Principal Coordinates Analysis

abbreviation PCoA from Legendre and Legendre, 3rd edition, 2012 (*The Bible*)

I. PCoA: The best of both worlds?

A. The advantages of PCA

* independent axes
* true orthogonal axes
* variance explained on each axis

B. The advantages of NMDS

* Can use any distance metric
* Largely assumption free

C. PCoA combines most of these advantages

* An eigenanalysis on a distance matrix

|  |  |  |  |
| --- | --- | --- | --- |
| TABLE 1 | | **Gradient Type** | |
| ***Linear*** | ***Unimodal*** |
| **Relationship with Explanatory Variables** | ***Unconstrained*** | PCA  NMDS  PCoA | NMDS  PCoA  ?? |
| ***Constrained*** | ?? | ?? |

*Distance metric required:*

Principal Components = correlation or covariance

Nonmetric Multidimensional Scaling = Any

Principal Coordinates = Any (sort of…)

II. Interesting historical notes

* Originally described by Gower in 1966
* Even though it was developed rather early in the use of multivariate statistics, PCoA largely fell by the wayside
* Other techniques like PCA, CA and constrained ordinations became much more popular
* Only recently (last 12 years) did PCoA start to make a comeback in ecology, largely due to publications by Marti Anderson
* Also because it became part of multiple analysis pipelines for microbial data (QIIME, mothur)
* Also frequently known as Metric Multidimensional Scaling or Classic Multidimensional Scaling

III. Steps in PCoA

A. Choosing a distance metric

* Pretty much the same considerations as NMDS: choose a metric appropriate for your data type
* Can use any distance that will form a square matrix
* However, occasionally, some matrices will not satisfactorily undergo singular value decomposition, i.e., the extraction of eigenvalues/vectors

B. Eigenanalysis

* Very much like the eigenanalysis in PCA
* One important difference: it is performed on a distance matrix based on distances between *observations (rows)*, whereas PCA is performed among *variables (columns)*
* Like with PCA, you get eigenvalues and eigenvectors
* Prior to eigenanalysis, the resemblance matrix is centered and scaled. This centering has the effect of positioning the origin of the new system of axes at the centroid of the original scatter of objects
* There are some complications regarding negative eigenvalues resulting in the need to scale the eigenvectors. I won’t delve into these extensively, but they result because eigenanalysis is an inherently Euclidean and linear process, whereas many distance metrics are *semimetric* (for example, Bray Curtis), i.e., their fundamental domain is not Euclidean

C. Scores

* Scores *are* the vectors from the eigenanalysis
* Stop to think why this is true vs. what happens with PCA
* The exact scores from PCA can be recovered by PCoA by 1) carrying out a PCoA on a Euclidean distance matrix among *objects*; 2) dividing the PCoA eigenvalues by n-1

D. Notes on the process: The problem of negative eigenvalues

* some distance metrics will result in negative eigenvalues
* As mentioned before, a product of trying to project a non-Euclidean object on to Euclidean space
* It is like the remaining fraction of variance from an imperfect solution (*The Bible*)
* Multiple solutions are available to this problem, the most common being by adding a constant to all values in a distance matrix. Exactly how this process works and can be optimized is, quite frankly, a bit over my pay grade, but is described by the equation in Gower and Legendre 1986, Theorem 7\*.
* All current manifestations of PCoA in software apply the Theorem 7\* transformation

IV. Why PCoA is awesome

* Clearly it combines some strengths of PCA (orthogonal, independent axes) with some strengths of NMDS (any distance metric)
* Given modern software, can be run as easily as other techniques

V. Why PCoA is not a panacea

* Still the assumptions of eigenanalysis
* There are no equivalent of PCA’s weights or loadings, which is often a desirable property
* NMDS is better at compressing the values of a resemblance matrix into a few dimensions (from *The Bible*)

VI. How we should think about PCoA as an option for ordination

* This one is probably handled best as a discussion….