SHORT TECHNICAL NOTE

A new simplified manual tour, with examples in mathematica

Alex Aumann[a], German Valencia[a], Ursula Laa[b], Dianne Cook[c]

[a]School of Physics and Astronomy, Monash University; [b]Institute of Statistics, University of Natural Resources and Life Sciences, Vienna; [b]Department of Econometrics and Business Statistics, Monash University

ARTICLE HISTORY
Compiled April 1, 2022

ABSTRACT
Something here

KEYWORDS
data visualisation; grand tour; statistical computing; statistical graphics; multivariate data; dynamic graphics

1. Introduction

From a statistical perspective it is rare to have data that are strictly 3D, and so unlike in most computer graphics applications, the more useful methods for data analysis show projections from an arbitrary dimensional space. These are dynamic data visualizations methods and are collected under the term *tours*. Tours involve views of high-dimensional (p) data in low-dimensional (d) projections. In his original paper on the grand tour, Asimov (1985) provided several algorithms for tour paths that could theoretically show the viewer the data *from all sides*. Prior to Asimov's work, there were numerous preparatory developments including **?**'s PRIM-9. PRIM-9 had user-controlled rotations on coordinate axes, allowing one to manually tour through low-dimensional projections. It is impractical to impossible to steer through all possible projections, unlike Asimov's tours which allows one to quickly see many, many different projections. After Asimov there have been many, many tour developments, as summarized in Lee et al. (2021).

One such direction of work develops the ideas from PRIM-9, to provide manual control of a tour. Cook and Buja (1997) describes controls for 1D (or 2D) projections, in a 2D (or 3D) manipulation space, allowing the user to select any variable axis, and rotate it into or out of or around the projection through horizontal, vertical, oblique, radial or angular changes in value. Spyrison and Cook (2020) refines this algorithm and implements them to generate animations.

Manual controls are especially useful for assessing sensitivity of structure to particular elements of the projection. There are many places where it is useful. In exploratory data analysis, where one sees clusters in a projection, can some variables be removed

CONTACT Alex Aumann. Email: aaum0002@student.monash.edu, German Valencia. Email: german.valencia@monash.edu, Ursula Laa. Email: ursula.laa@boku.ac.at, Dianne Cook. Email: dicook@monash.edu

from the projection without affecting the clustering. For interpreting models, one can reduce or increase a variable's contribution to examine the variable importance. These controls can also be used to interactively generate facetted plots (**?**), or spatiotemporal glyphmaps (**?**). Having the user interact with a projection is extremely valuable for understanding high-dimensional data. However, these algorithms have two problems: (1) the pre-processing of creating a manipulation space overly complicates the algorithm, (2) extending to higher dimensional control is difficult.

Through experiments with the relatively new interactive graphics capabilities in mathematica(?), we have realized that there is a simpler approach, which is more direct, and extensible for generating user interaction. This paper explains this, and is organized as follows. The next section describes the new algorithm for manual control. This is followed by details on implementation. The applications section illustrate how these can be used.

2. Manual tour

2.1. *Background*

Describe the manip space, out of manual tour

2.2. *New definition*

1. Provide a k-D projection
2. Select variable to control
3. Detect or provide direction of change
4. Update projection, and orthonormalise

Need to think about how the values for selected variable are made exact, that is constrained orthonormalisation.

Need to think about checks, and error catching, maybe in implementation section. If selected variable has 0 coefficient, will that generate an orthonormalisation error?

3. Implementation

4. Applications

5. Discussion

Acknowledgements

The authors gratefully acknowledge the support of the Australian Research Council. The paper was written in `rmarkdown` (Xie, Allaire, and Grolemund 2018) using `knitr` (Xie 2015).

Supplementary material

The source material and animated gifs for this paper are available at

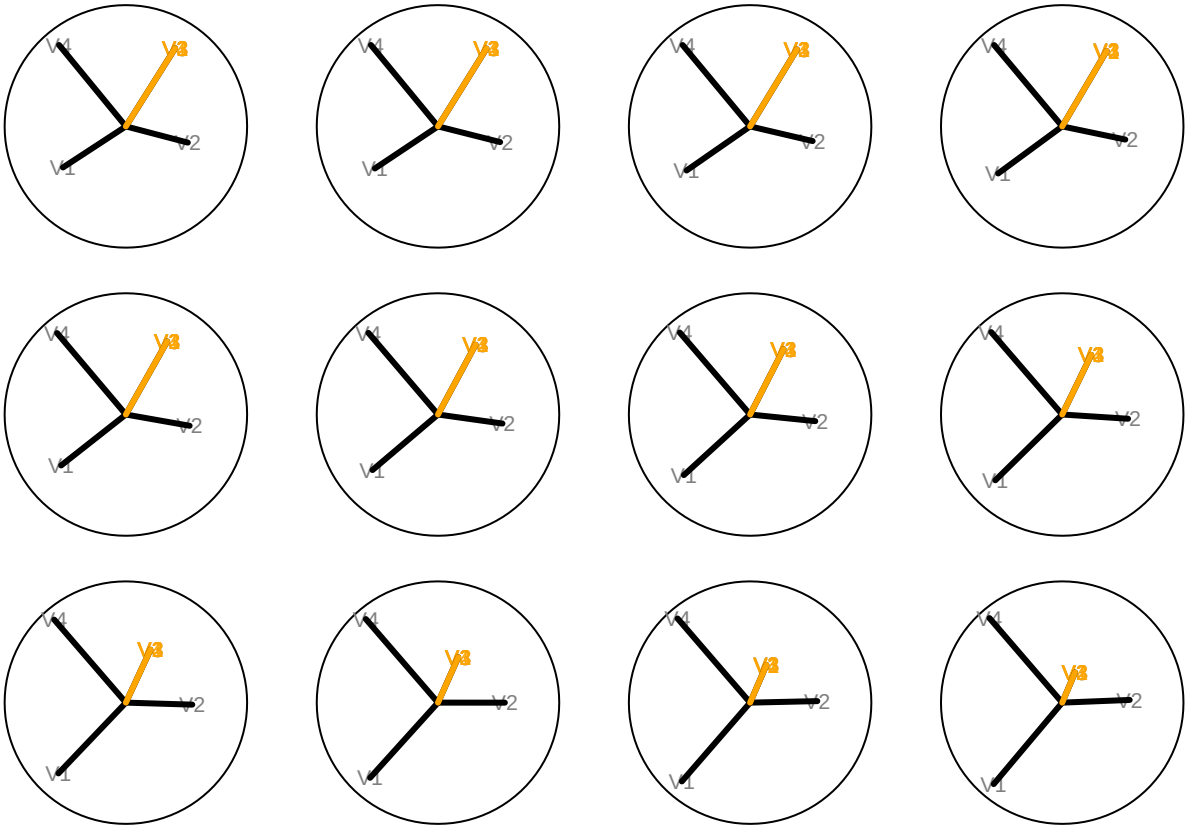

Figure 1. Sequence of projections where V3 contribution is changed.

References

Asimov, D. 1985. "The Grand Tour: A Tool for Viewing Multidimensional Data." *SIAM Journal of Scientific and Statistical Computing* 6 (1): 128–143.

Cook, Dianne, and Andreas Buja. 1997. "Manual Controls for High-Dimensional Data Projections." *Journal of Computational and Graphical Statistics* 6 (4): 464–480. http://www.jstor.org/stable/1390747.

Lee, Stuart, Dianne Cook, Natalia da Silva, Ursula Laa, Earo Wang, Nick Spyrison, and H. Sherry Zhang. 2021. "Advanced Review: The State-of-the-Art on Tours for Dynamic Visualization of High-dimensional Data." *arXiv:2104.08016 [cs, stat]* http://arxiv.org/abs/2104.08016.

Spyrison, Nicholas, and Dianne Cook. 2020. "spinifex: an R Package for Creating a Manual Tour of Low-dimensional Projections of Multivariate Data." *The R Journal* 12 (1): 243. https://journal.r-project.org/archive/2020/RJ-2020-027/index.html.

Xie, Yihui. 2015. *Dynamic Documents with R and knitr*. 2nd ed. Boca Raton, Florida: Chapman and Hall/CRC. https://yihui.name/knitr/.

Xie, Yihui, Joseph J. Allaire, and Garrett Grolemund. 2018. *R Markdown: The Definitive Guide*. Chapman and Hall/CRC. https://bookdown.org/yihui/rmarkdown.