Response to reviewer

Article ID: 2021-69

We really appreciate the helpful comments and feedback. Below is a point-by-point explanation of how we have changed the paper. The reviewer comments are in black and our responses are in red.

In our revision, we have used the new format for the journal and produced both a pdf format and an html format. The html format includes interactive graphics whenever appropriate.

Reviewer 1

The Review

This work presents visual diagnostic tools for different methods of projection pursuit optimization. They use line plots, PCA and the tour method to visually compare the behavior of different optimizers. Finally, they showcased the analysis of three optimizers using their tool.

The tool is likely to be of interest to the R community. Most of my comments are about the technical aspect of the content. The writing is overall good, but please see my comment under the "writing" regarding to the clarity of the text.

Major comment

- "Sometimes, two optimisers can start with the same basis but finish with bases of opposite sign… However, this creates difficulties for comparing the optimisers since the end bases will be symmetric to the origin. A sign flipping step is conducted… " and the "Reconciling the orientation" section
  - For 1D projections, flipping the sign is sufficient. For p-D projections of $\mathbb{R}^n$, one may need a little more. To my knowledge most index functions are invariant to any rotations that happens within the p-subspace. This makes me wonder that in general, one may want to align two bases so that their difference does not contain any "within-subspace" components. In other words, one may want to compare representations in the Grassmannian manifold (space of k-subspaces) instead of in the Stiefel manifold (space of orthonormal k-frames).

Your comment involves two different, and quite separate aspects, of projection pursuit. One is that index functions should be rotation invariant, that within the same plane any rotation of the index should have same value. There are exceptions where rotation variant indexes can still be very useful, see discussion in Laa and Cook (2020). And yes, operating on a Stiefel manifold can help - actually this is a work in progress to add to the tourr package. The same optimisation diagnostics reported in this paper should apply to Stiefel manifolds in addition to Grassmann manifolds.

The second element of your comment refers to the direction of projection vectors. The index value will be the same regardless of direction, but the representation of a vector as a dot on the surface of a sphere means that it may be located on either side depending on the direction. Technically it is simply that the signs of the coefficient are opposite. You encounter this in many different analyses, including PCA. In PCA, there is a convention applied to force the first element of the project vector to be positive. That is analogous to what we are doing with the flipping operation. The text in section "Reconciling the orientation" has been updated to reflect this explanation.

Minor comment

- The software seems to be specific to projection pursuit. I wish the authors could comment more on its potential generality to other optimizations. For example, can one generalizes the tour method to visualize any parameter trace of an optimizer?

It is feasible to generalise the tour method for diagnosing other optimisers, given that the parameter space can be simulated. In our examples, the orthonormality constraint regulates the parameter space of the projection basis ($p \times d$) to be a sphere, or torus, which can be simulated by the geozoo package. In other problems the parameter space may take different forms in the p-dimensional space. Nevertheless, this can be visualised and the ferrn package should be possible to use. Researchers need first to create a data collection object in a tibble structure, analgous to the one described in our paper, that tracks the parameter during the optimisation. Then the ferrn diagnostic plots could be used. Some additional examples of the data collection objects that we created are included in the ferrn package and should be referred to as the format for inputs into the explore functions. We have added several sentences in the conclusion explaining this.

- Figure 4 "…All the bases in PD have been flipped for easier comparison of the final bases…"
  - This is related to the major comment. In the Figure, PD flipped in sign but stopped at a basis representation similar to the destination of the CRS. Flipping the sign back, does it mean that PD and CRS diverged but stopped at the similar subspace with opposite basis representations? Also, given that PD and CRS have the same starting bases and one is flipped, why wouldn't the dashed line pass the origin (gray dot) in the figure?

Thanks for sharing your interpretation on the effects of flipping, that's right! The gray dot should land on the dashed line and this was a mistake in the code. This has now been corrected in the package and Figure 4 has been updated accordingly.

- "Recall that the columns in a 2D basis are orthogonal to each other, so the space of $p \times 2$ bases is a p-D torus".
  - It is not clear to me why "so". For p=2, the space of 2x2 bases is a pair of 1-D tori instead of just one (because you have two orientations for a pair of orthogonal vectors), right? For p=3, the space of 3x2 bases is SO(3), which is not a 3-D torus? Ref: https://en.wikipedia.org/wiki/Stiefel_manifold

We changed the wording of the sentence to: "a torus in the p-D space" and provide a reference for this statement.

Writing

- "Along with the bases recorded during the optimisation and a zero basis, the parameters (centre and radius) of the 2D space can be estimated by PCA. The centre of the circle is the first two PC coordinates of the zero matrix, and the radius is estimated as the largest distance from the centre to the basis."
  - It took me multiple passes to get what this means. Perhaps instead starting with something like: "We use PCA to project and visualize the parameters/bases in 2D. The centre of the 2D view is…"?

The sentence has been re-phrased as suggested.

- "Figure 6 shows some frames from the tour plot of the same two optimisations in its original 5D space."
  - I think this is the first mention of "5D space"/5D data in the main text (apart from the caption in Figure 4), so it may be worth clarifying what the original 5D space is.

The data set is first introduced in the subsection Simulation setup under Diagnosing an optimiser. To avoid confusion, we change the wording from "original 5D space" to "original space" and give more details about the original space in the figure caption: "The basis space in this example is a 5D unit sphere ...".

- "While the previous few examples have looked at the space of 1D bases … 2D bases"

- I don't think the terms "1D bases" and "2D bases" are standard. Perhaps (orthonormal) "1-frames" and "2-frames"?

We keep the term "1D bases" and "2D basis" in the paper since 1) "frame" is conventionally "basis", and 2) we use this term also in the software. To make this clearer, a sentence is added at the end of section "Projection pursuit guided tour" to note this.

Reviewer 2

These experiments are interesting and so is the methodology. The conclusions could do with beefing up What can we learn about how to optimise the holes index for 1 and 2d? Does this apply to other smooth indices? What about non-smooth indices? What about indices in other dimensions?

The conclusion has been expanded.

1. Should refer to earlier work on scagnostic indices, page 3

Two references of Wilkinson work on scagnostics have been added.

2. figure 2 caption "using the holes index" "and a boxplot". What is the colour scale for the points? The text related to figure 2 in the bottom of page 5 should mention the dataset and the index, giving references if needed.

The caption has now been updated and explanation for the color scale is added."

3. "with higher index value" -> "with a higher index value". The word "the" is missing in a few other places in the paper. Corrected.
4. In Figure 3 would be better if the y-axes scales were the same for the two graphs. Fixed.
5. What is the star in figures 4 and 5? This is not mentioned in the text or caption

The black star represents the theoretical best basis the optimisers aiming to find. This has been added in the caption of Figure 4, but not Figure 5 as this refers back to Figure 4 in the caption.

6. In the text you say things like "details=TRUE" and "animate=TRUE" but what functions are you referring to?

"details=TRUE" and "animate=TRUE" refers to the function "explore space pca". These two arguments are now mentioned in the implementation section under function "explore space pca", instead of subsection Diagnostic 3a and 3b.

7. In Figure 9, it seems CRS is doing better here than in Figure 8 – maybe this could be mentioned.

In the experiment of Figure 9, the parameter "max.tries" is set to 400 to allow guided tour to do its best instead of a default 25 as in Figure 8. A sentence in the Figure 9 captions is added to reflect this.

8. "not the case for all the indexes" missing a "." page 11 Added.
9. For the Kolmogorov index, why are you using randomly generated data instead of normal quantiles?

The Kolmogorov Smirnoff index here looks at the difference between a theoretical pdf and an empirical one based on the data. In our case we are not assuming normality on the theoretical pdf and that is the reason why we are not using normal quantiles. To make it clearer, we change the wording from "Normal Kolmogorov Index" to "Kolmogorov Index".

10. Fix references with "et al" Fixed!