

UM MÉTODO NÃO-SUPERVISIONADO PARA IDENTIFICAÇÃO DAS CAPACIDADES DE CHOQUET COM APLICAÇÃO AO APOIO À DECISÃO MULTICRITÉRIO

Leonardo Tomazeli Duarte

Faculdade Ciências Aplicadas (FCA), Universidade Estadual de Campinas (UNICAMP)
Rua Pedro Zaccaria, 1300, CEP:13484-350, Limeira-SP
leonardo.duarte@fca.unicamp.br

RESUMO

Um problema fundamental que surge na aplicação da integral de Choquet em apoio à decisão multicritério diz respeito à identificação das capacidades de Choquet, que, *grossa modo*, traduzem a maneira com a qual os diversos critérios levados em conta no processo decisório devem ser combinados. No presente trabalho, é proposto um novo método de identificação que opera de modo não-supervisionado, isto é, considerando unicamente a informação presente na matriz de decisão. O método proposto se fundamenta em matrizes de similaridade estatística e na classe de capacidades 2-aditivas. A realização de um conjunto de experimentos numéricos, considerando tanto dados sintéticos quanto reais, indica que o método proposto permite uma associação mais natural, se comparada a abordagens anteriores, entre os tipos de interação entre critérios que usualmente são modelados pela integral de Choquet e a estrutura de correlação dos dados observados na matriz de decisão.

PALAVRAS CHAVE. Apoio à Decisão Multicritério. Integral de Choquet. Aprendizado não-supervisionado.

Tópicos: Apoio à Decisão Multicritério (ADM); Estatística (EST).

ABSTRACT

In multi-criteria decision aiding (MCDA), the Choquet integral has been used as an aggregation operator to deal with the case of interacting decision criteria. In this context, a practical problem that arises is related to the identification of the Choquet capacities (fuzzy measure). In the present work, we propose a novel capacity identification method that works in an unsupervised fashion, thus requiring only knowledge of the decision matrix. The introduced method relies on statistical similarity matrices and on the class of 2-additive capacities. As attested by numerical experiments with both synthetical and actual data, the proposed approach is conceptually simple to be implemented and allows one to associate criteria interactions that are typically modeled by the Choquet integral with the data correlation structure.

KEYWORDS. Multicriteria decision. Choquet integral. Unsupervised learning.

Paper topics: Apoio à Decisão Multicritério (ADM); Estatística (EST).

1. Introdução

Os elementos centrais em diversas abordagens de apoio à decisão multicritério (ADM) compreendem um conjunto de n alternativas, $A = \{a_1, \dots, a_n\}$, e um conjunto de m critérios, $C = \{c_1, \dots, c_m\}$. Usualmente, as entradas de uma estratégia de ADM são dadas por um conjunto de n vetores, $\mathbf{p}_k = [p_{k1}, \dots, p_{km}]$, onde p_{kj} representa o valor da alternativa k com relação ao critério j . A matriz de decisão $\mathbf{P} \in \mathbb{R}^{n \times m}$ contém todos os vetores \mathbf{p}_k , de modo que \mathbf{p}_t corresponde à t -ésima linha de \mathbf{P} .

Dados os conjuntos A e C , e a matriz \mathbf{P} , um dos objetivos em ADM é estabelecer uma estrutura de preferência sobre A de tal maneira que se torne possível ordenar os elementos de A [Figueira et al., 2004]. Nesse sentido, uma etapa fundamental, conhecida como agregação [Bouyssou et al., 2006], visa obter uma medida de valor global, g_k , para uma dada alternativa a_k . Em sua forma mais simples, a agregação é conduzida por funções lineares e pode ser vista como um processo de cálculo de média ponderada, no qual as ponderações dos critérios são dadas por constantes de escalas representadas pelo vetor $\mathbf{w} = [w_1, \dots, w_m]$.

A agregação baseada na integral de Choquet generaliza a agregação linear e permite ao usuário lidar com interação entre critérios [Grabisch e Labreuche, 2010; Marichal, 2000; Grabisch, 2000]. Diferentemente do paradigma de agregação linear, no qual o usuário deve definir as constantes de escalas \mathbf{w} , o uso da integral de Choquet requer a definição de parâmetros conhecidos como capacidades de Choquet, que, em breves termos, podem ser vistos como ponderações para todas as coalizações de critérios, isto é, o conjunto das partes $\mathcal{P}(C)$.

Dado que todos elementos de $\mathcal{P}(C)$, com exceção do subconjunto vazio e do próprio conjunto C , apresentam uma capacidade associada, o número total de capacidades de Choquet a serem estimadas é de $2^m - 2$ Marichal e Roubens [2000]. Ou seja, uma primeira dificuldade na estimação das capacidades de Choquet é que há um crescimento exponencial do número de parâmetros com relação ao número de critérios. Além disso, o significado das capacidades de Choquet nem sempre admite uma fácil interpretação, de modo que o usuário do método pode encontrar dificuldades em atribuir diretamente valores a todos elementos de $\mathcal{P}(C)$.

O problema de estimação das capacidades de Choquet pode ser abordado por uma ampla gama de métodos que podem ser classificados como: supervisionados, parcialmente supervisionados, e não-supervisionados. Nos métodos supervisionados, a determinação das capacidades requer do usuário dois tipos de informação: i) a importância de todos os critérios, bem como as (possíveis) interações entre eles; ii) as preferências do decisor por meio de relações de ordem entre as alternativas. Após a atribuição de tais informações, o problema de identificação das capacidades pode ser formulado como um problema de otimização no qual a função objetivo é dada por uma distância entre a saída fornecida pela integral de Choquet e um valor que explica as preferências do tomador de decisão. Exemplos de métodos supervisionados podem ser encontrados em [Grabisch, 1996; Marichal e Roubens, 2000; Grabisch et al., 2008; Anderson et al., 2014]. Embora extremamente útil num contexto de desagregação, métodos supervisionados podem ser muito restritivos em certos problemas práticos [Grabisch e Labreuche, 2010].

Nos métodos parcialmente supervisionados não é necessário definir uma ordem *a priori* entre as alternativas, de modo que o usuário deve unicamente atribuir valores que modelam a importância e as interações entre os critérios. Exemplos de métodos parcialmente supervisionados incluem os trabalhos recentes baseados no digrama do diamante de comparação par a par [Wu et al., 2014, 2015].

No paradigma não-supervisionado, que é o foco do presente trabalho, a identificação das capacidades é conduzida levando em conta exclusivamente a matriz de decisão \mathbf{P} . Logo, nenhum tipo de conhecimento prévio sobre as preferências do decisor é considerado. Dado que, na abordagem não-supervisionada, as interações entre os critérios são estimadas a partir dos dados presentes em \mathbf{P} , torna-se necessário recorrer a ferramentas estatísticas. Assim sendo, métodos não-supervisionados são indicados para situações nas quais o problema de decisão é de larga escala e,

além disso, há interações entre critérios, porém que são desconhecidas pelo tomador de decisão. O problema de avaliação de estudantes [Grabisch, 1996] é um exemplo representativo de cenários para os quais uma abordagem não-supervisionada se torna interessante.

Os primeiros métodos não-supervisionados de identificação de capacidade foram propostos por Kojadinovic [2004, 2008] e consideram medidas oriundas da teoria da informação, como a entropia de Shannon [Cover e Thomas, 1991]. Embora traga ao problema informações interessantes, como estatísticas de ordem superior, o uso de medidas da teoria da informação não é adequado para lidar com cenários práticos, dado que a estimação de tais medidas não é simples mesmo para casos nos quais o número de critério m é pequeno. Diante de tal dificuldade, Rowley et al. [2015] propuseram um método não-supervisionado baseado na matriz de correlação de Pearson. Ao trocar medidas difíceis de serem estimadas por correlações, o algoritmo resultante se torna extremamente eficiente em termos computacionais.

No presente trabalho, um novo método não-supervisionado de estimação de capacidades é introduzido. A abordagem proposta se fundamenta em matrizes de similaridade estatística e na classe de capacidades conhecida como 2-aditiva. Diferentemente do método introduzido em [Rowley et al., 2015], a abordagem proposta permite obter uma associação natural entre os tipos de interação que usualmente são modelados pela integral de Choquet (complementaridade e substituição) e as similaridades estatísticas obtidos a partir dos dados. Além disso, a proposta apresentada é de simples implementação.

O trabalho é organizado da seguinte maneira. A Seção 2 revisa rapidamente os aspectos principais referentes à aplicação da integral de Choquet em ADM. Na Seção 3, a abordagem proposta é descrita. Com o intuito de avaliar a proposta, experimentos numéricos são realizados na Seção 4. Por fim, a Seção 5 apresenta as conclusões do trabalho.

2. Aplicação da integral de Choquet no apoio à decisão multicritério

Um elemento central na integral de Choquet é o conceito de capacidade (ou medida *fuzzy*) sobre o conjunto $C = \{c_1, \dots, c_m\}$. A capacidade é definida como a função $\mu : \mathcal{P}(C) \rightarrow [0, 1]$ que satisfaz: 1) $\mu(\emptyset) = 0$; 2) $\mu(C) = 1$; e 3) $\mu(G) \leq \mu(H)$ para todos $G, H \subseteq C$ tal que $G \subseteq H$. Em termos práticos, a capacidade $\mu(H)$ quantifica a importância da coalizão de critérios H [Grabisch, 2000].

Dado o conjunto C , suas respectivas capacidades μ , e um vetor \mathbf{p}_k , a integral de Choquet, denotada por $D_{\mathbf{p}_k}^\mu$, pode ser aplicada para agregar os elementos de \mathbf{p}_k num único número real, da seguinte maneira:

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

onde o operador de índice (\cdot) representa a permutação que contém a seguinte ordem $p_{k(1)} \leq p_{k(2)} \leq \dots \leq p_{k(m)}$, e $T_{(t)} = \{c_{(t)}, \dots, c_{(m)}\}$ (assume-se que $p_{k(0)} = 0$). A integral de Choquet generaliza a média ponderada e média ponderada ordenada quando as capacidades são aditivas e cardinais, respectivamente [Yager, 1988; Grabisch, 2000].

Uma ferramenta útil para interpretar a integral de Choquet é dada pelo índice generalizado de interação [Grabisch, 2000]. Este índice é definido para qualquer coalizão H obtida do conjunto $C = \{c_1, \dots, c_m\}$ e é dado por:

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

Quando calculado para um dado critério c_i , o índice generalizado de interação resulta no índice de importância de Shapley, doravante denotado por $I_i = I(c_i)$. O índice de Shapley quantifica a

contribuição média do critério c_i em todas as coalizões, fornecendo assim uma medida da relevância do critério c_i no processo de agregação [Grabisch, 1996]. Os índices de Shapley são sempre positivos e sempre somam 1, isto é, $\sum_{i=1}^m I_i = 1$ [Grabisch, 2000].

A interação entre dois critérios c_i e c_j pode ser quantificada a partir do cálculo de (2) para a coalização $H = \{c_i, c_j\}$. A medida resultante, representada por $I_{ij} = I(\{c_i, c_j\})$ e chamada de índice de interação, nos permite interpretar o tipo de interação existente entre c_i e c_j [Grabisch e Labreuche, 2010]. De fato, quando $I_{ij} > 0$, a relação entre c_i e c_j é chamada *positiva* ou *complementar*, pois, neste caso, a presença de boas avaliações, para um dada alternativa, nos critérios c_i e c_j impactam sinergicamente o resultado da agregação. Por outro lado, se $I_{ij} < 0$, o que caracteriza uma interação *negativa* ou *substitutiva*, um valor alto na agregação, no que diz respeito aos critérios c_i e c_j , pode ser atingido mesmo quando apenas um desses critérios apresenta uma boa avaliação, ou seja, há um tipo de redundância entre c_i e c_j . Finalmente, $I_{ij} = 0$ caracteriza o caso em que não há interação entre os critérios c_i e c_j , e, portanto, a contribuição de tais critérios na agregação se dá de modo independente.

Em termos práticos, as preferências do tomador de decisão podem ser representadas pelo ajuste (ou estimação no caso não-supervisionado) das capacidades $\mu(H)$ ou dos índices generalizados de interação $I(H)$ para todas as coalizões de critérios, pois $I(H)$ é uma transformada inversível de $\mu(H)$, dada pela Equação (2) [Grabisch, 2000]. Logo, o número total de parâmetros a serem estimados é igual a $2^m - 2$, de modo que pode ser difícil interpretá-los facilmente. Com o intuito de paliar esse problema, é comum considerar classes restritas de capacidade nas quais o grau de liberdade estabelecido é inferior a $2^m - 2$ e para as quais os parâmetros livres possuem um significado claro. Nesse contexto, merecem destaque as classes de capacidade k -aditiva — uma capacidade é dita k -aditiva se $I(H) = 0$ para todas as coalizões H cuja cardinalidade é superior a k (assumindo que há pelo menos uma coalizão G de cardinalidade k para a qual $I(G) \neq 0$). Agregação via média ponderada pode ser vista como um caso especial de capacidade 1-aditiva [Grabisch e Labreuche, 2010].

No presente trabalho, nosso foco se dá na classe de capacidades 2-aditivas. Neste caso, as capacidades são funções unicamente dos índices de Shapley I_i e dos índices de interação I_{ij} , e, assim sendo, o número de parâmetros livre se reduz a $\binom{m}{1} + \binom{m}{2}$. Apesar de tal redução, classes de capacidade 2-aditiva são suficientemente flexíveis para modelar interações do tipo positiva e negativa.

A integral de Choquet assume uma forma interessante no caso de capacidades 2-aditivas [Grabisch e Labreuche, 2010], dada por

$$D_{\mathbf{P}_k}^\mu = \sum_{I_{ij} > 0} (p_{ki} \wedge p_{kj}) I_{ij} + \sum_{I_{ij} < 0} (p_{ki} \vee p_{kj}) |I_{ij}| + \sum_{i=1}^m p_{ki} \left(I_i - \frac{1}{2} \sum_{j \neq i} |I_{ij}| \right). \quad (3)$$

Essa representação é deveras útil para compreensão dos parâmetros I_{ij} como termos associados à parte não-linear no processo de agregação, enquanto que os índices de Shapley I_i estão diretamente associados à parte linear. Finalmente, é importante destacar que, no caso de capacidades 2-aditivas, é necessário considerar a seguinte restrição de modo a garantir a monotonicidade da integral de Choquet:

$$I_i - \frac{1}{2} \sum_{j \neq i} |I_{ij}| \geq 0, \quad \forall i. \quad (4)$$

3. Um novo método não-supervisionado de identificação das capacidades

3.1. Motivação

Métodos não-supervisionados de identificação das capacidades são baseados em um modelo estatístico da matriz de decisão \mathbf{P} , no qual as colunas de \mathbf{P} são vistas como realizações de

variáveis aleatórias que serão representadas por P_1, \dots, P_m . Portanto, torna-se possível estimar (através de dados) possíveis relações entre os diferentes critérios via, por exemplo, abordagens baseadas em teoria da informação [Kojadinovic, 2004] ou em estatísticas de segunda ordem [Rowley et al., 2015]. O presente trabalho foca na solução recente de [Rowley et al., 2015], uma vez que a abordagem proposta em [Kojadinovic, 2004] requer estimativa de entropias marginais, que é uma tarefa complicada na prática.

A abordagem introduzida em [Rowley et al., 2015] é baseada no coeficiente de correlação de Pearson, dado por

$$\rho(P_i, P_j) = Cov(P_i, P_j)(\sigma_{P_i}\sigma_{P_j})^{-1}, \quad (5)$$

onde $\rho(P_i, P_j) \in [-1, 1]$; σ_{P_t} e $Cov(P_i, P_j)$ denotam o desvio padrão de P_t e a covariância entre P_i e P_j , respectivamente. Duas variáveis são ditas descorrelacionadas, positivamente correlacionadas e negativamente correlacionadas quando $\rho = 0$, $\rho > 0$ e $\rho < 0$, respectivamente. Dada a variável aleatória multidimensional $\Pi = [P_1, \dots, P_m]$, é possível definir uma matriz de correlação par a par, $\mathbf{R}_\Pi \in \mathbb{R}^{m \times m}$, onde $[\mathbf{R}_\Pi]_{ij} = \rho(P_i, P_j)$. Note que $[\mathbf{R}_\Pi]_{ij} = 1$ para todo $i = j$.

O método proposto em [Rowley et al., 2015] explora a informação presente nos autovalores de \mathbf{R}_Π , denotados por $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m$, e nos autovalores de sub-matrizes obtidas de \mathbf{R}_Π . Mais precisamente, dada uma coalizção de critérios H , Rowley et al. [2015] propuseram a seguinte atribuição para $\mu(H)$:

$$\mu(H) = \frac{J(H)}{J(C)}, \quad (6)$$

onde

$$J(H) = |\{\tilde{\lambda}_q | \tilde{\lambda}_q \geq 1\}| + \sum_{q: \tilde{\lambda}_q < 1} \tilde{\lambda}_q. \quad (7)$$

Nesta equação, $\tilde{\lambda}_1 \geq \tilde{\lambda}_2 \geq \dots \geq \tilde{\lambda}_v$ correspondem aos autovalores da matriz de correlação, \mathbf{R}_H , que é a sub-matriz de \mathbf{R}_Π formada a partir das linhas e colunas associadas com os critérios presentes em H . O termo $|\{\tilde{\lambda}_q | \tilde{\lambda}_q \geq 1\}|$ corresponde ao número de autovalores que são maiores ou iguais a um. $J(C)$ também é calculado de acordo com (7), porém levando em conta os autovalores de \mathbf{R}_Π .

Conforme discutido em [Rowley et al., 2015], a ideia por trás de (6) é que $J(H)$ fornece uma estimativa do número de critérios que não interagem na coalizção H . Desta forma, se \mathbf{R}_Π é uma matriz diagonal, o que representa o caso em que os critérios são descorrelacionados, então $J(H) = |H|$ e, conseqüentemente, a capacidade resultante é aditiva. Por outro lado, se há dois critérios correlacionados, então, para todas as coalizções H que contêm esses dois critérios, espera-se que $J(H)$ seja aproximadamente igual ao número de critérios menos 1.

A abordagem proposta em [Rowley et al., 2015] apresenta alguns elementos interessantes, como sua simplicidade de implementação e o fato de que é possível mostrar que (6) satisfaz as condições de uma capacidade de Choquet. No entanto, dado que tal método se baseia exclusivamente na informação dos autovalores de \mathbf{R}_Π , as capacidades são definidas considerando apenas a quantidade de critérios que interagem entre si e não a natureza dessas interações. Conseqüentemente, a abordagem de [Rowley et al., 2015] pode resultar em capacidades iguais mesmo em cenários onde os sinais dos coeficientes de correlação entre os dados da matriz \mathbf{P} são diferentes.

3.2. Proposta

Diferentemente de [Rowley et al., 2015], o objetivo principal do método proposto é estabelecer associações entre medidas de similaridade estatística (como o coeficiente de correlação) obtidas a partir de \mathbf{P} e os índices de interação a serem considerados na integral de Choquet. De fato, mesmo numa situação não-supervisionada, na qual o tomador de decisão não dispõe de conhecimento sobre as interações entre os critérios, torna-se fundamental que ele entenda ao menos como a informação estatística estimada a partir dos dados \mathbf{P} é levada em conta no processo de agregação.

Um primeiro elemento do método proposto diz respeito à possibilidade de se considerar medidas de similaridade outras que o coeficiente de correlação de Pearson. De modo geral, é

possível considerar uma matriz de similaridade $\mathbf{F}_\Pi \in \mathbb{R}^{m \times m}$ associada a $\Pi = [P_1, \dots, P_m]$, onde $[\mathbf{F}_\Pi]_{ij} = \varsigma(P_i, P_j)$ e $\varsigma(P_i, P_j)$ é uma medida geral de similaridade estatística entre as variáveis P_i e P_j . \mathbf{F}_Π é uma matriz simétrica, isto é, $\varsigma(P_i, P_j) = \varsigma(P_j, P_i)$.

No presente trabalho, a medida de similaridade adotada é o coeficiente de correlação de Spearman [Corder e Foreman, 2014], que, para um conjunto de n amostras conjuntamente obtidas de P_i e P_j , é definido como o coeficiente de correlação de Pearson entre os dados ordenados. As motivações para o uso da correlação de Spearman são as seguintes: primeiramente, assim como no caso da correlação de Pearson, é simples definir estimadores estatísticos para a correlação de Spearman. Além disso, a correlação de Spearman é mais adequada, em relação à correlação de Pearson, para detecção de correlações não-lineares entre duas variáveis. De fato, a correlação de Spearman também se encontra no intervalo $[-1, 1]$, porém atinge seus valores máximo e mínimo quando há uma perfeita relação monótona positiva e negativa, respectivamente, entre P_i e P_j — no caso da correlação de Pearson, os valores máximos e mínimos são atingidos quando há uma relação linear perfeita entre as variáveis. Finalmente, a correlação de Spearman é mais robusta a *outliers*.

O segundo elemento de nossa proposta é que consideramos a classe de capacidades 2-aditivas. Além de reduzir o número de parâmetros a serem ajustados, o uso da classe 2-aditiva é mais conveniente dado que buscamos o estabelecimento de um relação das estatísticas dos dados com os índices de interação. Conforme discutido anteriormente, no caso de classes 2-aditivas, a definição de $\mu(H)$ pode ser realizada no domínio transformado dado pelos índices de Shapley I_i e pelos índices de interação I_{ij} . Com o intuito de ajustar I_i , será considerado o princípio da máxima entropia [MacKay, 2003], que advoga que, na ausência de informações *a priori* sobre relevância de cada um dos critérios, uma possível solução é considerar que a influência de tais critérios na agregação deve ser igualmente distribuídas. Tal princípio pode ser implementado da seguinte maneira:

$$I_i = \frac{1}{m}, \quad i = 1, \dots, m. \quad (8)$$

Note que $\sum_{i=1}^m I_i = 1$, respeitando, assim, a propriedade de soma igual a um dos índices de Shapley.

O ajuste dos índices de interação I_{ij} é feito levando em conta a correlação de Spearman, representada por $\varsigma(P_i, P_j)$. Portanto, após a estimação de \mathbf{F}_Π a partir da matriz de decisão \mathbf{P} , os índices de interação são atribuídos da seguinte maneira:

$$I_{ij} = -\varsigma(P_i, P_j). \quad (9)$$

A ideia de tal atribuição pode ser justificada da seguinte maneira. Uma forte correlação positiva entre dois critérios i e j significa que eles caminham na mesma direção, ou seja, valores baixos para o critério i ocorrem, via de regra, junto com valores baixos para o critério j — o mesmo para valores altos. Logo, os critérios i e j estão, de certo modo, medindo um fator latente único, e, assim sendo, torna-se natural estabelecer uma interação de substituição entre eles — isto é feito atribuindo um índice de interação negativo. Por outro lado, uma correlação negativa significa que as alternativas tendem a apresentar um valor alto para o critério i e baixo para o critério j (ou vice-versa). Neste caso, parece natural favorecer alternativas que apresentam bom desempenho em ambos os critérios, o que pode ser feito através da definição de um índice de interação positivo (interação complementar). Finalmente, não havendo correlação entre os critérios, espera-se que o índice de interação seja nulo.

As definições dos índices I_i e I_{ij} de acordo com (8) e (9), respectivamente, nos permitem

calcular a integral de Choquet para a i -ésima alternativa da seguinte maneira:

$$D_{\mathbf{P}_k}^\mu = \sum_{\varsigma(P_i, P_j) > 0} (p_{ki} \vee p_{kj}) |\varsigma(P_i, P_j)| - \sum_{\varsigma(P_i, P_j) < 0} (p_{ki} \wedge p_{kj}) \varsigma(P_i, P_j) + \sum_{i=1}^m p_{ki} \left(\frac{1}{m} - \frac{1}{2} \sum_{j \neq i} |\varsigma(P_i, P_j)| \right). \quad (10)$$

Quando $\varsigma(P_i, P_j) = 0, \forall i \neq j$, a integral de Choquet se torna idêntica à agregação via média ponderada.

Conforme discutido na Seção 2, uma restrição sobre I_i e I_{ij} (Equação (4)) deve ser satisfeita para o caso de capacidades 2-aditivas. Assim sendo, quando a condição (4) não é satisfeita, é possível considerar uma aproximação de \mathbf{F}_Π , denotada por $\hat{\mathbf{F}}_\Pi$, onde $[\hat{\mathbf{F}}_\Pi]_{ij} = \hat{\varsigma}(P_i, P_j)$. Tal aproximação pode ser obtida pela resolução do seguinte problema de otimização:

$$\begin{aligned} \min_{\hat{\mathbf{F}}_\Pi} \quad & \|\hat{\mathbf{F}}_\Pi - \mathbf{F}_\Pi\|_F \\ \text{sujeito a} \quad & \frac{1}{m} - \frac{1}{2} \sum_{j \neq i} |\hat{\varsigma}(P_i, P_j)| \geq 0, \quad i = 1, \dots, m \\ & \hat{\mathbf{F}}_\Pi \geq 0 \\ & \hat{\varsigma}(P_i, P_j) - \hat{\varsigma}(P_j, P_i) = 0. \end{aligned} \quad (11)$$

A função custo neste caso é distância de Frobenius entre $\hat{\mathbf{F}}_\Pi$ e \mathbf{F}_Π . A primeira restrição garante que (4) é satisfeita. A segunda restrição é necessária dado que matrizes de coeficientes de correlação são necessariamente positiva-definidas. Finalmente, a terceira restrição garante que $\hat{\mathbf{F}}_\Pi$ é simétrica. O problema de otimização expresso em (11) é um problema de programação não-linear com restrições de igualdade e de desigualdade, e, no presente trabalho, consideramos uma resolução via força bruta, pois o número de parâmetros nos experimentos foi pequeno.

O método não-supervisionado de identificação de capacidades proposto é resumido no Algoritmo 1.

Algoritmo 1 Método proposto

- 1: **Entrada** Matriz de decisão \mathbf{P}
 - 2: Estime \mathbf{F}_Π a partir de \mathbf{P} . Na proposta, isto é equivalente a estimar os coeficientes de Spearman entre P_i e P_j
 - 3: Defina I_i , para $i = 1, \dots, m$, como sendo $I_i = 1/m$;
 - 4: **if** a Equação (4) é satisfeita, defina $I_{ij} = -\varsigma(P_i, P_j), \forall i, j$
else
Obtenha $\hat{\mathbf{F}}_\Pi$ resolvendo (11)
Defina $I_{ij} = -\hat{\varsigma}(P_i, P_j), \forall i, j$
end if
 - 5: Calcule a integral de Choquet (10) para todas as alternativas (e as ordene de acordo com os valores obtidos)
-

4. Experimentos numéricos

Nesta seção, um conjunto de experimentos é realizado considerando dados sintéticos (Seção 4.1) e dados reais (Seção 4.2). Três métodos de agregação são testados: i) média aritmética simples, ii) a proposta de [Rowley et al., 2015], e iii) a presente proposta, descrita no Algoritmo 1.

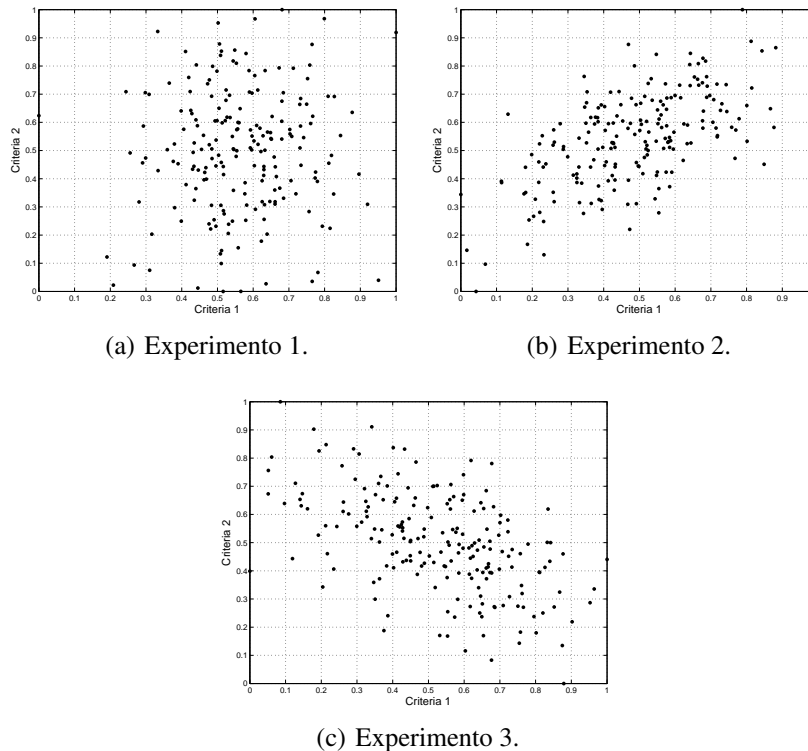


Figura 1: Gráficos de dispersão entre os critérios c_1 e c_2 .

4.1. Experimentos com dados sintéticos

Nessa situação, três cenários foram considerados, todos eles com $m = 3$ critérios e $n = 200$ alternativas. Em todas as situações, a matriz \mathbf{P} foi gerada a partir da amostragem de uma distribuição gaussiana multidimensional — a diferença entre os cenários considerados se encontram na matriz de correlação, definidas da seguinte maneira:

- *Experimento 1 (dados descorrelacionados)*: a matriz de correlação é diagonal;
- *Experimento 2 (dados correlacionados positivamente)*: o primeiro e o segundo critérios são correlacionados positivamente, isto é, $\rho(P_1, P_2) = 0.6$, porém descorrelacionados em relação ao terceiro critério — $\rho(P_1, P_3) = \rho(P_2, P_3) = 0$;
- *Experimento 3 (dados correlacionados negativamente)*: o primeiro e o segundo critérios são correlacionados negativamente, isto é, $\rho(P_1, P_2) = -0.6$, porém descorrelacionados com relação ao terceiro critério — $\rho(P_1, P_3) = \rho(P_2, P_3) = 0$;

Os dados observados foram normalizados entre $[0, 1]$, e, para cada critério, 0 e 1 correspondem aos valores mínimo e máximo, respectivamente (considerando um problema de maximização). Na Figura 1, os gráficos de dispersão dos critérios 1 e 2 para cada um dos experimentos são apresentados — cada ponto nestes gráficos corresponde a uma alternativa.

Dado que o modelo de geração de dados considerado é gaussiano, o que implica em dependência estatística linear, os coeficientes de Pearson e Spearman apresentam valores similares. De fato, no Experimento 1, as estimativas desses coeficientes foram dadas por $\rho(P_1, P_2) = 0.05$, $\rho(P_1, P_3) = -0.05$, $\rho(P_2, P_3) = -0.04$, $\varsigma(P_1, P_2) = 0.02$, $\varsigma(P_1, P_3) = -0.04$ e $\varsigma(P_2, P_3) = 0.00$. No Experimento 2, as seguintes estimativas foram obtidas $\rho(P_1, P_2) = 0.61$, $\rho(P_1, P_3) = 0.06$, $\rho(P_2, P_3) = 0.01$, $\varsigma(P_1, P_2) = 0.56$, $\varsigma(P_1, P_3) = 0.07$ e $\varsigma(P_2, P_3) = 0.01$. Finalmente, no Experimento 3, foram obtidas as seguintes estimativas: $\rho(P_1, P_2) = -0.52$, $\rho(P_1, P_3) = 0.06$, $\rho(P_2, P_3) = -0.08$, $\varsigma(P_1, P_2) = -0.52$, $\varsigma(P_1, P_3) = 0.07$ e $\varsigma(P_2, P_3) = -0.09$.

As capacidades obtidas para os métodos analisados são apresentadas na Tabela 1, que contem os resultados para os três experimentos considerados. Além disso, a Tabela 2 apresenta os índices de Shapley e os índices de interação obtidos nos experimentos — conforme discutido anteriormente, há uma transformada inversa que associa as capacidades aos índices de interação, de modo que a Tabela 1 pode ser obtida a partir da Tabela 2 e vice-versa.

Tabela 1: Capacidades estimadas (experimentos com dados sintéticos).

	Média aritmética			Rowley et al. [2015]			Proposta		
	Exp 1	Exp 2	Exp 3	Exp 1	Exp 2	Exp 3	Exp 1	Exp 2	Exp 3
$\mu(c_1)$	0.33	0.33	0.33	0.34	0.42	0.41	0.32	0.65	0.11
$\mu(c_2)$	0.33	0.33	0.33	0.34	0.42	0.41	0.34	0.62	0.03
$\mu(c_3)$	0.33	0.33	0.33	0.34	0.42	0.41	0.31	0.37	0.32
$\mu(c_1, c_2)$	0.67	0.67	0.67	0.67	0.58	0.60	0.64	0.71	0.66
$\mu(c_1, c_3)$	0.67	0.67	0.67	0.67	0.81	0.79	0.67	0.95	0.36
$\mu(c_2, c_3)$	0.67	0.67	0.67	0.68	0.83	0.78	0.66	0.98	0.44
$\mu(c_1, c_2, c_3)$	1	1	1	1	1	1	1	1	1

Tabela 2: Índices de Shapley e índices de interação (experimentos com dados sintéticos).

	Média aritmética			Rowley et al. [2015]			Proposta		
	Exp 1	Exp 2	Exp 3	Exp 1	Exp 2	Exp 3	Exp 1	Exp 2	Exp 3
I_1	0.33	0.33	0.33	0.33	0.29	0.30	0.33	0.33	0.33
I_2	0.33	0.33	0.33	0.34	0.30	0.30	0.33	0.33	0.33
I_3	0.33	0.33	0.33	0.33	0.41	0.39	0.33	0.33	0.33
I_{12}	0	0	0	-0.01	-0.24	-0.18	-0.02	-0.56	0.52
I_{13}	0	0	0	0.00	-0.01	0.00	0.04	-0.07	-0.07
I_{23}	0	0	0	0.01	0.01	-0.01	0.00	-0.01	0.08

Uma primeira observação sobre as Tabelas 1 e 2 é que, no Experimento 1, os métodos investigados obtiveram resultados similares, o que pode ser atribuído a inexistência de correlação entre os dados. Por outro lado, devido à existência de uma correlação positiva entre os critérios c_1 e c_2 no Experimento 2, as capacidades estimadas por [Rowley et al., 2015] e pelo método proposto são claramente não-aditivas, de modo que se observa um índice de interação I_{12} negativo, o que conduz ao estabelecimento de uma interação de substituição entre c_1 e c_2 . Portanto, conforme observado na Tabela 3, que mostra as dez primeiras alternativas com maiores valores globais, as ordenações obtidas pelos métodos baseados na integral de Choquet são menos próximas à ordenação obtida pela média aritmética.

Novamente, no Experimento 3, tanto o método de [Rowley et al., 2015] quanto a presente proposta resultaram em capacidades não-aditivas. Todavia, enquanto o método de [Rowley et al., 2015] forneceu um modelo de interação muito próximo ao obtido, pelo mesmo método, no Experimento 2 (interação substitutiva entre c_1 e c_2), a presente proposta culminou num índice de interação I_{12} positivo e, logo, numa interação complementar entre c_1 e c_2 . De fato, as matrizes de correlação associadas aos Experimentos 2 e 3 possuem autovalores idênticos, e, assim sendo, o método de [Rowley et al., 2015] não é capaz de distinguir a natureza da relação estatística existente entre os critérios c_1 e c_2 . Tal característica explica as diferentes ordenações obtidas no Experimento 3, conforme mostra a Tabela 3.

4.2. Experimentos com dados reais

Nesta seção, consideramos o problema de ordenação de países e grupos de países (por exemplo, zona Euro) considerando a maximização dos seguintes critérios: 1) produto interno bruto

Tabela 3: Ordenação das alternativas (experimentos com dados sintéticos).

Ordem	Média aritmética			Rowley et al. [2015]			Proposta		
	Exp 1	Exp 2	Exp 3	Exp 1	Exp 2	Exp 3	Exp 1	Exp 2	Exp 3
1	a_{132}	a_{48}	a_{37}	a_{132}	a_{48}	a_{37}	a_{132}	a_{48}	a_{37}
2	a_{42}	a_{143}	a_{101}	a_{42}	a_{43}	a_{62}	a_{42}	a_{20}	a_{101}
3	a_{159}	a_{43}	a_{62}	a_{159}	a_{20}	a_{13}	a_{159}	a_{143}	a_{25}
4	a_{176}	a_{20}	a_{89}	a_{176}	a_{135}	a_{89}	a_{176}	a_{43}	a_{97}
5	a_{55}	a_{135}	a_{13}	a_{55}	a_{143}	a_{178}	a_{88}	a_{135}	a_{62}
6	a_{88}	a_{196}	a_{25}	a_2	a_{163}	a_{84}	a_2	a_{46}	a_{124}
7	a_2	a_{75}	a_{84}	a_{88}	a_{196}	a_{101}	a_{55}	a_{163}	a_{19}
8	a_{195}	a_{141}	a_9	a_{195}	a_{46}	a_{65}	a_{195}	a_{199}	a_{185}
9	a_{19}	a_{199}	a_{178}	a_{19}	a_{199}	a_{107}	a_4	a_{196}	a_{27}
10	a_4	a_{163}	a_{19}	a_4	a_{40}	a_9	a_{19}	a_{141}	a_{13}

(PIB) *per capita*; 2) inverso das emissões de gás CO₂; 3) Porcentagem da terra ocupada por florestas¹. Os dados, que foram obtidos da página do Banco Mundial (<http://data.worldbank.org/>) e se referem ao ano de 2011, foram normalizados entre [0, 1]. Os gráficos de dispersão entre os critérios são exibidos na Figura 2. Há um forte correlação negativa entre o PIB *per capita* e o inverso de emissões de CO₂, que é confirmada pelos coeficientes de correlação: $\rho(P_1, P_2) = -0.64$, $\rho(P_1, P_3) = 0.03$, $\rho(P_2, P_3) = 0.05$, $\zeta(P_1, P_2) = -0.89$, $\zeta(P_1, P_3) = 0.09$ e $\zeta(P_2, P_3) = 0.00$. Note que, neste caso, como os dados não seguem uma distribuição gaussiana, os coeficientes de Sperman assumem valores mais elevados se comparados aos coeficientes de Pearson.

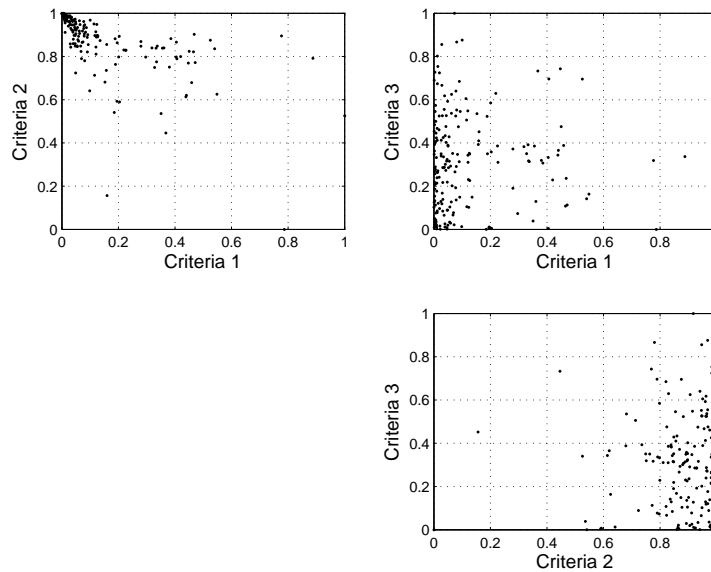


Figura 2: Gráficos de dispersão entre os critérios (dados reais).

Na Tabela 4, os índices de Shapley e os índices de interação obtidos por cada um dos métodos são apresentados. A presente proposta resultou num valor positivo de I_{12} , o que define portanto uma interação complementar entre o PIB *per capita* e o inverso das emissões de CO₂. Ou seja, dado que altos valores de PIB estão associados a altos valores de emissão de CO₂, a presente proposta favorece países que apresentam um bom desempenho em ambos os critérios.

¹Esses critérios foram adotados meramente para ilustrar as diferenças entre os métodos sob investigação, de modo que não se argumenta, no presente trabalho, que tal escolha é a melhor maneira para se ordenar países.

Diferentemente da presente proposta, o método de [Rowley et al., 2015] resultou num valor negativo de I_{12} (interação substitutiva) e, logo, de acordo com esse método, um país precisa apenas apresentar um bom desempenho ou com relação ao PIB *per capita* ou com relação às emissões de CO₂. Consequentemente, conforme exibido na Tabela 5, os dez primeiros países selecionados pelos métodos baseados na integral de Choquet foram consideravelmente diferente. Um exemplo interessante é o caso da Austrália. Este país foi ordenado nas posições 175, 90, e 10 pela média aritmética, pelo método de [Rowley et al., 2015] e pela presente proposta, respectivamente. Tal diferença pode ser explicada pelos valores obtidos por esse país em cada um critérios, que foram dados por $\mathbf{p}_8 = [0.55 \ 0.62 \ 0.16]$. Dessa forma, a despeito do desempenho ruim com relação ao critério ligado à porcentagem de área floresta, a Austrália obteve uma boa posição pela presente proposta pois foi um dos poucos países que apresentaram bom desempenho no PIB *per capita* e na emissão de CO₂.

Tabela 4: Índices de Shapley e índices de interação (experimentos com dados reais).

	Média aritmética	Rowley et al. [2015]	Proposta
I_1	0.33	0.29	0.33
I_2	0.33	0.29	0.33
I_3	0.33	0.42	0.33
I_{12}	0	-0.26	0.65
I_{13}	0	0.00	0.00
I_{23}	0	0.03	0.00

Tabela 5: Ordenação das alternativas (experimentos com dados reais).

Ordem	Média aritmética	Rowley et al. [2015]	Proposta
1	Suécia	Suriname	Noruega
2	Noruega	Gabão	Suíça
3	Suíça	Guiana	Suécia
4	Suriname	Ilhas Salomão Islands	Finlândia
5	Finlândia	Suécia	Japão
6	Gabão	Papua-Nova Guiné	Brunei
7	Japão	Guiné-Bissau	Luxemburgo
8	Luxemburgo	Butão	Áustria
9	Guiana	Finlândia	Canadá
10	Ilhas Salomão	Rep. Dem. Congo	Austrália

5. Conclusões

Este trabalho introduziu um novo método não-supervisionado para identificação das capacidades de Choquet. A proposta realizada se fundamenta na classe de capacidades 2-aditivas e, diferentemente de outros métodos disponíveis, possibilita uma associação direta entre a estrutura de correlação dos dados e os tipos de interação usualmente modelados pela integral de Choquet. Experimentos com dados sintéticos e reais confirmaram o interesse da proposta.

Referências

Anderson, D. T., Price, S. R., e Havens, T. C. (2014). Regularization-based learning of the Choquet integral. In *IEEE International Conference on Fuzzy Systems*, p. 2519–2526. IEEE.

- Bouyssou, D., Marchant, T., Pirlot, M., Tsoukias, A., e Vincke, P. (2006). *Evaluation and decision models with multiple criteria: Stepping stones for the analyst*, volume 86. Springer.
- Corder, G. W. e Foreman, D. I. (2014). *Nonparametric statistics: a step-by-step approach*. John Wiley & Sons.
- Cover, T. M. e Thomas, J. A. (1991). *Elements of information theory*. Wiley-Interscience.
- Figueira, J., Greco, S., e Ehrgott, M., editors (2004). *Multiple criteria decision analysis: state of the art surveys*, volume 78. Springer.
- Grabisch, M. (1996). The application of fuzzy integrals in multicriteria decision making. *European journal of operational research*, 89(3):445–456.
- Grabisch, M. (2000). A graphical interpretation of the Choquet integral. *IEEE Transactions on Fuzzy Systems*, 8(5):627–631. ISSN 1063-6706.
- Grabisch, M., Kojadinovic, I., e Meyer, P. (2008). A review of methods for capacity identification in Choquet integral based multi-attribute utility theory: Applications of the kappalab R package. *European Journal of Operational Research*, 186(2):766–785.
- Grabisch, M. e Labreuche, C. (2010). A decade of application of the Choquet and Sugeno integrals in multi-criteria decision aid. *Annals of Operations Research*, 175(1):247–286.
- Kojadinovic, I. (2004). Estimation of the weights of interacting criteria from the set of profiles by means of information-theoretic functionals. *European Journal of Operational Research*, 155(3): 741–751.
- Kojadinovic, I. (2008). Unsupervised aggregation of commensurate correlated attributes by means of the Choquet integral and entropy functionals. *International Journal of Intelligent Systems*, 23 (2):128–154.
- MacKay, D. J. C. (2003). *Information theory, inference, and learning algorithms*. Cambridge University Press.
- Marichal, J.-L. (2000). An axiomatic approach of the discrete Choquet integral as a tool to aggregate interacting criteria. *IEEE Transactions on Fuzzy Systems*, 8(6):800–807.
- Marichal, J.-L. e Roubens, M. (2000). Determination of weights of interacting criteria from a reference set. *European journal of operational Research*, 124(3):641–650.
- Rowley, H. V., Geschke, A., e Lenzen, M. (2015). A practical approach for estimating weights of interacting criteria from profile sets. *Fuzzy Sets and Systems*, 272:70 – 88.
- Wu, J., Yang, S., Zhang, Q., e Ding, S. (2015). 2-additive capacity identification methods from multicriteria correlation preference information. *IEEE Transactions on Fuzzy Systems*, 23(6): 2094–2106. ISSN 1063-6706.
- Wu, J., Zhang, Q., Du, Q., e Dong, Z. (2014). Compromise principle based methods of identifying capacities in the framework of multicriteria decision analysis. *Fuzzy Sets and Systems*, 246: 91–106.
- Yager, R. R. (1988). On ordered weighted averaging aggregation operators in multicriteria decision-making. *IEEE Transactions on Systems, Man and Cybernetics*, 18(1):183–190.