

Aprendizagem de preferências através do método TOPSIS: uma aplicação para ratings de liberdade econômica

Diogo F. de Lima Silva¹, Julio Cezar S. Silva², Adiel T. de Almeida Filho²

¹Departamento de Engenharia de Produção – Universidade Federal de Pernambuco (UFPE)
Caixa Postal 7457 – 50.630-971 – Recife – PE – Brasil

²Centro de Informática – Universidade Federal de Pernambuco (UFPE)
Av. Jornalista Anibal Fernandes, s/n, Cidade Universitária (Campus Recife), CEP: 50.740-
560 – Recife – PE – Brasil

diogo@greefo.org.br, jcass4@cin.ufpe.br, adielfilho@cin.ufpe.br

Abstract. *This article presents the application of a multicriteria classification method, which uses preference learning, to allocate a set of 180 countries in 5 classes of economic freedom. The classes were pre-defined and ordered, which characterizes a sorting problem. In the application, choices of reference alternatives were simulated. Then, the results were compared to the classification carried out by a specialized agency. The method obtained consistent results, with a similarity percentage higher than 85%.*

Resumo. *Esse artigo apresenta a aplicação de um método multicritério de classificação, que utiliza aprendizagem de preferências, para alocar um conjunto de 180 países em 5 classes de liberdade econômica. As classes são pré-definidas e ordenadas, o que caracteriza o problema como de “sorting”. Para a aplicação, escolhas de alternativas de referência foram simuladas. Em seguida, os resultados foram comparados à classificação realizada por uma agência especializada. O método obteve resultados consistentes, com um percentual de semelhança superior a 85%.*

1. Introdução

Aprendizagem de preferências tem se tornado um chave em apoio a decisão multicritério (Greco & Kadziński, 2018). Segundo Mousseau & Slowinski (1998), em vários casos não é realista admitir que decisores sejam capazes de definir uma grande quantidade de parâmetros de maneira exata. Assim, o uso de técnicas de desagregação de preferências (Jacquet-Lagrèze & Siskos, 2001), incorporando características de programação matemática e inteligência artificial, para fazer sistemas aprendam a partir de exemplos de decisão, está presente em muitas variações de métodos multicritério (Greco et al., 2016).

Na problemática de classificação (Roy, 1996), um conjunto de alternativas deve ser alocado a classes pré-definidas. Pode-se ainda distinguir entre classificação ordinal (sorting), classificação nominal e clusterização. Enquanto em problemas sorting as classes são pré-

definidas e ordenadas em termos de preferências, problemas de classificação lida com classes pré-definidas porém não ordenadas, e em problemas de clusterização as classes não são pré-definidas nem ordenadas (Ouenniche et al., 2018). Nos métodos de desagregação de preferências usados para sorting, os exemplos de decisão são alocações de alternativas (fictícias ou não), chamadas de referências, para as classes pré-determinadas. Alguns exemplos de métodos são o UTADIS (Pardalos et al., 1995), o DRSA (Greco et al., 2001), a adaptação do ELECTRE-TRI (Mousseau & Slowinski, 1998), e o recentemente proposto PDTOPSIS-Sort (de Lima Silva et al., 2020).

A problemática de classificação é particularmente interessante em problemas financeiros (Douplos & Zopounidis, 2014; Zopounidis et al., 2015) envolvendo temas tais como risco de crédito, risco país, fusões e aquisições. Um problema de sorting ainda pouco explorado na literatura de decisão multicritério em finanças diz respeito à classificação de países em níveis de liberdade econômica. Esse problema de decisão é estudado pela fundação Heritage, que anualmente gera relatórios com métricas para um conjunto de países (Heritage Foundation, 2020). Esse problema foi recentemente tratado por De Lima Silva & de Almeida Filho (2020), que aplicaram os métodos baseados em TOPSIS para alocar países em classes de liberdade econômica. Por outro lado, a aplicação não utilizou nenhuma técnica de desagregação de preferências e os parâmetros foram totalmente baseados nos parâmetros utilizados pela fundação Heritage.

Nesse artigo, o método PDTOPSIS-Sort é aplicado para explorar a técnica de aprendizado de preferências no problema de avaliação de liberdade econômica. Exemplos de decisão foram simulados a partir de conjuntos de alternativas reais e os resultados da alocação foram comparados às alocações originais da Heritage no ano de 2019.

2. PDTOPSIS-Sort

O método PDTOPSIS-Sort foi proposto por (de Lima Silva et al., 2020) para problemas de sorting com base em desagregação de preferências e na estrutura do popular TOPSIS (Behzadian et al., 2012), vastamente aplicado na literatura.

Seja $A = \{a_1, a_2, \dots, a_m\}$ um conjunto de m alternativas; $G = \{g_1, g_2, \dots, g_n\}$ um conjunto de n critérios; $w = \{w_1, w_2, \dots, w_n\}$ um vetor de pesos para os critérios; $C = \{C_1, C_2, \dots, C_q\}$ um conjunto de q classes pré-definidas e ordenadas, com $C_1 > C_2 > \dots > C_q$; e $P = \{p_1, p_2, \dots, p_q\}$ um conjunto de $p = q - 1$ perfis de fronteira. Sejam $G^+ \subset G$ e $G^- \subset G$ respectivamente subconjuntos de critérios de benefício e custo, domínio dos critérios representado por $\{a^*, a^-\}$, onde a_j^* e a_j^- são respectivamente o maior e menor valor relacionado ao critério g_j . Sejam $a_{i,j}$ e $P_{k,j}$ respectivamente a avaliação da alternativa a_i e do perfil P_k considerando g_j . Além disso, seja $R = \{ref_1, ref_2, \dots, ref_r\}$ um conjunto de r alternativas de referência e r_k o número de referências alocadas na classe C_k . O Algoritmo 1 detalha a rotina do método PDTOPSIS-Sort.

Algoritmo 1: Rotina do PDTOPSIS-Sort

Etapa 1: Determinar a Matriz de Decisão $X = [a_{i,j}]_{m \times n}$

Etapa 2: Definir alternativas de referência $R = [ref_{h,j}]_{r \times n}$

Etapa 3: Determinar o domínio de cada critério, representado por $D = \begin{bmatrix} a_j^* \\ a_j^- \end{bmatrix}_{2 \times n}$

Etapa 4: Inferir perfis $P = [P_{k,j}]_{q-1 \times n}$ e pesos $w = [w_1, \dots, w_n]$, usados no processo de alocação. A inferência utiliza um procedimento de desagregação de preferências com minimização de variáveis de erro: $\sigma_{i,k}^+$ and $\sigma_{i,k}^-$.

Etapa 5: Validar os parâmetros inferidos na Etapa 4, antes de classificar as alternativas.

Etapa 6: Processo de alocação.

Etapa 6.1: Determinar a matriz de decisão $M = [M_{i,j}]_{(m+q+1) \times n} = \begin{bmatrix} X \\ P \\ D \end{bmatrix}$, obtida concatenando $X = [a_{i,j}]_{m \times n}$, $P = [P_{k,j}]_{q-1 \times n}$ and $D = \begin{bmatrix} a_j^* \\ a_j^- \end{bmatrix}_{2 \times n}$

Etapa 6.2: Calcular a matriz de decisão normalizada $R = [r_{i,j}]_{(m+q+1) \times n}$

$$r_{ij} = \frac{M_{ij}}{\max_{1 \leq i \leq m+q+1} M_{ij}}$$

Etapa 6.3: Calcular a matriz de decisão normalizada pelos pesos $V = [v_{i,j}]_{(m+q+1) \times n}$

$$v_{i,j} = w_j r_{i,j}, i = 1, 2, \dots, (m + q + 1); j = 1, 2, \dots, n$$

Etapa 6.4: Determinar a solução ideal e a anti-ideal.

$$v^* = [v_1^*, v_2^*, \dots, v_n^*], \quad v_j^* = \begin{cases} \max_i v_{i,j}, & g_j \in G^+ \\ \min_i v_{i,j}, & g_j \in G^- \end{cases}$$
$$v^- = [v_1^-, v_2^-, \dots, v_n^-], \quad v_j^- = \begin{cases} \min_i v_{i,j}, & g_j \in G^+ \\ \max_i v_{i,j}, & g_j \in G^- \end{cases}$$

Etapa 6.5: Calcular distâncias Euclidianas para as soluções ideal e anti-ideal.

$$d_{a_i}^* = \sqrt{\sum_{j=1}^n (v_{i,j} - v_j^*)^2}, \quad i = 1, 2, \dots, m.$$

$$d_{a_i}^- = \sqrt{\sum_{j=1}^n (v_{i,j} - v_j^-)^2}, \quad i = 1, 2, \dots, m.$$

$$d_{P_k}^* = \sqrt{\sum_{j=1}^n (v_{i,j} - v_j^*)^2}, \quad k = 1, 2, \dots, q-1; i = k + m.$$

$$d_{P_k}^- = \sqrt{\sum_{j=1}^n (v_{i,j} - v_j^-)^2}, \quad k = 1, 2, \dots, q-1; i = k + m.$$

Etapa 6.6: Calcular o coeficiente de proximidade das alternativas e dos perfis

$$Cl(a_i) = \frac{d_{a_i}^-}{d_{a_i}^* + d_{a_i}^-}, \quad i = 1, 2, \dots, m$$

$$Cl(P_k) = \frac{d_{P_k}^-}{d_{P_k}^* + d_{P_k}^-}, \quad k = 1, 2, \dots, q-1$$

Etapa 6.7: Classificar as alternativas comparando seus coeficientes de proximidade $Cl(a_i)$ com os obtidos para os perfis $Cl(P_k)$.

$$a_i \in C_1 \text{ iff } Cl(a_i) \geq Cl(P_1)$$

$$a_i \in C_k \text{ iff } Cl(P_{k-1}) > Cl(a_i) \geq Cl(P_k), \quad k = 2, \dots, (q-1)$$

$$a_i \in C_q \text{ iff } Cl(a_i) < Cl(P_{q-1})$$

Etapa 7: Realizar uma análise de sensibilidade para os resultados da classificação.

Nesse método, a classificação das alternativas é realizada com base em suas distâncias relativas a uma alternativa ideal e uma alternativa anti-ideal. Um coeficiente de proximidade é então calculado para cada alternativa, e comparado aos coeficientes de proximidades de perfis de fronteira distribuídos entre classes consecutivas. Dessa forma, o perfil P_k indica ao mesmo tempo o limite inferior para a classe P_k e superior para a classe P_{k+1} .

Na etapa 4 do método, acontece a aprendizagem de preferências. Uma inferência baseada em conjunto de alternativas de referências (usadas como exemplos de decisão) determinadas na Etapa 2. Nessa etapa, os perfis referenciais para as classes e os pesos dos critérios são inferidos com um modelo de programação matemática. O modelo objetiva reduzir ao máximo variáveis de erro, que podem ser necessárias caso o conjunto de exemplos de decisão, ou os critérios escolhidos, não façam sentido em termos da ordenação das classes. Mais detalhes sobre cada etapa do método e sobre o modelo de otimização podem ser

encontrados em (de Lima Silva et al., 2020). O modelo de otimização aplicado neste trabalho é apresentado na Seção 3.

3. Aplicação

O problema consiste em avaliar um conjunto de $m = 180$ países em relação à liberdade econômica a partir de 12 critérios, descritos nos indicadores utilizados para Fundação Heritage e apresentados na Tabela 1.

Tabela 1. Conjunto de critérios

Indicadores de liberdade econômica	
g_1	Direitos de propriedade
g_2	Eficácia Judicial
g_3	Integridade do governo
g_4	Carga tributária
g_5	Gastos públicos
g_6	Saúde Fiscal
g_7	Liberdade nos Negócios
g_8	Liberdade trabalhista
g_9	Liberdade Monetária
g_{10}	Liberdade comercial
g_{11}	Liberdade de investimento
g_{12}	Liberdade financeira

Para avaliar cada país, a Heritage utiliza uma média simples das avaliações do mesmo nos doze indicadores, que tem valores variando entre 0 e 100. Dessa maneira, a avaliação final das alternativas fica situada no mesmo domínio. Adicionalmente, a fundação define pontos de corte para classificar as alternativas em 5 classes, descritas na Tabela 2.

Tabela 2. Conjunto de classes

Classe		Perfis de corte
C_1	Livre	$80 \leq \text{avaliação global} \leq 100$
C_2	Predominantemente livre	$70 \leq \text{avaliação global} < 80$
C_3	Moderadamente livre	$60 \leq \text{avaliação global} < 70$
C_4	Predominantemente sem liberdade	$50 \leq \text{avaliação global} < 60$
C_5	Reprimido	$0 \leq \text{avaliação global} < 50$

Por outro lado, a definição precisa de tais parâmetros pode deixar margem para indagações. Por exemplo, por que utilizar o valor 80 e não 79,9 ou 80,1? Além disso, por que

os critérios foram identificados como tendo igual importância? Um critério não poderia ter um peso um pouco maior do que o outro? Um decisor real poderia tomar tais decisões de maneira precisa para julgar o país no qual abrirá uma nova sede de sua empresa, qual país investir, ou qualquer outra decisão relacionada?

Modelos que utilizam aprendizagem de preferência são interessantes nesses casos por diminuir consideravelmente o esforço cognitivo necessário para modelar as preferências de um decisor. Neste trabalho, utilizamos como parâmetro inicial para a escolha de alternativas de referência as classificações originais da fundação Heritage. Assim, para cada classe, referências foram escolhidas aleatoriamente respeitando a alocação inicial da fundação. Ao todo, 500 simulações foram realizadas.

O modelo de inferência de parâmetros do método PDTOPSIS-Sort aplicado para este trabalho é apresentado a seguir:

$$Min. ERRO = \frac{1}{r_3} \left(\sum_{i \in C_1} \sigma_{i,1}^+ \right) + \sum_{k=2}^4 \frac{1}{r_3} \left(\sum_{i \in C_k} (\sigma_{i,k}^+ + \sigma_{i,k}^-) \right) + \frac{1}{r_3} \left(\sum_{i \in C_q} \sigma_{i,q}^- \right) \quad (1)$$

Subject to:

$$Cl(ref_i) + \sigma_{i,k}^+ \geq Cl(P_k), \quad \forall i \in C_k, k = 1, 2, 3, 4, \quad (2)$$

$$Cl(ref_i) - \sigma_{i,k}^- < Cl(P_{k-1}), \quad \forall i \in C_k, k = 2, 3, 4, 5 \quad (3)$$

$$P_{1,j} \geq P_{2,j} \geq P_{3,j} \geq P_{4,j}, \quad j = 1, \dots, 12 \quad (4)$$

$$P_{k,j} \geq \mu_{jC_{k,k+1}} - 0,5 \times |\mu_{jC_{k,k+1}}|, \quad \forall k, j \quad (5)$$

$$P_{k,j} \leq \mu_{jC_{k,k+1}} + 0,5 \times |\mu_{jC_{k,k+1}}|, \quad \forall k, j \quad (6)$$

$$P_{k,j} - P_{k+1,j} \geq 5, \quad k = 1, 2, 3 \quad (7)$$

$$w_j \leq \frac{1}{2 \sum_{t=1}^n w_t}, \quad j = 1, 2, \dots, n \quad (8)$$

$$\sum_{j=1}^n w_j = 1 \quad (9)$$

$$w_j \geq 0, \quad j = 1, 2, \dots, n \quad (10)$$

$$\sigma_{ik}^+ \geq 0 \text{ and } \sigma_{ik}^- \geq 0, \quad \forall i, k \quad (11)$$

Como explicado na etapa 4 do Algoritmo 1, os parâmetros inferidos através do modelo de otimização são os perfis de fronteira entre as classes e os pesos dos critérios., sendo esta inferência baseada na minimização de variáveis de erro. Após a inferência dos parâmetros, a classificação é realizada. Os resultados obtidos são apresentados na Seção 4.

4. Resultados e Discussões

A Tabela 3 apresenta resultados das similaridades encontradas entre as classificações obtidas nas simulações e as classificações da fundação Heritage. Além do resultado geral, também são apresentados os resultados para cada classe, incluindo máximo, mínimo, média e desvio padrão das similaridades. Observa-se que os resultados alcançaram um bom desempenho, sendo consistentes com o que era esperado. Com apenas 3 exemplos de referência por classe, obteve-se semelhança média de 86.4%, e uma semelhança máxima de 97.8% em relação à classificação original das alternativas.

Bons resultados também foram obtidos dentro das classes, com um percentual de semelhança médio acima de 90% para as classes C_3 e C_4 . A menor semelhança foi observada nos países originalmente alocados para a classe C_5 que, mesmo assim, superou o percentual de 67% de acertos.

Tabela 3. Semelhança entre as classificações

		Global	C_1	C_2	C_3	C_4	C_5
Heritage Foundation	Média	0.864	0.799	0.84	0.909	0.903	0.676
	Desvio Padrão	0.0592	0.172	0.125	0.0719	0.0927	0.211
	Máximo	0.978	1	1	1	1	1
	Mínimo	0.683	0.333	0.345	0.644	0.484	0.136

A Tabela 4 ilustra o comportamento das variáveis de decisão obtidas nas simulações. Apesar dos percentuais de semelhança relevantes apresentados na Tabela 3, observa-se que não necessariamente os perfis médios estiveram próximos dos utilizados pela fundação Heritage. Além disso, percebe-se uma variabilidade de valores de um mesmo perfil em relação a diferentes critérios. Por exemplo, $\bar{P}_{2,6} - \bar{P}_{2,12} > 20$. Este tipo de diferenciação de um perfil dentro dos critérios não é realizado pela fundação Heritage (que apenas utiliza pontos de corte no valor global), porém, está presente em vários métodos de classificação multicritério.

Em relação aos pesos, observa-se que houve variações, mas a média estiveram próximos ao valor de 1/12, utilizado como solução inicial na modelagem. É importante destacar que a equação (9) limita os pesos superiormente ao valor de 0.5. Porém, este valor não foi atingido na simulação, como descrito na última linha da tabela.

Tabela 4. Perfis médios

	g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8	g_9	g_{10}	g_{11}	g_{12}	
P_1	Min	75.9	65.57	64.37	74.3	53.68	79.4	75.78	65.69	82.17	83.77	70.83	68.33
	Média	84.2	75.29	78.17	84.11	71.98	90.8	83.77	74.92	85.34	88.12	81.43	76.03
	Max	91.2	83.97	90.98	93.53	88.54	97.53	90.95	84.38	90.02	90.95	89.17	85
P_2	Min	56.17	43.67	39.9	69.3	47.4	59.98	63.44	53.33	75.82	72.6	57.5	48.33
	Média	68.21	57.64	55.56	79.11	66.63	82.41	74.27	65.14	80.29	81.79	72.73	62.22
	Max	80.65	69.88	73.65	88.53	83.54	92.53	82.65	77.7	85.02	85.95	83.33	73.33
P_3	Min	36.45	29.58	29.11	64.3	42.13	38.65	51.95	46.8	69.63	63.97	37.5	31.67
	Média	49.77	42.31	36.76	74.11	61.58	66.99	63.51	58.91	75.02	74.2	56.55	46.45
	Max	62.5	55.46	47.8	83.53	78.54	87.29	73.13	72.13	80.02	80.2	71.67	56.67
P_4	Min	22.5	20.37	20.3	52.62	31.78	5.633	34.77	32.48	43.38	48.37	17.5	16.67
	Média	36.04	31.03	28	68.6	55.51	45.54	49.87	50.29	67.41	64.67	38.64	31.72
	Max	45.57	44.33	37.95	78.53	73.54	80.08	61.78	61.43	75.02	74.3	59.58	45
W	Min	0.0032	7e-07	6e-17	0	0	1e-05	1e-16	3e-17	0	1e-17	0	0
	Média	0.088	0.084	0.086	0.066	0.08	0.085	0.081	0.092	0.084	0.082	0.085	0.086
	Max	0.23	0.21	0.29	0.12	0.16	0.18	0.21	0.24	0.3	0.18	0.27	0.18

5. Considerações Finais

A aprendizagem de preferências tem ganhado espaço ao longo dos anos e hoje é vista como uma necessidade em diversas aplicações. A principal vantagem de usar uma técnica pertencente a esse escopo é a diminuição iminente do esforço cognitivo de um usuário ou decisor para a utilização de um método formal de apoio a decisão. Assim, parâmetros de métodos podem ser estimados a partir de exemplos de referência. Além disso, é possível aprender com exemplos de decisões que estruturar melhor sistemas de recomendações através de abordagens que considerem os aspectos de racionalidade e dominância.

Esse artigo apresentou uma aplicação de um método de aprendizado de preferências baseado em TOPSIS para alocação de países em classes de liberdade financeira. Escolhas de exemplos de decisão foram realizadas através de uma simulação. Então, os resultados foram comparados àqueles disponibilizados por uma fundação especializada. Observou-se que o método PDTOPSIS-Sort teve resultados consistentes e coerentes com as expectativas. Foi possível obter uma média de acertos superior a 85%, com utilização de poucas alternativas de referência.

Trabalhos futuros podem explorar outros métodos multicritério que utilizam desagregação de preferências no contexto de liberdade econômica. Além disso, pode-se testar diferentes quantidades de exemplos de referências para testar a sensibilidade dos resultados da aplicação.

Agradecimentos

Esta pesquisa foi parcialmente financiada pela Facepe (IBPG-0753-3.08/17, IBPG-0373-1.03/19), CNPq (311140/2017-3, 425544/2016-8) e CAPES (001).

Referências

- Behzadian, M., Khanmohammadi Otagh Sara, S., Yazdani, M., & Ignatius, J. (2012). A state-of-the-art survey of TOPSIS applications. *Expert Systems with Applications*, *39*(17), 13051–13069.
- de Lima Silva, D. F., & de Almeida Filho, A. T. (2020). Sorting with TOPSIS through boundary and characteristic profiles. *Computers & Industrial Engineering*, *141*, 106328.
- de Lima Silva, D. F., Ferreira, L., & de Almeida-Filho, A. T. (2020). A new preference disaggregation TOPSIS approach applied to sort corporate bonds based on financial statements and expert's assessment. *Expert Systems with Applications*, *IN PRESS*.
- Doumpos, M., & Zopounidis, C. (2014). *Multicriteria Analysis in Finance (1st ed.)*. Springer International Publishing.
- Greco, S., Ehrgott, M., & Figueira, J. R. (Eds.). (2016). *Multiple Criteria Decision Analysis: State of Art Surveys (2nd ed., Vol. 233)*. New York, NY: Springer New York.
- Greco, S., & Kadziński, M. (2018). Feature cluster: Learning perspectives in Multiple Criteria Decision Analysis. *European Journal of Operational Research*, *264*(2), 403–404.
- Greco, S., Matarazzo, B., & Slowinski, R. (2001). Rough sets theory for multicriteria decision analysis. *European Journal of Operational Research*, *129*(1), 1–47.
- Heritage Foundation (2020). Index of Economic Freedom, accessed February 24, 2020, available at <https://www.heritage.org/index/>.
- Jacquet-Lagrèze, E., & Siskos, Y. (2001). Preference disaggregation: 20 Years of MCDA experience. *European Journal of Operational Research*, *130*(2), 233–245.
- Mousseau, V., & Slowinski, R. (1998). Inferring an ELECTRE TRI Model from Assignment Examples. *Journal of Global Optimization*, *12*(2), 157–174.
- Ouenniche, J., Pérez-Gladish, B., & Bouslah, K. (2018). An out-of-sample framework for TOPSIS-based classifiers with application in bankruptcy prediction. *Technological Forecasting and Social Change*, *131*(May 2017), 111–116.
- Pardalos, P. M., Siskos, Y., & Zopounidis, C. (Eds.). (1995). *Advances in Multicriteria Analysis (Vol. 5)*. Boston, MA: Springer US.
- Roy, B. (1996). *Multicriteria methodology for decision aiding*. Kluwer Academic Publishers.
- Zopounidis, C., Galariotis, E., Doumpos, M., Sarri, S., & Andriopoulos, K. (2015). Multiple criteria decision aiding for finance: An updated bibliographic survey. *European Journal of Operational Research*, *247*(2), 339–348.