Subject: Principal component analysis
Posted by Haje Korth on Wed, 05 Dec 2007 14:00:11 GMT

View Forum Message <> Reply to Message

Hi,
I am puzzled by principal component analysis. I calculated the eigenvalues
using both PCOMP and IMSP_PRINC_COMP routines. Could someone enlighten me
why the results are completely different? I have tried different keywords to
see whether I can match them by trial and error, but I had no success. There
must be someone out there who undertstands this much better than I do.

Thanks so much,
Haje


```
IDL> a=[[1,-2,-6],[-2,1,-3],[-6,-3,5]]
IDL> pca=pcomp(a,eigenvalues=ev) & print,transpose(ev)
      2.24227    0.757732    0.000000
IDL> ev=imsl_princ_comp(a) & print,ev
      9.53359    -5.19751    2.66392
```


Subject: Re: Principal component analysis
Posted by Haje Korth on Wed, 05 Dec 2007 16:51:05 GMT

View Forum Message <> Reply to Message

yes, but the imsl routine had a finite third eigenvalue while PCOMP showed
this value as zero. That was difficult to understand. Vince came up with the
call that will produce the same result as pcomp:
IMSL_PRINC_COMP(CORRELATE(a)). Again, I am not quite sure I understand this.
Shouldn't the correlation analysis be part of the PCA?

"David Fanning" <news@dfanning.com> wrote in message
news:MPG.21c07edf91c4354298a12d@news.frii.com...
> Haje Korth writes:
>
>> thanks for validating this. I tend to just go with PCOMP since I don't
>> really know what the IMSL routine actually does. As I wrote in the
>> response
>> to David I just got thrown off by not being able to reconcile the output
>> from the different routines.
>
> It looks to me like passing the CORRELATE results
> to EIGENQL just scales the eigenvalues into -1 to 1. That
> would seem to be a sensible choice to me.
>
> Cheers,
>

> David
> --
> David Fanning, Ph.D.
> Fanning Software Consulting, Inc.
> Coyote's Guide to IDL Programming: http://www.dfanning.com/
> Sepore ma de ni thui. ("Perhaps thou speakest truth.")

Subject: Re: Principal component analysis
Posted by Vince Hradil on Wed, 05 Dec 2007 16:57:06 GMT

View Forum Message <> Reply to Message

On Dec 5, 10:47 am, "Haje Korth" <haje.ko...@nospam.jhuapl.edu> wrote:
> Yup, that'll do it. I am still not sure I understand the logic behind this.
> I though the correlation is part of the PCA.
>
> "Vince Hradil" <hrad...@yahoo.com> wrote in message
>
>   news:8362380a-217a-45d2-b7c4-0198e5931b39@y5g2000hsf.googleg roups.com...
>
>> On Dec 5, 10:08 am, "Haje Korth" <haje.ko...@nospam.jhuapl.edu> wrote:
>>> I have tried that, it gives
>>> IDL> ev=imsl_princ_comp(correlate(a,/cov)) & print,ev
>>>   45.2906   3.70938-2.65683e-006
>
>>> These EVs are the same as you get using PCOMP with /COV keyword.
>
>>> "Vince Hradil" <hrad...@yahoo.com> wrote in message
>
>>> news:54fc6ed8-ccd7-4ac6-8e0d-09f5d190eeac@o6g2000hsd.googleg roups.com...
>
>>>> On Dec 5, 9:12 am, Vince Hradil <hrad...@yahoo.com> wrote:
>>>> > On Dec 5, 8:00 am, "Haje Korth" <haje.ko...@nospam.jhuapl.edu> wrote:
>
>>>> > > Hi,
>>>> > > I am puzzled by principal component analysis. I calculated the
>>>> > > eigenvalues
>>>> > > using both PCOMP and IMSP_PRINC_COMP routines. Could someone
>>>> > > enlighten
>>>> > > me
>>>> > > why the results are completely different? I have tried different
>>>> > > keywords to
>>>> > > see whether I can match them by trial and error, but I had no
>>>> > > success.
>>>> > > There
>>>> > > must be someone out there who undertstands this much better than I
>>>> > > do.
>
>

>>>> > > Thanks so much,
>>>> > > Haje
>
>>>> > > IDL> a=[[1,-2,-6],[-2,1,-3],[-6,-3,5]]
>>>> > > IDL> pca=pcomp(a,eigenvalues=ev) & print,transpose(ev)
>>>> > >   2.24227   0.757732   0.000000
>>>> > > IDL> ev=imsl_princ_comp(a) & print,ev
>>>> > >   9.53359   -5.19751   2.66392
>
>>>> > From the HELP:
>
>>>> > Syntax
>>>> > Result = IMSL_PRINC_COMP(covariances [, /COV_MATRIX]
>>>> > [, /CORR_MATRIX] [, CORRELATIONS=variable] [, CUM_PERCENT=variable] [,
>>>> > DF=variable] [, /DOUBLE] [, EIGENVECTORS=variable] [,
>>>> > STDEV=variable])
>
>>>> > Note that IMSL_PRINC_COMP requires that you pass the covariance or
>>>> > correlation matrix - not the vectors.
>
>>>> so maybe try
>>>> ev=imsl_princ_comp(correlate(a,/covariance) & print, ev
>>>> (I don't have an analyst license)
>
>> There you go 8^)
>> How about
>> ev=imsl_princ_comp(correlate(a)) & print, ev

Oh, yes correlation IS part of PCA, it's just that IMSL decided to let
the user do that part him/herself. IMSL_PRINC_COMP calculates the
principal components of the cov/cor matrix. Calculating these
principal components is just part of "Principle Components Analysis".
IMSL leaves it up to the user to decide how to implement the principal
components in his/her analysis.

Subject: Re: Principal component analysis
Posted by Haje Korth on Wed, 05 Dec 2007 17:02:28 GMT
View Forum Message <> Reply to Message

THanks, I guess I was expecting the whole package for a routine with that
name. And the help was not helping either. :-)

"Vince Hradil" <hradilv@yahoo.com> wrote in message
 news:d447b810-a195-4e34-8dd6-a353ba4a3ac9@b40g2000prf.google groups.com...
> On Dec 5, 10:47 am, "Haje Korth" <haje.ko...@nospam.jhuapl.edu> wrote:
>> Yup, that'll do it. I am still not sure I understand the logic behind

>> this.
>> I though the correlation is part of the PCA.
>>
>> "Vince Hradil" <hrad...@yahoo.com> wrote in message
>>
>> news:8362380a-217a-45d2-b7c4-0198e5931b39@y5g2000hsf.googleg roups.com...
>>
>>> On Dec 5, 10:08 am, "Haje Korth" <haje.ko...@nospam.jhuapl.edu> wrote:
>>>> I have tried that, it gives
>>>> IDL> ev=imsl_princ_comp(correlate(a,/cov)) & print,ev
>>>>   45.2906   3.70938-2.65683e-006
>>
>>>> These EVs are the same as you get using PCOMP with /COV keyword.
>>
>>>> "Vince Hradil" <hrad...@yahoo.com> wrote in message
>>
>>>> news:54fc6ed8-ccd7-4ac6-8e0d-09f5d190eeac@o6g2000hsd.googleg roups.com...
>>
>>>> > On Dec 5, 9:12 am, Vince Hradil <hrad...@yahoo.com> wrote:
>>>> >> On Dec 5, 8:00 am, "Haje Korth" <haje.ko...@nospam.jhuapl.edu>
>>>> >> wrote:
>>
>>>> >> > Hi,
>>>> >> > I am puzzled by principal component analysis. I calculated the
>>>> >> > eigenvalues
>>>> >> > using both PCOMP and IMSP_PRINC_COMP routines. Could someone
>>>> >> > enlighten
>>>> >> > me
>>>> >> > why the results are completely different? I have tried different
>>>> >> > keywords to
>>>> >> > see whether I can match them by trial and error, but I had no
>>>> >> > success.
>>>> >> > There
>>>> >> > must be someone out there who undertstands this much better than
>>>> >> > I
>>>> >> > do.
>>
>>>> >> > Thanks so much,
>>>> >> > Haje
>>
>>>> >> > IDL> a=[[1,-2,-6],[-2,1,-3],[-6,-3,5]]
>>>> >> > IDL> pca=pcomp(a,eigenvalues=ev) & print,transpose(ev)
>>>> >> >   2.24227   0.757732   0.000000
>>>> >> > IDL> ev=imsl_princ_comp(a) & print,ev
>>>> >> >   9.53359   -5.19751   2.66392
>>
>>>> >> From the HELP:
>>

>>>> >> Syntax
>>>> >> Result = IMSL_PRINC_COMP(covariances [, /COV_MATRIX]
>>>> >> [, /CORR_MATRIX] [, CORRELATIONS=variable] [, CUM_PERCENT=variable]
>>>> >> [,
>>>> >> DF=variable] [, /DOUBLE] [, EIGENVECTORS=variable] [,
>>>> >> STDEV=variable])
>>
>>>> >> Note that IMSL_PRINC_COMP requires that you pass the covariance or
>>>> >> correlation matrix - not the vectors.
>>
>>>> > so maybe try
>>>> > ev=imsl_princ_comp(correlate(a,/covariance) & print, ev
>>>> > (I don't have an analyst license)
>>
>>> There you go 8^)
>>> How about
>>> ev=imsl_princ_comp(correlate(a)) & print, ev
>
> Oh, yes correlation IS part of PCA, it's just that IMSL decided to let
> the user do that part him/herself. IMSL_PRINC_COMP calculates the
> principal components of the cov/cor matrix. Calculating these
> principal components is just part of "Principle Components Analysis".
> IMSL leaves it up to the user to decide how to implement the principal
> components in his/her analysis.