

Integral de Choquet Discreta: Perspectiva Pedagógica

Luiz Alves de Souza Neto *
Universidade Federal do Maranhão
São Luís - MA - Brasil

Marcos Nascimento Azevedo †
Universidade Federal do Maranhão
São Luís - MA - Brasil

28 de outubro de 2022

Resumo

Tomada de decisão multicritério envolve um conjunto de métodos que ajudam a resolver problemas que possuem vários objetivos frequentemente conflitantes, com múltiplos critérios, incertezas, várias etapas, diferentes agentes e um conjunto finito de alternativas e diversos indivíduos afetados pela decisão. Dessa maneira, o método da integral de Choquet, dentro dos problemas de tomadas de decisão, pode fornecer uma estimativa esperada de um evento incerto. Diversas aplicações, na literatura, mostram que a função de agregação da integral de Choquet em processamento de imagem, processamento de vídeo e visão computacional, na teoria da decisão comportamental. Uma aplicação crescente, da integral de Choquet, envolve a ordenação de um conjunto de alternativas de informações de n amostras catalogadas com o propósito de mapear cada entrada de alternativas i com um *score* de saída. Dessa forma, neste artigo apresentará a integral de Choquet como uma média aritmética e como o agregador OWA (Ordered weighted averaging) e a ordenação preferencial na escolha de alunos dado o conjunto de notas. Outras duas aplicações consideram uma abordagem supervisionada para estimar as medidas de capacidades de Choquet aplicando um modelo de otimização. O primeiro resultado identificou pesos das notas das disciplinas e as suas interações com outros tipos de critérios como o número de ausências. O segundo resultado é em relação ao uso da integral de Choquet como um preditor em um estudo de casos de COVID-19.

*luiz.alves@ufma.br

†marcos.azevedo@ufma.br

Como perspectiva, a aplicação da regularização baseados na norma l_p .

Palavras-Chaves: multicritério, integral de Choquet, otimização.

Abstract

Multicriteria decision making involves a set of methods that help solve problems that have several often conflicting goals, with multiple criteria, uncertainties, several stages, different agents and a finite set of alternatives and several individuals affected by the decision. In this way, the Choquet integral method, within decision-making problems, can provide a way to measure the expected utility of an uncertain event. Several applications in the literature show that the aggregation function of the Choquet integral in image processing, video processing and computer vision, in behavioral decision theory. A growing application of the Choquet integral involves ordering a set of information alternatives from n cataloged samples for the purpose of mapping each input of alternatives i with an output *score*. Thus, the approach of the application of this work focuses, from a numerical base of student data, to verify the composition of the evaluations in the disciplines in the final grade of the students, assigning weights as the measure of Choquet's ability.

Keywords: multicriteria, Choquet integral, optimization.

1 Introdução

Desde a sua introdução em 1974 por Sugeno (Sugeno), o conceito de medida fuzzy tem sido frequentemente utilizado na tomada de decisão multicritério. A primeira aplicação, do próprio Sugeno, foi sobre avaliação de rostos de mulheres (Sugeno). As medidas fuzzy podem modelar em algum sentido um tipo de interação entre os critérios.

A integral de Choquet é uma extensão da integral clássica de Lebesgue, e, portanto, da conhecida soma ponderada. As propriedades das integrais de Choquet e Sugeno como um operador de agregação foram estudadas em profundidade, e sua conexão com os operadores OWA (Ordered weighted averaging) em sua forma aditiva usual ou forma mínima ponderada (GRABISCH; ROUBENS, 2008).

A integral de Choquet foi apresentada na década de 1950 por Gustave Choquet (CHOQUET, 1954; GRABISCH; ROUBENS, 2000) e começou a ser aplicada em Métodos de Análise Multicritério, (MCDA - Multiple criteria decision analysis), na década de 1990, onde os primeiros trabalhos permaneceram num nível intuitivo e experimental (GRABISCH; LABREUCHE, 2010).

Um aspecto central no uso da integral de Choquet em MCDA consiste no problema da estimação das medidas fuzzy associadas com a integral, problema esse que também é conhecido como identificação das capacidades de Choquet, que, consequentemente, cai em um problema de otimização.

A Integral de Choquet é intensamente utilizada em MCDA como operador não-linear de agregação (PELEGRINA et al., 2020). Dessa forma, é possível modelar sinergia e/ou redundância entre critérios, que poderá ser aplicada para considerar informações adicionais sobre os dados, evitando resultado enviesado (GRABISCH; ROUBENS, 2008).

Uma das dificuldades na identificação das capacidades é a de que o problema inverso resultante é mal-posto. Portanto, a solução obtida após a otimização pode não ser única, ou seja, diferentes conjuntos de capacidade de Choquet podem resultar num mesmo valor após a etapa de agregação. Uma abordagem comum para resolver problemas inversos mal-postos é considerar regularização, do tipo l_1 ou l_2 , ver (TIBSHIRANIT, 1996).

2 Metodologia

O presente artigo tem o propósito de realizar uma aplicação elementar ou introdutória da técnica da integral de Choquet enfatizando o uso das principais definições e propriedades básicas para o problema.

Vamos fazer algumas breves menções a aplicações, que estão relacionadas ao nosso contexto e logo em seguida, apresentaremos os conceitos básicos da integral de Choquet e sua generalização com o agregador OWA e média aritmética e aplicações com a ordenação preferencial de alunos a partir das notas em determinadas disciplinas.

A identificação das medidas de capacidades de Choquet é amplamente estudada nos artigos (OLIVEIRA; DUARTE; ROMANO, 2020; PELEGRINA et al., 2020; GRABISCH, 2016) e suas aplicações com experimentos computacionais em interações com até dois critérios, ou seja, capacidade 2-aditivas. Em (PELEGRINA et al., 2020) eficiência da integral de Choquet 2-aditiva é comparada com o modelo multilinear.

Nos artigos pesquisados, o método para a extração das medidas de capacidades da integral de Choquet é dada com a utilização de programação quadrática supervisionada com um *solver*.

3 Trabalhos Relacionados

Em diversos estudos relacionam a aplicação da integral de Choquet em contextos de várias áreas do conhecimento, uma delas é a proposta por (BUENO; DIMURO; BORGES, 2018) em redução e dimensionamento de imagens. Como ferramenta de dinâmica de populacional mostrado modelo de crescimento logístico via equação diferencial com relação a uma medida fuzzy no trabalho de (LAIATE; BARROS; ESMI, 2016).

A integral de Choquet quando o conjunto de amostras são discretas foi relatada no estudo de (MAYAG; BOUYSSOU, 2020). A combinação de métodos entre o número fuzzy triangular, método MULTIMORA e integral de Choquet para superar as limitações na avaliação de riscos ocupacionais (WANG; LIU; QIN, 2018).

A integral de Choquet também foi utilizada em combinação com método aditivo de análise envoltória de dados DEA (Data Envelopment Analysis) para julgar preferências descrito em (PEREIRA; FIGUEIRA; MARQUES, 2020). Em seleção de instrumentos e dispositivos médicos perante ambiente de incertezas (BüYüKÖZKAN; GÖÇER, 2019). E avaliação da satisfação geral dos pacientes é exposta como aplicação da integral Choquet discreta com medida fuzzy no artigo (HUANG et al., 2010).

Neste artigo iremos abordar a integral de Choquet na sua forma discreta. Uma abordagem superficial e elementar sobre a integração via Lebesgue será suficiente nesse momento, uma vez que o principal objetivo principais deste trabalho é mostrar como manipular um operador da integral de Choquet para aplicações computacionais.

Em (DUARTE; TOREZZAN, 2020) está definido as condições de fronteiras e monotonicidade de uma medida discreta fuzzy para existência de uma integral discreta de Choquet. Já em (GRABISCH; ROUBENS, 2000) temos uma aplicação com três critérios senfo a medida μ a medida de capacidade de Choquet de um modelo de interação de critérios e relacionada com a importância da coalizão entre os critérios de um determinado conjunto.

Definição 1 *Seja o conjunto $N = \{1, 2, 3, \dots, n\}$. Uma medida fuzzy discreta é uma Função $\mu : 2^N \rightarrow [0, 1]$ que é monotona (i.e $\mu(A) \leq \mu(B)$ nem sempre $A \subset B$) e satisfaz $\mu(\emptyset) = 0$ e $\mu(N) = 1$ (condição de fronteira).*

Definição 2 *Seja μ uma medida fuzzy em um conjunto N cujos elementos são denotados como $P_{i1}, P_{i2}, \dots, P_{im}$, $i = \{1, 2, \dots, n\}$. A integral discreta de Choquet de uma função $f : X \rightarrow R^+$ em relação a μ é dada por:*

$$CI_i = \sum_{j=1}^m (P_{i(j)} - P_{i(j-1)}) \mu(D_{(j)}), \quad (1)$$

P_{ij} é o valor de uma amostra i para o critério j . O operador de índice representa uma permutação com a seguinte ordem: $P_{i(1)} \leq P_{i(2)} \leq \dots \leq P_{i(m)}$ e $D_{(j)} = \{c_{(t)}, c_{(t+2)}, \dots, c_{(m)}\}$ e o m é o número de critérios de uma determinada amostra i e os elementos do conjunto D representam os critérios da alternativa ou amostra.

Devido ao caráter combinatório dessa medida, para um problema com m critérios, teremos 2^m elementos que a compõem.

A propriedade de monotonicidade, significa que o procedimento de se adicionar um novo critério em uma coalizão não pode diminuir o valor da capacidade para aquela coalizão.

Seja o exemplo com três critérios (GRABISCH; ROUBENS, 2000), $D = \{c_1, c_2, c_3\}$. Sendo $\mu(\cdot)$ a medida de capacidade de Choquet de um modelo de interação

de critérios e relacionada com a importância da coalizão entre os critérios do conjunto D é dada por:

- $\mu(c_1, c_2) < \mu(c_1) + \mu(c_2)$ (Redundância)
- $\mu(c_1, c_2) > \mu(c_1) + \mu(c_2)$ (Sinergia)
- $\mu(c_1, c_2) = \mu(c_1) + \mu(c_2)$ (Ausência de interação)

Onde a capacidade de Choquet é uma função μ que satisfaz as seguintes propriedades:

- $\mu(c_1) \leq \mu(c_1, c_2)$
- $\mu(c_2) \leq \mu(c_1, c_2)$
- $\mu(c_1, c_2) \leq \mu(c_1, c_2, c_3)$
- ...
- $\mu(c_1, c_2, c_3) = 1$ valor máximo
- $\mu(\emptyset) = 0$ valor mínimo

3.1 Multicritérios e Agregação

Conforme o trabalho publicado em (OLIVEIRA; DUARTE; ROMANO, 2020), podemos considerar a integral de Choquet como um operador agregador. Podemos, sem perda de generalidade, considerar uma matriz como um conjunto de entradas cujos elementos são variáveis de tomada de decisão multicritérios. Dessa maneira, considere $P \in R_{n \times m}$ como uma matriz de decisão, onde n é o número de amostras coletadas, pertencentes ao conjunto $A = \{a_1, a_2, \dots, a_n\}$, e m é o número de critérios, pertencentes ao conjunto $D = \{c_1, c_2, \dots, c_m\}$. Cada elemento P_{ij} da matriz representa a utilidade da alternativa i com relação ao critério j . O objetivo do processo de agregação é encontrar, para uma dada alternativa a_i , o seu *score*, dada por s_i . Isso é feito utilizando-se um operador de agregação $F(\cdot)$, aplicado nos critérios referentes a essa alternativa. Esses critérios estão associados a um conjunto de parâmetros w , que podem representar, por exemplo, as ponderações de cada critério no processo de agregação.

$$\begin{bmatrix} P_{11} & P_{12} & P_{13} & \dots & P_{1n} \\ P_{21} & P_{22} & P_{23} & \dots & P_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ P_{n1} & P_{n2} & P_{n3} & \dots & P_{nm} \end{bmatrix} \longrightarrow s_i = F(P_{i1}, \dots, P_{im}, w)$$

De acordo com (GRABISCH, 2016), o objetivo do tomador de decisão é modelar as preferências entre as alternativas, da seguinte forma: $a_i \succ a_{i'} \Leftrightarrow s_i > s_{i'}$, para todo $a_i, a_{i'} \in A$.

Tomando a matriz de decisão conforme já ilustrado nesta seção, podemos fazer algumas adaptações simbólicas para a integral de Choquet dada uma amostra i , da seguinte forma dada por (OLIVEIRA; DUARTE; ROMANO, 2020):

$$\gamma_i^\mu(P_{i1}, P_{i2}, \dots, P_{im}) = \sum_{j=1}^m (P_{i(j)} - P_{i(j-1)})\mu(D_j). \quad (2)$$

Podemos definir a forma vetorial da integral de Choquet, que para uma amostra i é dada por:

$$\gamma_i^\mu(P_{i1}, \dots, P_{im}) = m_i^t \mu \quad (3)$$

Onde,

$$m_i = \begin{bmatrix} 0 \\ \vdots \\ P_{i(m)} - P_{i(m-1)} \\ \vdots \\ P_{i(2)} - P_{i(1)} \\ \vdots \\ P_{i(i)} - P_{i(0)} \end{bmatrix}$$

As diferenças nos elementos da matriz m devem estar na mesma posição correspondentes às suas capacidades que são definidas numa sequência lexicográfica, como segue:

$$\mu = [\mu(\emptyset), \mu(\{c_1\}), \mu(\{c_2\}), \dots, \mu(\{c_1, c_2\}), \dots, \mu(\{c_1, c_2, \dots, c_m\})]^t \quad (4)$$

Por simplicidade, vamos trocar algumas notações e podemos assumir que $\mu(c_j) = \mu_j$. Dessa maneira, temos:

$$\mu = [\mu_\emptyset, \mu_1, \mu_2, \mu_3, \dots, \mu_{12\dots m}]^t. \quad (5)$$

Então a integral de Choquet pode ser dada, com a matriz de decisão M , pode ser representada, para todas as alternativas, como:

$$\gamma^\mu(P_{i1}, P_{i2}, \dots, P_{im}) = M^t \mu. \quad (6)$$

e a matriz M é:

$$\begin{bmatrix} 0 & 0 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ P_{1(m)} - P_{1(m-1)} & P_{2(m)} - P_{2(m-1)} & \dots & P_{n(m)} - P_{n(m-1)} \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ P_{1(1)} - P_{1(0)} & P_{2(2)} - P_{2(0)} & \dots & P_{n(1)} - P_{n(0)} \end{bmatrix}$$

A propriedade de monotonicidade, já definida anteriormente, significa que o fato de se adicionar um novo critério em uma coalizão não pode diminuir o valor da capacidade para aquela coalizão (GRABISCH; ROUBENS, 2000). Por exemplo, usando a notação, temos que $\mu_1 \leq \mu_{12}$.

4 Outros Operadores de agregação

Nesta seção veremos o agregador OWA, empregado em combinações lineares de variáveis ordenadas. Essa técnica será dada como um caso particular da integral de Choquet.

4.1 Agregação OWA (Ordered weighted averaging)

O operador OWA foi desenvolvido por (YAGER; FILEV, 1999) para possibilitar soluções intermediárias, que não são completamente restritivas ou aditivas.

O operador OWA de dimensão n corresponde a uma função de mapeamento $[0, 1]^n \rightarrow [0, 1]$, que agrega um conjunto de n valores normalizados a_1, a_2, \dots, a_n .

Esse tipo de agregação realiza uma ordenação de um conjunto de alternativas de entradas de um conjunto. Esse método é basicamente uma combinação linear dos valores ordenados. Os pesos não estão mais associados aos valores dos critérios, isto é, w_1 não é mais peso do critério 1, o w_1 será o peso do critério de menor valor.

$$OWA_i(P_{i1}, \dots, P_{im}) = \sum_{j=1}^m w_j P_{i(j)}. \quad (7)$$

Onde $\sum_j w_j = 1$ e $P_{i(1)} \leq P_{i(2)} \leq \dots \leq P_{i(m)}$. Entre os membros dessa classe de operadores de agregação estão os operadores de máximo (max), mínimo (min) e a média aritmética (MA). Essas medidas podem ser todas generalizadas conforme (YAGER; FILEV, 1999).

Se $w_m = 1$ e $w_j = 0$, para todo $j \neq m$,

$$OWA_i(P_{i1}, \dots, P_{im}) = \max(P_{i1}, \dots, P_{im}). \quad (8)$$

Se $w_1 = 1$ e $w_j = 0$, para todo $j \neq 1$.

$$OWA_i(P_{i1}, \dots, P_{im}) = \min(P_{i1}, \dots, P_{im}). \quad (9)$$

Se $w_j = \frac{1}{m}$, para todo $j \in [1, m]$. Chamamos média.

5 Relação da Integral de Choquet com o OWA

Uma propriedade interessante da integral de Choquet é que ela generaliza o agregador linear e o OWA.

$$\mu(G) = \sum_j^{|G|} w_j. \quad (10)$$

Onde $|G|$ representa a cardinalidade do conjunto G . Dessa forma, podemos afirmar que o operador OWA com pesos w_1, \dots, w_m é uma integral de Choquet, cujas capacidades são definidas por (GRABISCH; LABREUCHE, 2007):

$$\mu(G) = \sum_j^{|G|-1} w_{m-j}. \quad (11)$$

5.1 Aplicações da integral de Choquet na forma discreta

Nesta seção iremos exemplificar algumas aplicações com situações hipotéticas com integral de Choquet como uma média aritmética, OWA e a ordenação de alunos.

5.2 Integral de Choquet e Média Aritmética

Vamos considerar para o nosso exemplo os seguintes pesos, $w_1 = w_2 = w_3 = \frac{1}{3}$, e $P_1 = 1$, $P_2 = 3$, e $P_3 = 5$. Utilizando a equação para a média aritmética (MA), temos o que segue.

$$MA = \left(\frac{1}{3} * 1\right) + \left(\frac{1}{3} * 3\right) + \left(\frac{1}{3} * 5\right) = 3$$

Agora, procedendo para o cálculo da integral de Choquet, encontramos as capacidades por meio da equação já mostrada nas seções anteriores.

$$\mu(\{c_1\}) = \mu(\{c_2\}) = \mu(\{c_3\}) = \sum_1^1 w_j = w_1 = \frac{1}{3}$$

$$\mu(\{c_1, c_2\}) = \mu(\{c_1, c_3\}) = \mu(\{c_2, c_3\}) = \sum_{j=1}^2 w_j = w_1 + w_2 = \frac{2}{3}$$

$$\mu(\{c_1, c_2, c_3\}) = \sum_{j=1}^3 w_j = w_1 + w_2 + w_3 = 1$$

Então, com essas informações em mãos podemos calcular a integral de Choquet como segue:

$$\gamma^\mu(P_1, P_2, P_3) = (P_{(1)} - P_{(0)})\mu(\{c_1, c_2, c_3\}) + (P_{(2)} - P_{(1)})\mu(\{c_2, c_3\}) + (P_{(3)} - P_{(2)})\mu(\{c_3\})$$

$$\gamma^\mu(P_1, P_2, P_3) = (1 - 0) * 1 + (3 - 1) * \frac{2}{3} + (5 - 3) * \frac{1}{3} = 3$$

Logo, pelos cálculos realizados, verificamos que $MA = \gamma^\mu(P_1, P_2, P_3) = 3$

5.3 Integral de Choquet e o OWA

Para exemplificarmos essa aplicação, vamos considerar os pesos seguintes, $w_1 = 0.2$, $w_2 = 0.3$, $w_3 = 0.5$, $P_{(1)} = 1$, $P_{(2)} = 3$ e $P_{(3)} = 5$. Então o cálculo do OWA é:

$$OWA = (0.2 * 1) + (0.3 * 3) + (0.5 * 5) = 3.6$$

Agora procedendo com o cálculo por integral de Choquet, encontramos os valores para as capacidades como segue:

$$\mu(\{c_1\}) = \mu(\{c_2\}) = \mu(\{c_2\}) = \mu(\{c_3\}) = \sum_{j=0}^0 w_{3-j} = w_3 = 0.5;$$

$$\mu(\{c_1, c_2\}) = \mu(\{c_1, c_3\}) = \mu(\{c_2, c_3\}) = \sum_{j=0}^1 w_{3-j} = w_3 + w_2 = 0.8;$$

$$\mu(\{c_1, c_2, c_3\}) = \sum_{j=0}^2 w_{3-j} = w_3 + w_2 + w_1 = 1$$

Aplicando a integral de Choquet conforme a equação já mostrada aqui em seções anteriores, temos:

$$\gamma^\mu(P_1, P_2, P_3) = (P_{(1)} - P_{(0)})\mu(\{c_1, c_2, c_3\}) + (P_{(2)} - P_{(3)})\mu(\{c_2, c_3\}) + (P_{(2)} - P_{(3)})\mu(\{c_3\})$$

$$\gamma^\mu(P_1, P_2, P_3) = (1 - 0) * 1 + (3 - 1) * 0.8 + (5 - 3) * 0.5 = 3.6$$

Pelos cálculos para generalização do OWA pela integral de Choquet, temos, nesta aplicação, $OWA = \gamma^\mu(P_1, P_2, P_3) = 3.6$

5.4 Integral de Choquet para ordenação das médias escolares dos alunos

Neste exemplo, vamos realizar uma aplicação para ordenação de alunos conforme as notas atribuídas individualmente. Aplicando agora a formulação da integral de Choquet. Nesse caso específico, vamos utilizar uma modelagem que considera as interações entre os critérios. Definindo as capacidades de Choquet, que terão o mesmo papel dos pesos no agregador linear. Como temos 3 critérios, teremos que definir 8 ($= 2^3$) capacidades de Choquet. A coluna MP é a média ponderada.

Aluno	Física (60 horas)	Cálculo (60 horas)	Ética (30 horas)	MP
A	10	9.5	3.5	0.85
B	5	6	4.5	0.53
C	7	7	8	0.72

Tabela 1 – Notas de alunos e média aritmética ponderada MP

- $\mu(\emptyset) = 0$;
- $\mu(\{Física\}) = \frac{60}{150} = 0.4$;
- $\mu(\{Cálculo\}) = \frac{60}{150} = 0.4$;
- $\mu(\{Ética\}) = \frac{30}{150} = 0.2$.

Nessas coalizões, devemos verificar quais critérios há sinergia ou redundância entre eles.

- $\mu(\{Física, Cálculo\}) = 0.5 < 0.4 + 0.4$, pois há redundância entre esses dois critérios;
- $\mu(\{Física, Ética\}) = 0.9 > 0.4 + 0.2$, pois há sinergia entre esses dois critérios;

- $\mu(\{Cálculo, Ética\}) = 0.9 > 0.4 + 0.2$, pois também há sinergia entre esses dois critérios;
- $\mu(\{Física, Cálculo, Ética\}) = 1$, por definição.

Para o aluno $A(i = 1)$, temos:

$$\gamma_1^\mu(P_{11}, P_{12}, P_{13}) = (P_{1(1)} - P_{1(0)})\mu(D(1)) + (P_{1(2)} - P_{1(1)})\mu(D(2)) + (P_{1(3)} - P_{1(2)})\mu(D(3))$$

Ordenando as notas, temos: $P_{1(0)} = 0$, $P_{1(1)} = 3.5$ (Ética), $P_{1(2)} = 9.5$ (Cálculo) e $P_{1(3)} = 10$ (Física). Vamos agora verificar quais capacidades de Choquet utilizaremos nesse caso:

$$\mu(D_{(1)}) = \mu(\{c_{(1)}, c_{(2)}, c_{(3)}\}) = \mu(\{Ética, Cálculo, Física\}) = 1;$$

$$\mu(D_{(2)}) = \mu(\{c_{(2)}, c_{(3)}\}) = \mu(\{Cálculo, Física\}) = 0.5;$$

$$\mu(D_{(3)}) = \mu(\{c_{(3)}\}) = \mu(\{Física\}) = 0.4$$

Então, temos para o aluno A :

$$\gamma_1^\mu(P_{11}, P_{12}, P_{13}) = \frac{(3.5 - 0)}{10} * 1 + \frac{(9.5 - 3.5)}{10} * 0.5 + \frac{(10 - 9.5)}{10} * 0.4 = 0.67$$

A Figura 1 mostra a interpretação geométrica da agregação linear e da agregação feita utilizando-se a integral de Choquet, para o aluno A . Nela, aparecem os valores dos critérios para cada caso, assim como os pesos para o agregador linear. Na interpretação geométrica da integral de Choquet, os critérios já foram previamente ordenados.

Para o aluno $B(i = 2)$, temos:

$$\gamma_2^\mu(P_{21}, P_{22}, P_{23}) = (P_{2(1)} - P_{2(0)})\mu(D(1)) + (P_{2(2)} - P_{2(1)})\mu(D(2)) + (P_{2(3)} - P_{2(2)})\mu(D(3))$$

Ordenando as notas, temos: $P_{2(0)} = 0$, $P_{2(1)} = 4.5$ (Ética), $P_{2(2)} = 5$ (Física) e $P_{2(3)} = 6$. (Cálculo). Vamos agora verificar quais capacidades de Choquet utilizaremos nesse caso:

$$\mu(D_{(1)}) = \mu(\{c_{(1)}, c_{(2)}, c_{(3)}\}) = \mu(\{Ética, Física, Cálculo\}) = 1;$$

$$\mu(D_{(2)}) = \mu(\{c_{(2)}, c_{(3)}\}) = \mu(\{Física, Cálculo\}) = 0.5;$$

$$\mu(D_{(3)}) = \mu(\{c_{(3)}\}) = \mu(\{Cálculo\}) = 0.4.$$

Então, temos para o aluno B :

$$\gamma_2^{\mu^2}(P_{21}, P_{22}, P_{23}) = \frac{(4.5 - 0)}{10} * 1 + \frac{(5 - 4.5)}{10} * 0.5 + \frac{(6 - 5)}{10} * 0.4 = 0.515.$$

Para o aluno $C(i = 3)$, temos:

$$\gamma_3^\mu(P_{31}, P_{32}, P_{33}) = (P_{3(1)} - P_{3(0)})\mu(D(1)) + (P_{3(2)} - P_{3(1)})\mu(D(2)) + (P_{3(3)} - P_{3(2)})\mu(D(3))$$

Ordenando as notas, temos: $P_{3(0)} = 0$, $P_{3(1)} = 7$ (Física), $P_{3(2)} = 7$ (Cálculo) e $P_{3(3)} = 8$ (Ética). Vamos agora verificar quais capacidades de Choquet utilizaremos nesse caso:

$$\mu(D_{(1)}) = \mu(\{c_{(1)}, c_{(2)}, c_{(3)}\}) = \mu(\{\text{Física}, \text{Cálculo}, \text{Ética}\}) = 1;$$

$$\mu(D_{(2)}) = \mu(\{c_{(2)}, c_{(3)}\}) = \mu(\{\text{Cálculo}, \text{Ética}\}) = 0.9;$$

$$\mu(D_{(3)}) = \mu(\{c_{(3)}\}) = \mu(\{\text{Ética}\}) = 0.2.$$

Então, temos para o aluno C:

$$\gamma_3^\mu(P_{31}, P_{32}, P_{33}) = \frac{(7-0)}{10} * 1 + \frac{(7-7)}{10} * 0.9 + \frac{(8-7)}{10} * 0.2 = 0.72.$$

Podemos verificar a forma matricial da integral de Choquet dada pela matriz M

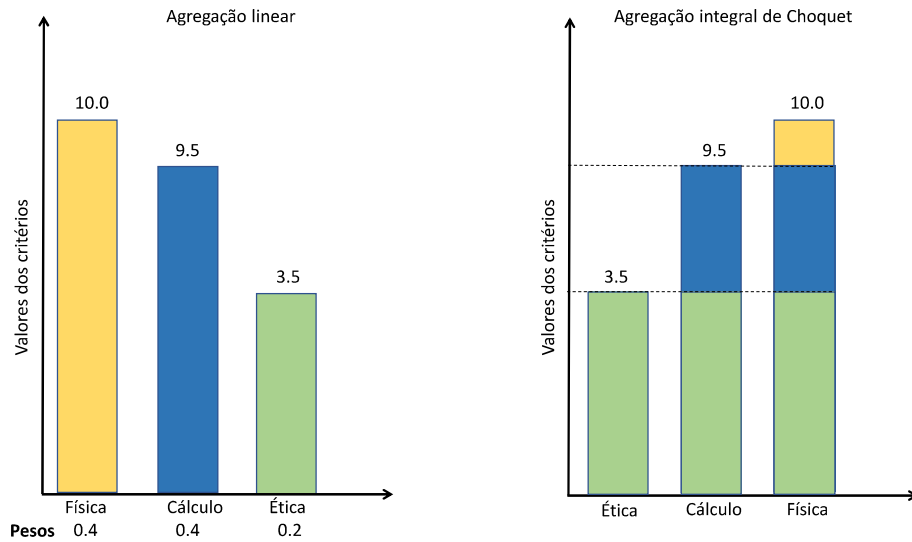


Figura 1 – Representação gráfica para agregação linear e de Choquet

para esse exemplo e a Figura 1 representa a agregação linear e agregação por integral de Choquet.

$$M = \begin{bmatrix} 0 & 0 & 0 \\ 10 - 9.5 & 0 & 0 \\ 0 & 6 - 5 & 0 \\ 0 & 0 & 8 - 7 \\ 9.5 - 3.5 & 5 - 3.5 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 7 - 7 \\ 3.5 - 0 & 3.5 - 0 & 7 - 0 \end{bmatrix},$$

e o vetor de capacidades é dado por:

$$\mu = \begin{bmatrix} 0 \\ 0.4 \\ 0.4 \\ 0.2 \\ 0.5 \\ 0.9 \\ 0.9 \\ 1 \end{bmatrix},$$

de tal modo que:

$$\gamma = M^t \mu = \begin{bmatrix} 0.67 \\ 0.515 \\ 0.72 \end{bmatrix}.$$

A Tabela 2 mostra novamente os três alunos A, B e C, com suas notas nas três disciplinas: Física, Cálculo e Ética. Nela aparece os valores calculados utilizando a integral de Choquet e, também, a média ponderada, a fim de se comparar os resultados.

Aluno	Física (60 horas)	Cálculo (60 horas)	Ética (30 horas)	MP	γ
A	10	9.5	3.5	0.85	0.67
B	5	6	4.5	0.53	0.515
C	7	7	8	0.72	0.72

Tabela 2 – Resultado com a média ponderada MP e o *score* λ

Vemos que, utilizando a integral de Choquet, a ordem ficou Aluno C \succ Aluno A \succ Aluno B, ou seja, o Aluno C tem preferência sobre o Aluno A, que tem preferência sobre o Aluno B. Portanto, utilizando a formulação da integral de Choquet, que permite modelar as interações que existem entre critérios, vemos que a ordem foi alterada, privilegiando aquele aluno que manteve uma regularidade de notas nas três disciplinas.

Um problema que encontramos nesse cenário é o de identificação das capacidades de Choquet. Pelos Exemplos podemos verificar que para um operador de agregação como a média ponderada temos m pesos, enquanto para a integral de Choquet temos $2m$ coeficientes fuzzy. Sabemos, pelas propriedades da medida fuzzy, que $\mu(\emptyset) = 0$ e $\mu(C) = 1$ (GRABISCH, 2016; ANZILLI; GIOVE, 2020), portanto num problema de identificação temos $(2^m - 2)$ coeficientes para determinar, levando-se em consideração a monotonicidade existente entre esses coeficientes. Na próxima seção, será mostrada uma maneira uma outra aplicação em que a abordagem do problema será dada através de identificação de capacidades por programação quadrática e o seu respectivo procedimento metodológico.

6 Abordagem Supervisionada para identificação de medidas de Capacidade de Choquet

Programação quadrática são particularidade de problemas de otimização numérica que possui variadas aplicações, como ajuste de curvas em estatística, aprendizado de máquina para máquinas de vetores de suporte (SVMs) entre outros (ZANGHIRATI; ZANNI, 2003).

Vamos considerar uma abordagem supervisionada cujo objetivo é minimizar o erro quadrático médio entre as avaliações obtidas pela estimação na função $\Gamma(P) = \gamma^\mu(P_{i1}, P_{i2}, \dots, P_{im}) = M\mu$ e os valores reais de saída em u_i .

Pretendemos minimizar o erro cometido ao realizar a estimativa para os valores de u_i , por $\min_{\mu} \sum (u_i - \Gamma(P))^2$. As restrições são da monotonicidade e a condição de fronteira.

A função $\sum (u_i - \Gamma(P))^2$ pode ser expressa por,

$$\begin{aligned} \|M\mu - u\|_2^2 &= (M\mu - u)^T(M\mu - u) \\ &= \mu^T M^T M \mu - \mu^T M^T u - u^T M \mu + u^T u \\ &\propto (1/2)\mu^T M^T M \mu - (1/2)\mu^T M^T u - (1/2)u^T M \mu \\ &= (1/2)\mu^T (M^T M) \mu + (-M^T u)^T \mu. \end{aligned} \quad (12)$$

Dessa forma, fazendo $Q = M^T M$ e $q = -M^T u$, o problema pode ser escrito na forma padrão de um problema de programação quadrática:

$$\begin{aligned} \underset{x}{\text{minimizar}} \quad & (1/2)\mu^T Q \mu + q^T \mu \\ \text{sujeito a:} \quad & Gx \leq h \\ & 0 \leq \mu \leq 1 \\ & \mu(\emptyset) = 0 \quad \mu(D) = 1. \end{aligned} \quad (13)$$

Neste caso, a matriz de restrições Q é simétrica e semi-definida positiva, a matriz G de dimensão $m(2^{m-1} - 1) + 4$ 2^m vetor h devem ser adequadamente construídos para representar as restrições do problema quadrático (ADEYEBA; ANDERSON; HAVENS, 2015).

7 Aplicações da Abordagem Supervisionada da Integral de Choquet Discreta

Nesta seção vamos realizar algumas aplicações com a abordagem supervisionada. A primeira é a determinação das medidas de capacidades de Choquet a partir de sinais coletados de pacientes que apresentaram sintomas de COVID-19. A segunda aplicação será a utilização como preditor de COVID-19 em pacientes com entradas binárias de sinais e sintomas gripais. Vale ressaltar que os experimentos ocorrem como uma simulação computacional.

7.1 Identificação Medidas de Capacidade como Pesos de Sinais em Pacientes

Assim, a análise feita nessa seção terá abordagem semelhante ao desenvolvido pelo exemplo da subseção 5.4. Os pesos ou medida de capacidade de Choquet foram extraídos pelo código-fonte disponível em (NETO, 2022).

Os sinais são dados fornecidos por exames médicos e são geralmente dados de forma contínua. Para fins de experimentação computacional, abordamos cinco entradas que é CORIZA, FEBRE, DOR DE CABEÇA, TOSSE, DISPNEIA e como variável alvo COVID-19.

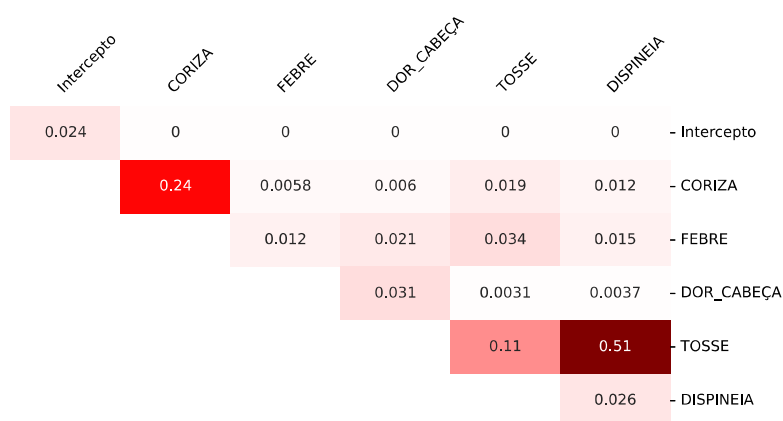


Figura 2 – Pesos da medida de Capacidade de Choquet

Agora, observando as interações entre as variáveis de sintomas, ocorre a redundância, pois, no momento em que se acrescenta sintomas, cresce a possibilidade resultado negativo para o *score* final.

Os *scores* com tonalidade em vermelho são os que apresentam maior importância dentro do modelo de diagnóstico. A diagonal fornece as informações para as medidas individuais e os outros valores indicam as interações par-a-par entre os sintomas. A variável intercepto representa o viés do modelo, em que mostra o enviesamento do conjunto de dados.

7.2 Predição de COVID-19

Vamos, nesta aplicação, e, de forma didática e com adaptações, realizar a predição de COVID-19 com base em dados históricos de pacientes que deram entradas em hospitais e catalogadas por equipes médicas.

A base possui 4 pacientes e 5 sintomas e os valores estão dispostos no formato binário, sendo 1 para quem apresenta os sintomas e 0 para o caso contrário.

Após realizar o processo de encontrar os pesos através do método de programação quadrática, podemos expressar os pesos pela Figura 2, cuja diagonal representa as medidas individuais, no nosso caso, vamos considerar pesos e os outros elementos da matriz superior são as interações par-a-par entre os sintomas. Para

	CORIZA	FEBRE	DOR_CABEÇA	TOSSE	DISPINEIA	COVID19
P_1	1	1	0	1	1	1
P_2	0	1	1	1	0	1
P_3	1	0	0	1	0	0
P_4	1	0	0	0	0	0
P_5	0	0	1	1	1	1
P_6	1	0	1	1	0	1

Tabela 3 – Amostras de Sintomas de Pacientes com COVID-19

o problema de predição podemos considerar apenas os pesos individuais e realizar o produto entre a matriz de sintomas e o vetor de pesos. A saída serão valores no intervalo de $[0, 1]$, pois estão relacionados com as restrições impostas da integral de Choquet.

Vamos verificar qual a probabilidade de um novo paciente está com COVID-19, com base nos pesos extraídos da matriz de pesos.

	CORIZA	FEBRE	DOR_CABEÇA	TOSSE	DISPINEIA
P	0	1	1	0	0

Pelo produto ponderado, verificamos o valor de 0.124, e pela coalizão entre os sintomas *Febre* e *DOR_CABEÇA*, é acrescentado em 0.394, então as chances de estar com a doença é de 0.52.

8 Conclusão

Pela seção 5.4 os *scores* atribuídos pela integral de Choquet a cada um dos alunos pode ser uma alternativa para o cálculo da média, pois assim, quando se acrescenta critérios de notas, os pesos ou medidas de capacidade de Choquet sinalizam um possível desfecho, ou conteúdos em disciplinas onde o aluno requer reforço para um resultado positivo. Isso pode ser interessante quando se deseja um aluno com habilidades diferentes e não o todo da matéria.

Nos casos preditivos com o uso de integral de Choquet, as medidas de capacidade de Choquet podem ser uma forma de visualizar as interações entre os critérios, o que pode aumentar ou diminuir as possibilidades de predição das classes em casos de aplicação de uma regressão linear.

Para os pesos individuais ou as interações dos pesos, podem ocorrer muito próximos de zeros, isso acontece devido aos valores que podem estar fora de escalas. Outra alternativa, e fica como perspectiva, seria criar uma matriz esparsa através de regularização l_1 . Assim, tornando o problema de otimização relacionado com abordagem conhecida como LASSO (Least Absolute Shrinkage and Selection Operator) tal qual o método de variável não-negativa (TIBSHIRANIT, 1996).

Além da regularização LASSO, podemos expressar uma capacidade ou os pesos das interações de um modelo em termos do índice de interação de Banzhaf como descrito em (PELEGRINA et al., 2020) e também por potência de índices de interação correspondentes ao valor de Shapley conforme descrito nos trabalhos (GRABISCH, 1996; GRABISCH; ROUBENS, 2000; SHAPLEY, 1953).

Referências

ADEYEBA, T. A.; ANDERSON, D. T.; HAVENS, T. C. Insights and characterization of l1-norm based sparsity learning of a lexicographically encoded capacity vector for the choquet integral. In: *2015 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*. [S.l.: s.n.], 2015. p. 1–7. Citado na página 13.

ANZILLI, L.; GIOVE, S. Multi-criteria and medical diagnosis for application to health insurance systems: a general approach through non-additive measures. *Decisions in Economics and Finance*, Springer International Publishing, v. 43, p. 559–582, 2020. ISSN 11296569. Disponível em: <<https://doi.org/10.1007/s10203-020-00302-x>>. Citado na página 12.

BUENO, J. C. S.; DIMURO, G. P.; BORGES, E. N. *REDUÇÃO DIMENSIONAL DE IMAGENS: APLICAÇÃO DE FUNÇÕES DE (PRÉ) AGREGAÇÃO BASEADAS NA INTEGRAL DE CHOQUET*. 2018. Citado na página 3.

BüYÜKÖZKAN, G.; GÖÇER, F. Smart medical device selection based on intuitionistic fuzzy choquet integral. *Soft Computing*, Springer Berlin Heidelberg, v. 23, p. 10085–10103, 2019. ISSN 14337479. Disponível em: <<https://doi.org/10.1007/s00500-018-3563-5>>. Citado na página 4.

CHOQUET, G. Theory of capacities. In: *Annales de l'institut Fourier*. [S.l.: s.n.], 1954. v. 5, p. 131–295. Citado na página 2.

DUARTE, L. T.; TOREZZAN, C. *Introduction to the Choquet integral and its application to MCDA Introduction*. [S.l.]: Unicamp, 2020. 51-61 p. Citado na página 4.

GRABISCH, M. The application of fuzzy integrals in multicriteria decision making. *European Journal of Operational Research*, v. 89, p. 445–456, 3 1996. ISSN 03772217. Citado na página 16.

GRABISCH, M. Bases and transforms of set functions. *Bases and Transforms of Set Functions. S. Saminger-Platz and R. Mesiar. On Logical, Algebraic and Probabilistic Aspects of Fuzzy Set Theory*, p. 10, 2016. Disponível em: <<https://hal.archives-ouvertes.fr/hal-01302376>>. Citado 3 vezes nas páginas 3, 5 e 12.

GRABISCH, M.; LABREUCHE, C. Bi-capacities-part ii: the choquet integral. 2007. Disponível em: <<https://hal.archives-ouvertes.fr/hal-00187191>>. Citado na página 7.

GRABISCH, M.; LABREUCHE, C. A decade of application of the choquet and sugeno integrals in multi-criteria decision aid. *Annals of Operations Research*, Springer, v. 175, n. 1, p. 247–286, 2010. Citado na página 2.

GRABISCH, M.; ROUBENS, M. Application of the choquet integral in multicriteria decision making. *Fuzzy Measures and Integrals*, p. 348 – 374, 2000. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/summary;jsessionid=7B0CA6B5703F222B23BDA69D589EA2DE?doi=10.1.1.105.3433>>. Citado 4 vezes nas páginas 2, 4, 6 e 16.

GRABISCH, M.; ROUBENS, M. *Application of the Choquet integral in multicriteria decision making Multi-attribute decision making by consumers associations View project Application of the Choquet Integral in Multicriteria Decision Making*. 2008. Disponível em: <<https://www.researchgate.net/publication/228597117>>. Citado 2 vezes nas páginas 2 e 3.

HUANG, K. K. et al. Applying a generalized choquet integral with signed fuzzy measure based on the complexity to evaluate the overall satisfaction of the patients. *2010 International Conference on Machine Learning and Cybernetics, ICMLC 2010*, IEEE, v. 5, p. 2377–2382, 2010. Citado na página 4.

LAIATE, B.; BARROS, L. C. D.; ESMI, E. Choquet calculus and applications to population dynamics view project fuzzy differential equations under fréchet derivative view project. 2016. Disponível em: <<https://www.researchgate.net/publication/326315213>>. Citado na página 3.

MAYAG, B.; BOUYSSOU, D. Necessary and possible interaction between criteria in a 2-additive choquet integral model. *European Journal of Operational Research*, Elsevier B.V., v. 283, p. 308–320, 2020. ISSN 03772217. Disponível em: <<https://doi.org/10.1016/j.ejor.2019.10.036>>. Citado na página 4.

NETO, L. A. S. *Código do Método de integral de Choquet discreta*. 2022. Disponível em: <<https://colab.research.google.com/drive/1nYVq6zzDI6XFkHRWJl8qbUZ1eBXsCfNe>>. Citado na página 14.

OLIVEIRA, H. E. de; DUARTE, L. T.; ROMANO, J. M. T. Identification of the choquet integral parameters in the interaction index domain by means of sparse modeling. *Expert Systems with Applications*, Elsevier Ltd, v. 187, 1 2020. ISSN 09574174. Citado 3 vezes nas páginas 3, 5 e 6.

PELEGRINA, G. D. et al. The multilinear model in multicriteria decision making: The case of 2-additive capacities and contributions to parameter identification. *European Journal of Operational Research*, v. 282, p. 945–956, 2020. ISSN 03772217. Citado 2 vezes nas páginas 3 e 16.

PEREIRA, M. A.; FIGUEIRA, J. R.; MARQUES, R. C. Using a choquet integral-based approach for incorporating decision-maker's preference judgments in a data envelopment analysis model. *European Journal of Operational Research*, Elsevier B.V., v. 284, p. 1016–1030, 2020. ISSN 03772217. Citado na página 4.

SHAPLEY, L. S. *A Value for n-Person Games*. [S.l.]: Princeton University Press, 1953. Volume II. 307-318 p. Citado na página 16.

TIBSHIRANIT, R. Regression Shrinkage and Selection via the Lasso. *J. R. Statist. Soc. B*, v. 58, p. 267–288, 1996. Citado 2 vezes nas páginas 3 e 15.

WANG, W.; LIU, X.; QIN, Y. A fuzzy fine-kinney-based risk evaluation approach with extended multimooora method based on choquet integral. *Computers and Industrial Engineering*, Elsevier, v. 125, p. 111–123, 2018. ISSN 03608352. Disponível em: <<https://doi.org/10.1016/j.cie.2018.08.019>>. Citado na página 4.

YAGER, R.; FILEV, D. Induced ordered weighted averaging operators. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, v. 29, n. 2, p. 141–150, 1999. Citado na página 7.

ZANGHIRATI, G.; ZANNI, L. A parallel solver for large quadratic programs in training support vector machines. *Parallel computing*, Elsevier, v. 29, n. 4, p. 535–551, 2003. Citado na página 13.