
Subject: Principal Components Analysis
Posted by [Denis](#) on Wed, 11 Feb 1998 08:00:00 GMT
[View Forum Message](#) <> [Reply to Message](#)

I am looking for a PCA program that can handle 1500 images (2048x1024 pixels). Does someone have any idea where to find this.

Regards,
Denis

--
Denis Nadeau E-Mail nadeau@carbon.marine.usf.edu
Department of Marine Science Phone: (813) 553-3987
University of South Florida FAX: (813) 553-1103
140 7th Avenue St.Petersburg, FL 33701

Subject: Re: Principal components analysis
Posted by [wmc](#) on Wed, 03 Mar 1999 08:00:00 GMT
[View Forum Message](#) <> [Reply to Message](#)

Mark McGillion <mm@fs1.co.umist.ac.uk> wrote:
> What I need to know is.....what contribution is made to these derived
> variables by each of the original 500 variables. Is it possible to
> determine
> which of the original 500 variables provided the greatest contribution
> to each derived variable?

This isn't really an IDL question... there is a lot about principal components (aka empirical orthogonal functions) in the meteorological literature. Try, for example, Richman (1986, J. Climatology) for some examples of how ignorance can make you go badly wrong ;-). It also has lots of refs to earlier studies...

- William

--
William M Connolley | wmc@bas.ac.uk | <http://www.nbs.ac.uk/public/icd/wmc/>
Climate Modeller, British Antarctic Survey | Disclaimer: I speak for myself

Subject: Re: Principal components analysis
Posted by [menakkis](#) on Wed, 03 Mar 1999 08:00:00 GMT
[View Forum Message](#) <> [Reply to Message](#)

Mark McGillion <mm@fs1.co.umist.ac.uk> wrote:
> I am computing the principal components and derived variables of a set

- > of power spectral data (using IDL v5.1). The power spectral data is a
- > matrix of high dimensionality [500,400], the power spectra has 500
- > points, and there are 400 examples of this spectra. The 500 derived
- > variables calculated from the principle components describe the data
- > completely (100% of the variance). In my case, I find that the first
- > two derived variables describe over 95% of the variance, therefore I can
- > easily visualise my original data in 2-dimensions. Great!
- >
- > What I need to know is.....what contribution is made to these derived
- > variables by each of the original 500 variables. Is it possible to determine
- > which of the original 500 variables provided the greatest contribution
- > to each derived variable?

The values of your eigenvectors will show the contributions from each variable. The way the principal components rotation is done on one sample SAMP (of 500 channels or variables) is like this: The first principal component channel is $TOTAL((samp - global_mean) * evec1)$. (Each variable here is an array[500].) The second principal component channel is done the same, except using evec2 instead of evec1. And so on up to a possible 500 principal component channels. (Of course this is usually done in one matrix mult rather than a series of array dot products.) If you are using PCOMP, you can get the (scaled) eigenvectors by using the /COEFFICIENTS switch. e.g., The first (scaled) eigenvector will be in the first row of the returned matrix.

I'm not sure that you can always rely on an easy interpretation of the eigenvectors, especially when you have so many channels that are obviously highly correlated.

Peter Mason

-----== Posted via Deja News, The Discussion Network ==-----
<http://www.dejanews.com/> Search, Read, Discuss, or Start Your Own
