Error in the Euclidean Preference Model

Extended Abstract

Luke Thorburn
King's College London
London, United Kingdom
luke.thorburn@kcl.ac.uk

Maria Polukarov
King's College London
London, United Kingdom
maria.polukarov@kcl.ac.uk

Carmine Ventre
King's College London
London, United Kingdom
carmine.ventre@kcl.ac.uk

ABSTRACT

Spatial models of preference, in the form of vector embeddings, are learned by many deep learning and multiagent systems, including recommender systems. Often these models are assumed to approximate a Euclidean structure, where an individual prefers alternatives positioned closer to their "ideal point", as measured by the Euclidean metric. However, Bogomolnaia and Laslier [3] showed that there exist ordinal preference profiles that cannot be represented with this structure if the Euclidean space has two fewer dimensions than there are individuals or alternatives. We extend this result, showing that there are realistic situations in which almost all preference profiles cannot be represented with the Euclidean model, and derive a theoretical lower bound on the expected error when using the Euclidean model to approximate non-Euclidean preference profiles. Our results have implications for the interpretation and use of vector embeddings, because in some cases close approximation of arbitrary, true ordinal relationships can be expected only if the dimensionality of the embeddings is a substantial fraction of the number of entities represented.

KEYWORDS

preference representation; spatial preferences; embeddings

ACM Reference Format:
Luke Thorburn, Maria Polukarov, and Carmine Ventre. 2023. Error in the Euclidean Preference Model: Extended Abstract. In *Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), London, United Kingdom, May 29 – June 2, 2023*, IFAAMAS, 3 pages.

1 INTRODUCTION

Spatial models of preference, in the form of vector embeddings, are widely used in deep learning systems. In user-facing contexts such as recommender systems, user embeddings contain information about literal preferences [15], but embeddings used to capture degrees of similarity between elements of a set can also be viewed as preference models. For example, each word in a language could be considered to have a "preference" over all other words, "preferring" those with similar meanings. Spaces of word embeddings [1] can thus be viewed abstractly as models of preference.

A canonical spatial model is the Euclidean model, where both individuals and alternatives are represented as points in Euclidean space, and each individual prefers nearer alternatives, as measured by the standard Euclidean metric [3]. A preference profile of I individuals over A alternatives is said to be d-Euclidean if it can be represented with a d-dimensional Euclidean model. The Euclidean model is often used explicitly, and even spatial preference models that are not strictly Euclidean are often assumed to have an approximately Euclidean structure, such as when embeddings are compared using cosine similarity [12, Ch. 6.4], which induces the same ordinal relationships as the Euclidean metric when applied to normalized vectors. Given the prevalence of Euclidean preference models, it is important to understand their limitations.

In this paper, we assume a "ground truth" preference structure where each individual's preference is a strict order over the available alternatives. Consider the expressiveness of the Euclidean preference model relative to this ordinal model. There are three questions one might ask. For fixed positive integers I, A and d:

1. Are there any preference profiles of I individuals over A alternatives that are not d-Euclidean?
2. What proportion of such profiles are not d-Euclidean?
3. How large is the expected error when approximating arbitrary preferences with a d-dimensional Euclidean model?

Prior work answers question 1 [3]. We address questions 2 and 3. A preprint of our complete paper is available online [18].

Related Work. Important context is provided by Peters [16] who showed that the problem of determining whether a preference profile is d-Euclidean is, in general, NP-hard, and that some ordinal profiles require exponentially many bits to be represented in the Euclidean model. Because of these hardness results, it is usually not feasible to check whether a given ordinal profile is d-Euclidean, or to compute its best approximation with a d-dimensional Euclidean model (a task known as *multidimensional unfolding* [2, 6, 14]).

A related line of work discusses the limitations of the Euclidean preference model from the perspective of measurement theory and psychometric validity [4, 5, 10, 11, 17]. A general review of structured preference models is given by Elkind et al. [7].

Notation. The preference of individual i is denoted π_i (a ranked list) or $>_i$ (the corresponding order relation). The list of all preferences in a population of I individuals is called a *profile*. For given values of A and I, the set of all possible profiles is denoted $\mathcal{P}_{A,I}$. The number of *unique* preferences in a profile is denoted I^*.

2 THREE QUESTIONS

2.1 Are All Profiles Euclidean?

Bogomolnaia and Laslier [3] previously identified the minimum dimensionality required to represent all profiles of a given size.

Theorem 1 (Bogomolnaia and Laslier 2007). *All profiles $\Pi \in \mathcal{P}_{A,I}$ are Euclidean of dimension d if and only if $d \geq M$ where either $M = \min\{I-1, A-1\}$ or $M = \min\{I, A-1\}$, depending on A and I.*

Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), A. Ricci, W. Yeoh, N. Agmon, B. An (eds.), May 29 – June 2, 2023, London, United Kingdom.

Thus, the answer to the first question is no, not all profiles are d-Euclidean. If $d < \min\{I-1, A-1\}$, then there exists at least one profile $\Pi \in \mathcal{P}_{A,I}$ which cannot be losslessly represented with a d-Euclidean model. Is this really that big a deal? Maybe the profiles that are not d-Euclidean are only a small number of pathological edge cases that are unlikely to be encountered in the real world.

2.2 How Common are Non-Euclidean Profiles?

For a given $d < \min\{I-1, A-1\}$, what proportion of preference profiles in $\mathcal{P}_{A,I}$ are not d-Euclidean? Bogomolnaia and Laslier [3] identified a class of pathological sub-profiles that, if present, causes a profile to not be d-Euclidean for some d.

Definition 1. *A circulant pathology of size k is a preference sub-profile of k alternatives $a_1, \ldots, a_k$ and individuals $1, \ldots, k$ such that*

$$\begin{array}{ccccccccc} a_1 & \succ_1 & a_2 & \succ_1 & \cdots & \succ_1 & a_{k-1} & \succ_1 & a_k \\ a_2 & \succ_2 & a_3 & \succ_2 & \cdots & \succ_2 & a_k & \succ_2 & a_1 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_k & \succ_k & a_1 & \succ_k & \cdots & \succ_k & a_{k-2} & \succ_k & a_{k-1}. \end{array}$$

If a profile $\Pi \in \mathcal{P}_{A,I}$ contains a circulant pathology of size k as a sub-profile, then Π is not d-Euclidean for any $d \le k-2$ [3]. By calculating the probability that a circulant pathology arises in a profile constructed uniformly at random, we derive a lower bound on the proportion of profiles that are not d-Euclidean.

Theorem 2 (lower bound on probability of circulant pathology)**.** *Let A, I, and d be fixed positive integers such that $d < \min\{I, A-1\}$, and $\mathbf{P}(C)$ be the probability that a profile chosen uniformly from $\mathcal{P}_{A,I}$ contains a circulant pathology of size $k \ge d+2$. Then,*

$$\mathbf{P}(C) \ge 1 - \left(1 - \sum_{k=d+2}^{I} B_k\right)^{\left\lfloor \frac{A}{d+2} \right\rfloor},$$

where $B_k = \binom{I}{k}\left\{ {k \atop d+2} \right\}(d+2)!\left(\frac{1}{(d+2)!}\right)^k\left(1 - \frac{d+2}{(d+2)!}\right)^{I-k}$ and $\left\{ {k \atop d+2} \right\}$ denotes a Stirling number of the second kind.

Numerically evaluating this expression for various A, I and d shows that when $d \ll \min\{I, A\}$, almost all profiles are not d-Euclidean, and for fixed d and A the proportion seems to approach 1 as $I \to \infty$. So what? If we *approximate* them with a Euclidean preference model, is the approximation error big enough to worry about?

2.3 How Large is the Expected Error?

We consider this question from the perspective of individual preferences, rather than complete profiles. Specifically, we (1) take an arbitrary profile $\Pi \in \mathcal{P}_{A,I^*}$ consisting of I^* unique preferences; (2) approximate Π as well as possible in a d-dimensional Euclidean model, and call this model Euclidean(Π); (3) observe a new preference π generated uniformly from among all $A!$ preferences; and (4) let $\hat{\pi}$ be the preference representable in Euclidean(Π) that minimizes $m(\hat{\pi}, \pi)$ for some error measure m. That is, $\hat{\pi}$ is the closest possible approximation to π in such a Euclidean preference model. As our measure of error m, we use the Kendall tau distance: $m(\pi, \pi') = \#$ pairwise disagreements between π and π' [13]. We are interested in both $\mathbf{E}[m(\pi, \hat{\pi})]$ and $\mathbf{E}[m(\pi, \hat{\pi})]/\binom{A}{2}$, which is the

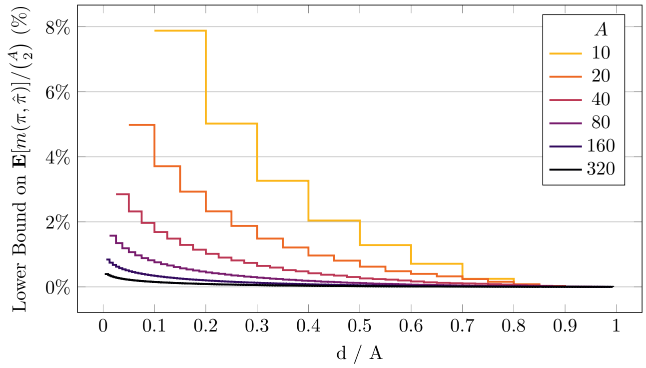

Figure 1: Lower bound on the expected error, as a percentage of the maximum possible error.

expected number of adjacent swaps required as a proportion of the maximum possible number of swaps between any two preferences. Let $r_{A,d} \le A!$ be the maximum number of unique preferences that can be simultaneously represented in Euclidean model of dimension d. In a lemma too long to state here [18, Lemma 1], we derive an upper bound $\hat{r} \ge r_{A,d}$. This bound, along with a geometric structure called a *permutohedron* [8], allows us to prove the following result.

Theorem 3 (lower bound on expected error)**.** *Let A, d be fixed positive integers such that $d < A-1$, $\Pi \in \mathcal{P}_{A,I^*}$ consist of I^* unique preferences, π be a preference chosen uniformly at random from the set of $A!$ possible preferences, $\hat{\pi}$ be the nearest preference to π that is representable in Euclidean(Π) (that is, the representable preference that can be reached in the fewest number of adjacent swaps), and K be a positive integer such that $K \le \binom{A}{2}$. If $I^* \ge r_{A,d}$, then*

$$\mathbf{E}[m(\pi, \hat{\pi})] \ge \sum_{k=0}^{K} \frac{(A! - n_{k,A})_{\hat{r}}}{(A!)_{\hat{r}}} \mathbf{1}(\hat{r} < A! - n_{k,A}),$$

where $(\cdot)_{\hat{r}}$ denotes a falling factorial and $n_{k,A} = \min\{(A-1)^k,\ A!\}$.

Figure 1 plots some values of the lower bound on $\mathbf{E}[m(\pi, \hat{\pi})]/\binom{A}{2}$.

3 CONCLUSIONS

Our theoretical bounds show that when $d \ll \min\{I, A\}$ almost all preference profiles are not d-Euclidean and, in some cases, the expected error when approximating a newly observed preference in the Euclidean model is at least 7% of the maximum possible error, as measured by the Kendall tau distance. The assumptions of our results may be met in social choice models of national elections, in text embeddings used for natural language processing, and in embeddings in recommender systems. To the extent that there are true ordinal relationships in such domains, our results suggest that these relationships may not be able to be closely approximated by relatively low dimensional embeddings if the use of those embeddings assumes the Euclidean structure. This includes cases where embeddings are compared using cosine similarity, and when the SoftMax function [9, Sec. 6.2.2.3] is used to recover a probability distribution over a set of alternatives. Our bound also provides a means by which to quantify a trade-off between dimensionality and accuracy, and hence inform the choice of d.

ACKNOWLEDGMENTS

Luke Thorburn was supported by UK Research and Innovation [grant number EP/S023356/1], in the UKRI Centre for Doctoral Training in Safe and Trusted Artificial Intelligence (safeandtrustedai.org). Carmine Ventre acknowledges funding from the UKRI Trustworthy Autonomous Systems Hub (EP/V00784X/1).

REFERENCES

[1] Felipe Almeida and Geraldo Xexéo. 2019. Word Embeddings: A Survey. https://doi.org/10.48550/ARXIV.1901.09069 arXiv:1901.09069

[2] Joseph F. Bennett and William L. Hays. 1960. Multidimensional Unfolding: Determining the Dimensionality of Ranked Preference Data. *Psychometrika* 25, 1 (01 Mar 1960), 27–43. https://doi.org/10.1007/BF02288932

[3] Anna Bogomolnaia and Jean-François Laslier. 2007. Euclidean Preferences. *Journal of Mathematical Economics* 43, 2 (Feb. 2007), 87–98. https://doi.org/10.1016/j.jmateco.2006.09.004

[4] Jon X. Eguia. 2011. Foundations of Spatial Preferences. *Journal of Mathematical Economics* 47, 2 (2011), 200–205. https://doi.org/10.1016/j.jmateco.2010.12.014

[5] Jon X. Eguia. 2013. Challenges to the Standard Euclidean Spatial Model. In *Advances in Political Economy: Institutions, Modelling and Empirical Analysis*. Springer Berlin Heidelberg, Berlin, Heidelberg, 169–180. https://doi.org/10.1007/978-3-642-35239-3_8

[6] Edith Elkind and Piotr Faliszewski. 2014. Recognizing 1-Euclidean Preferences: An Alternative Approach. In *Algorithmic Game Theory (Lecture Notes in Computer Science)*, Ron Lavi (Ed.). Springer, Berlin, Heidelberg, 146–157. https://doi.org/10.1007/978-3-662-44803-8_13

[7] Edith Elkind, Martin Lackner, and Dominik Peters. 2017. Structured Preferences. In *Trends in Computational Social Choice*, Ulle Endriss (Ed.). AI Access, (online), Chapter 10, 187–207.

[8] P. Gaiha and S. K. Gupta. 1977. Adjacent Vertices on a Permutohedron. *SIAM J. Appl. Math.* 32, 2 (March 1977), 323–327. https://doi.org/10.1137/0132025

[9] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. 2016. *Deep Learning*. MIT Press, Cambridge, MA, USA. http://www.deeplearningbook.org.

[10] Marc Henry and Ismael Mourifié. 2013. Euclidean Revealed Preferences: Testing the Spatial Voting Model. *Journal of Applied Econometrics* 28, 4 (2013), 650–666. https://doi.org/10.1002/jae.1276 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/jae.1276

[11] Macartan Humphreys and Michael Laver. 2010. Spatial Models, Cognitive Metrics, and Majority Rule Equilibria. *British Journal of Political Science* 40, 1 (2010), 11–30. https://doi.org/10.1017/S0007123409990263

[12] Daniel Jurafsky and James H. Martin. 2022. Speech and Language Processing. https://web.stanford.edu/~jurafsky/slp3/ed3book_jan122022.pdf. Accessed: 2022-08-09.

[13] Ravi Kumar and Sergei Vassilvitskii. 2010. Generalized Distances between Rankings. In *Proceedings of the 19th International Conference on World Wide Web* (Raleigh, North Carolina, USA) *(WWW '10)*. Association for Computing Machinery, New York, NY, USA, 571–580. https://doi.org/10.1145/1772690.1772749

[14] Oscar Luaces, Jorge Díez, Thorsten Joachims, and Antonio Bahamonde. 2015. Mapping Preferences into Euclidean Space. *Expert Systems with Applications* 42, 22 (Dec. 2015), 8588–8596. https://doi.org/10.1016/j.eswa.2015.07.013

[15] Shimei Pan and Tao Ding. 2019. Social Media-based User Embedding: A Literature Review. https://doi.org/10.48550/ARXIV.1907.00725 arXiv:1907.00725

[16] Dominik Peters. 2017. Recognising Multidimensional Euclidean Preferences. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI'17)*. AAAI Press, San Francisco, CA, USA, 642–648.

[17] Gonzalo Rivero Rodríguez. 2011. *Integrality and Separability in Multidimensional Voting Models: Ideology and Nationalism in Spanish Regional Elections*. Technical Report 265. Juan March Institute Center for Advanced Study in the Social Sciences.

[18] Luke Thorburn, Maria Polukarov, and Carmine Ventre. 2022. Error in the Euclidean Preference Model. https://doi.org/10.48550/ARXIV.2208.08160 arXiv:2208.08160