa-se ainda, no caso da Figura 2-(b), que a mistura após branqueamento também apresenta um alongamento do diagrama de dispersão na direção do eixo azul em relação ao diagrama das fontes independentes. Isso representa a ambiguidade de escala que ocorre na solução do problema cego, como previsto pelo Teorema 4 da Seção 3. Em outras palavras, não é possível recuperar a energia das fontes.

É importante notar que, para fontes conjuntamente gaussianas como na Figura 2-(c), o diagrama de dispersão das fontes é completamente simétrico. Nesse caso, o branqueamento da mistura é suficiente para que as misturas após branqueamento sejam estatisticamente independentes. Poderia se pensar que o problema de BSS, nesse caso, estaria então resolvido, já que dois sinais independentes foram recuperados a partir da decorrelação das misturas. No entanto, qualquer rotação que seja aplicada nas misturas após branqueamento resultará, também, em sinais estatisticamente independentes. Como o problema é cego, fica impossível determinar qual é a rotação adequada que recupera as fontes independentes que se deseja determinar. Portanto, no caso de fontes conjuntamente gaussianas, o passo seguinte de separação cega de fontes não é capaz de descobrir as fontes independentes originais. Essa é uma forma alternativa, e também complementar, de se convencer da condição do Teorema 4 de que no máximo uma das fontes pode ser gaussiana para se ter separação adequada.

5 Conclusões

Desde o início dos anos 1980, várias soluções interessantes vêm sendo propostas para resolver o problema de separação cega de fontes. Muitas dessas soluções estão baseadas em princípios de extração de componentes independentes conhecidos coletivamente como ICA. Os limites teóricos e os princípios de funcionamento da ICA foram esclarecidos por Comon em [9], graças ao teorema de Darmois-Skitovich [10]. Esse teorema estabelece a relação muito pouco evidente entre independência estatística e não gaussianidade. Quando aplicado ao contexto de BSS, esclarece a necessidade de adicionar a restrição de não gaussianidade das fontes para que a recuperação de independência garanta a separação.

Nota-se que, até os dias de hoje, o teorema de Darmois-Skitovich ainda é assunto de pesquisa (veja, por exemplo, [20] e [21]). Particularmente, o objetivo desse artigo foi auxiliar no entendimento de conceitos teóricos fundamentais de BSS, a partir da apresentação e interpretação do teorema de Darmois-Skitovich. Vislumbra-se, em trabalhos futuros, o entendimento aprofundado de técnicas atuais de BSS com níveis de sofisticação elevados.

Agradecimento

Esse trabalho foi parcialmente financiado pela Comissão de Aperfeiçoamento de Pessoal do Nível Superior (CAPES) do

Ministério da Educação.

Referências

- [1] Y. Li, Z. L. Yu, N. Bi, Y. Xu, Z. Gu, and S.-I. Amari, "Sparse Representation for Brain Signal Processing: A tutorial on methods and applications," *IEEE Signal Processing Magazine*, vol. 31, pp. 96–106, maio 2014.
- [2] E. Vincent, N. Bertin, R. Gribonval, and F. Bimbot, "From Blind to Guided Audio Source Separation: How models and side information can improve the separation of sound," *IEEE Signal Processing Magazine*, vol. 31, pp. 107–115, maio 2014.
- [3] P. Chevalier and A. Chevreuil, "Application to telecommunications," in *Handbook of Blind Source Separation: Independent Component Analysis and Applications* (P. Comon and C. Jutten, eds.), ch. 17, pp. 683–725, Oxford: Academic Press, 2010.
- [4] J.-F. Cardoso, "Blind signal separation: statistical principles," *Proceedings of the IEEE*, vol. 86, pp. 2009–2025, out. 1998.
- [5] A. Hyvärinen and E. Oja, "Independent component analysis: algorithms and applications," *Neural Networks*, vol. 13, no. 4–5, pp. 411–430, 2000.
- [6] C. Jutten and P. Comon, "Introduction," in *Handbook of Blind Source Separation: Independent Component Analysis and Applications* (P. Comon and C. Jutten, eds.), ch. 1, pp. 1–28, Oxford: Academic Press, 2010.
- [7] J. Héroult, C. Jutten, and 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), pp. 1017–1022, Groupe d'Etudes du Traitement du Signal et des Images (GRETSI), maio 1985.
- [8] P. Comon, "Independent component analysis," in *Higher-Order Statistics* (J. L. Laucome, ed.), pp. 29–38, Amsterdam: Elsevier, 1992.
- [9] P. Comon, "Independent component analysis, A new concept?," *Signal Processing*, vol. 36, pp. 287–314, abr. 1994.
- [10] G. Darmois, "Analyse générale des liaisons stochastiques: étude particulière de l'analyse factorielle linéaire," *Revue de l'Institut International de Statistique*, vol. 21, no. 1/2, pp. 2–8, 1953.
- [11] V. Skitovich, "Linear forms of independent random variables and the normal distribution law," *Izvestiya Akademii Nauk SSSR. Seriya Matematicheskaya*, vol. 18, no. 2, pp. 185–200, 1954.
- [12] H. W. Gutch and F. J. Theis, "Uniqueness of linear factorizations into independent subspaces," *Journal of Multivariate Analysis*, vol. 112, pp. 48–62, 2012.
- [13] G. Chabriel, M. Kleinstuber, E. Moreau, H. Shen, P. Tichavsky, and A. Yeredor, "Joint Matrices Decompositions and Blind Source Separation: A survey of methods, identification, and applications," *IEEE Signal Processing Magazine*, vol. 31, pp. 34–43, maio 2014.
- [14] A. M. Kagan, Y. V. Linnik, and C. R. Rao, *Characterization Problems in Mathematical Statistics*. John Wiley & Sons, Inc., 1973.
- [15] M. Babaie-Zadeh, "Darmois-Skitovich theorem and its proof," *Sharif University of Technology*, jan. 2002. Disponível em: <http://ee.sharif.edu/~bss/DarmoisTheorem.pdf>. Acesso em: 30 jul. 2017.
- [16] H. Cramér, *Random Variables and Probability Distributions*. Cambridge University Press, 1937.
- [17] A. Papoulis and S. U. Pillai, *Probability, Random Variables and Stochastic Processes*. London: McGraw-Hill Europe, 4 ed., 2002.
- [18] D. Dugué, "Analyticité et convexité des fonctions caractéristiques," *Annales de l'Institut Henri Poincaré*, vol. 12, pp. 565–608, 1951.
- [19] J. V. Stone, *Independent Component Analysis: A Tutorial Introduction*. A Bradford Book, Cambridge, Massachusetts: MIT Press, 2004.
- [20] F. J. Theis, "Uniqueness of complex and multidimensional independent component analysis," *Signal Processing*, vol. 84, no. 5, pp. 951–956, 2004.
- [21] I. P. Mazur, "Skitovich–Darmois Theorem for Discrete and Compact Totally Disconnected Abelian Groups," *Ukrainian Mathematical Journal*, vol. 65, pp. 1054–1070, dez. 2013.