Ms. No.: JMVA-14-443
Title: Scale and Dimension Reduction in Spaces of Positive Definite Matrices
Corresponding Author: Mr. Drew Lazar
Authors:

Dear Mr. Lazar,

Thank you for submitting your manuscript to the Journal of Multivariate Analysis. The reviewers have made suggestions which the Editor feels would improve your manuscript. The Editor encourages you to consider these comments and make an appropriate revision of your manuscript. The reviewers' comments are below.

Please submit your revision online within 8 weeks by logging onto the Elsevier Editorial System for the Journal of Multivariate Analysis:
http://ees.elsevier.com/jmva/
Your username is: lazar2d@cmich.edu
If you need to retrieve password details, please go to: http://ees.elsevier.com/JMVA/automail_query.asp.

The manuscript record can be found in the "Submissions Needing Revision" menu.

NOTE: Upon submitting your revised manuscript, please upload the source files for your article. For additional details regarding acceptable file formats, please refer to the Guide for Authors at: http://www.elsevier.com/journals/journal-of-multivariate-analysis/0047-259X/guide-for-authors

When submitting your revised paper, we ask that you include the following items:

Response to Reviewers (mandatory)

This should be a separate file labeled "Response to Reviewers" that carefully addresses, point-by-point, the issues raised in the comments appended below. You should also include a suitable rebuttal to any specific request for change that you have not made. Mention the page, paragraph, and line number of any revisions that are made.

Manuscript and Figure Source Files (mandatory)

We cannot accommodate PDF manuscript files for production purposes. We also ask that when submitting your revision you follow the journal formatting guidelines. Figures and tables may be embedded within the source file for the submission as long as they are of sufficient resolution for Production. For any figure that cannot be embedded within the source file (such as *.PSD Photoshop files), the original figure needs to be uploaded separately. Refer to the Guide for Authors for additional information.

http://www.elsevier.com/journals/journal-of-multivariate-analysis/0047-259X/guide-for-authors

Highlights (Optional)

Highlights consist of a short collection of bullet points that convey the core findings of the article and should be submitted in a separate file in the online submission system. Please use 'Highlights' in the file name and include 3 to 5 bullet points (maximum 85 characters, including spaces, per bullet point). See the following website for more information

http://www.elsevier.com/highlights

Please note that this journal offers a new, free service called AudioSlides: brief, webcast-style presentations that are shown next to published articles on ScienceDirect (see alsohttp://www.elsevier.com/audioslides). If your paper is accepted for publication, you will automatically receive an invitation to create an AudioSlides presentation.

Thank you, and we look forward to receiving your revised manuscript.

With kind regards,

Kavitha Balu
Journal Manager
Journal of Multivariate Analysis
Elsevier
E-mail: jmva@elsevier.com

Reviewers' comments:

Reviewer #1: Summary:

This paper presents a new method for dimensionality reduction in the space of symmetric positive-definite matrices, P(n), under the standard GL(n)-invariant Riemannian metric. The proposed method is recursive principal geodesic analysis (RPGA), which is an alternative to other PGA methods that seek to find lower-dimensional geodesic subspaces that minimize residual variance (or maximize projected variance, as discussed in the paper). RPGA recursively removes variance in one geodesic dimension at a time, via the GL(n) group action, followed by removing the mean from the data (which may change, unlike the linear case). Experiments are presented comparing linearized PGA (LPGA), PGA, and RPGA on synthesized data.

Comments:

1. There is no motivation given in the abstract/introduction for RPGA. Are there advantages to it over PGA, LPGA, or GPCA? And under what certain scenarios? Relatedly, there is no real-world motivation for why one would need to do data analysis on P(n).

2. One major limitation of RPGA is that its definition and computations as presented in this paper are limited to the specific manifold P(n). PGA, on the other hand, is defined for any Riemannian manifold (at least up to the cut locus of the mean). There should be a discussion about the generality of RPGA. Is it extendible to other manifolds, and what are the assumptions needed? It seems that it would at least be limited to homogeneous spaces, i.e., spaces with a transitive group of isometries in order for the main operations of demeaning and variance removal, (13), (14), to be defined.

3. Another disadvantage of RPGA is the need to demean the data in each iteration. This is opposed to PGA, where there is a single mean as the base point for geodesics. The reason this is a disadvantage is that in order to encode, or reconstruct, a data point (projected in the lower-dimensional space), one needs to reverse this process. So, for RPGA one would need to store the demeaning direction of each iteration, which is tangent vector (representing the geodesic segment that was the difference to the mean). So, to encode the data, we now need double the dimensions for each principal component. This makes RPGA a much less efficient representation than PGA, which only requires a single mean along with the principal components.

4. It is also not clear what RPGA is optimizing. PGA directly minimizes sum-of-squared residuals (7). In Euclidean spaces, RPCA clearly does minimize this same objective (i.e., it is equivalent to PCA). However, it is not clear that RPGA on P(n) minimizes (7) or some other objective. This should be discussed. Is RPGA an

approximation to PGA? Is there another criteria that RPGA is optimizing instead. It seems to be just an algorithm that is motivated by the RPCA algorithm, but not motivated by a rigorous optimality criteria. Perhaps it could be justified as a greedy algorithm approximation to PGA?

5. Finally, the experimental results are somewhat limited and unconvincing. First, there are only simulated experiments and no real-world examples. While real-world examples are not absolutely necessary, it does leave the reader wondering how RPGA compares in realistic data situations. Are the variances used in the simulation indicative of any realistic data scenarios? Second, it is not clear why RPGA would be chosen over PGA or LPGA in these results. In most scenarios at dimension 2 or 3, it has slightly higher error than PGA with slightly faster computation time. But at dimension 4, it has worse error and computation time. If you want lower error, it seems PGA is the way to go, and you don't save enough time to make RPGA worth it. If you need very fast computation and can deal with more error, it seems LPGA is the way to go. Third, this is not very much dimensionality reduction in a space that was initially 6-dimensional. I imagine dimensionality reduction would be more
important in P(n) spaces with larger n.

6. There are several missing references to the literature on PGA:

First, there is the following paper on exact PGA for the specific case of SO(3):
Said et al., EUSIPCO 2007
http://www.eurasip.org/Proceedings/Eusipco/Eusipco2007/Papers/c5p-h02.pdf

Next, is the following paper on exact PGA also by Sommer, which discusses how Jacobi fields can be used to optimize the principal geodesics. I think the Jacobi field approach should be discussed in relation to the numerical gradients computed for RPGA.
Sommer et al., Advances in Computational Mathematics, Volume 40, Issue 2, pp 283-313, 2014.
http://arxiv.org/abs/1008.1902

Finally, there is a recent probabilistic interpretation of PGA that should be referenced and discussed:
Zhang and Fletcher, Neural Information Processing Systems (NIPS) 2014
http://papers.nips.cc/paper/5133-probabilistic-principal-geodesic-analysis.pdf

There are many more references that should probably be included, and I recommend the author go through the cited references of these papers as well.

Overall, I think the RPGA idea is an interesting addition to the growing work on manifold statistical analysis. The analysis of curvature effects with increasing \epsilon radius of the data is particularly interesting and well done. However, there are some concerns over the methodology listed above that I think need to be carefully addressed.

Reviewer #2: The paper concerns dimensionality reduction in non-linear Riemannian manifolds,
particularly spaces of positive definite matrices. In Euclidean space, the various different
formulations of principal component analysis are all equivalent
as a result of the inner product structure and the Pythagorean theorem. This is not the case
in non-linear spaces. The paper proposes a new generalization - recursive
principal geodesic analysis (RPGA) - that generalizes the fact that variability
in low dimensional subspaces can be removed recursively when performing PCA in

Euclidean space. RPGA thus treats a problem similar to the existing PCA generalizations to the Riemannian manifold case (PGA, GPCA, PNS, HCA) but in a new way. As a second contribution, the effect of curvature and data spread on the differences between operations used many of those methods and their linear counterparts are quantified using expansions in a scaling parameter.

General comments:

The RPGA construction is a solid idea and its property of removing variance in low-dimensional subspaces is important. This latter fact gives it some similarity with the HCA construction (see below). The paper gives very little motivation for the construction, and the single synthetic experiment could be extended to give a much more detailed view of properties of the method. The contribution would stand out much more clearly if properly motivated and tested.

The exploration of the link between curvature and the linearized/non-linearized operations is very novel. The ref [9] touched upon this subject but the rigorous exploration using Taylor expansions in the scale parameter gives a much improved picture of the differences. One often observes that statistical tools that properly models the non-linear geometry give comparable results to the linearized versions such as LPGA. The exploration presented in the paper allows theoretical quantification of such empirical observations. The subject is of great interest to the shape statistics community. This part of the paper is a major part of the contribution and it could be given more weight in the introduction and abstract.

Based on the above, I recommend that (1) the presentation of the RPGA construction is accompanied by further discussion of situations where the procedure is expected to improve analyses (that is, the reasons for making the construction) and experiments validating this; and (2) that the introduction puts more emphasis on the scale expansions and linearized/non-linearized operations.

Further comments:
The paper focuses solely on the space P(n) of symmetric positive definite matrices. A discussion on possibilities of generalizing the method to other Lie groups with a transitive action would strengthen the paper.

P. 4: "\pi_{(S(V)}(p) is also guarantied to exist and be unique": Please add a reference.

Section 5: An illustration showing the effect of \Gamma and \Phi would be very helpful in interpreting the maps.

The idea of removing the mean bears resemblance with the HCA procedure (see http://link.springer.com/chapter/10.1007/978-3-642-40020-9_7) that removes the effect of variability in lower dimensional subspaces by moving the fitted low dimensional linear subspace V\subset T_IM to the projections of the data points in the generated subspaces. With RPGA, removing variability makes the mean a better basepoint for measuring the remaining variability. The symmetry of P(n) allows this. With HCA, the fitted subspaces at the mean are transported to the projections to obtain a similar effect.

The first sentence of the abstract directs the focus to data with non-linear constraints. This view is continued in section 2. Data that can only be properly modeled in non-linear spaces are not limited to this class (for example, in medical imaging, LDDMM landmark matching and shape modeling).

Section 7: As the author notes in the conclusion, the expansions are very related to curvature terms. A discussion on the possibility of extending the analysis to manifolds besides the symmetric space P(n) would be interesting.

Section 8:
As I understand the description of the experiment with simulated data in P(3), the data are generated by first sampling a normal distribution in T_IP(3) and then mapping the data to the manifold using the exponential map. This basically means that LPGA should provide the best model as the Log map would map the data back to the tangent space where it would be normally distributed. It is quite hard to see how the non-linearized methods can improve the analysis for data that are essentially already linear. Even if lower SSR are measured, it doesn't mean that a more precise picture of the data are obtained. A different way of measuring the performance could for example be to see how well the estimated directions align with the eigenvectors of the covariance matrix used when sampling the normal distribution.

This point relates to the lack of proper motivation for the RPGA procedure. If a data set (either synthetic or real) where RPGA would be an obviously better procedure was presented, the paper would be stronger. Analyzing data uniformly distributed around a geodesic (or higher dimensional subspace) could be a suggestion.

Typos and minor comments:
eq. (8) and beyond: in tr(v_jLog_i(p)), I believe Log_i(p) is considered a row vector. Though the meaning is clear, the notation when using traces instead of dot-products is not perfectly consist.

P. 5: in non-linear -> is non-linear
P. 5: ladder -> latter