Bilan financier de groupes pétroliers
Période 1969-1984

This document is a mix of:
indications to run the analysis,
indications about R features.

Outils utilisés : Excel, R

Table des matières

RAW DATA

The ACP analysis is done from the following raw data:

Année	NET	INT	SUB	LMT	DCT	IMM	EXP	VRD
1969	17.93	3.96	0.88	7.38	19.86	25.45	5.34	19.21
1970	16.21	3.93	0.94	9.82	19.11	26.58	5.01	18.40
1971	19.01	3.56	1.91	9.43	17.87	25.94	5.40	16.88
1972	18.05	3.33	1.73	9.72	18.83	26.05	5.08	17.21
1973	16.56	3.10	2.14	9.39	20.36	23.95	6.19	18.31
1974	13.09	2.64	2.44	8.10	25.05	19.48	11.61	17.59
1975	13.43	2.42	2.45	10.83	22.07	22.13	11.17	15.49
1976	9.83	2.46	1.79	11.81	24.10	22.39	11.31	16.30
1977	9.46	2.33	2.30	11.46	24.45	23.07	11.16	15.77
1978	10.93	2.95	2.25	10.72	23.16	24.17	9.64	16.20
1979	13.02	3.74	2.21	7.99	23.04	19.53	12.60	17.87
1980	13.43	3.60	2.29	7.09	23.59	17.61	16.67	15.72
1981	13.37	3.35	2.58	6.76	23.94	18.04	15.42	16.54
1982	11.75	2.74	3.11	7.37	25.04	18.11	14.71	17.18
1983	12.59	3.05	3.85	7.12	23.40	19.17	11.86	18.97
1984	13.00	3.00	4.00	7.00	24.00	20.00	12.00	17.00

NET	Situation nette ; représente l'ensemble des capitaux propres de l'entreprise.
INT	Intérêts ; représente l'ensemble des frais financiers supportés par l'entreprise.
SUB	Subventions ; représente le montant total des subventions accordées par l'Etat.
LMP	Dettes à long et moyen terme.
DCT	Dettes à court terme.
IMM	Immobilisations ; représente l'ensemble des terrains et du matériel de l'entreprise.
EXP	Valeurs d'exploitation.
VRD	Valeurs réalisables et disponibles ; ensemble des créances à court terme de l'entreprise.

BASIC STATISTICS

From the raw data, basic statistical indicators can be computed. The following results are obtained from R.

```
> bfgp<-read.table(file="bfgp.csv",sep=";",header=T,row.name=1)
> bfgp

       NET  INT  SUB   LMT   DCT   IMM   EXP   VRD
1969 17.93 3.96 0.88  7.38 19.86 25.45  5.34 19.21
1970 16.21 3.93 0.94  9.82 19.11 26.58  5.01 18.40
1971 19.01 3.56 1.91  9.43 17.87 25.94  5.40 16.88
1972 18.05 3.33 1.73  9.72 18.83 26.05  5.08 17.21
1973 16.56 3.10 2.14  9.39 20.36 23.95  6.19 18.31
1974 13.09 2.64 2.44  8.10 25.05 19.48 11.61 17.59
1975 13.43 2.42 2.45 10.83 22.07 22.13 11.17 15.49
1976  9.83 2.46 1.79 11.81 24.10 22.39 11.31 16.30
1977  9.46 2.33 2.30 11.46 24.45 23.07 11.16 15.77
1978 10.93 2.95 2.25 10.72 23.16 24.17  9.64 16.20
1979 13.02 3.74 2.21  7.99 23.04 19.53 12.60 17.87
1980 13.43 3.60 2.29  7.09 23.59 17.61 16.67 15.72
1981 13.37 3.35 2.58  6.76 23.94 18.04 15.42 16.54
1982 11.75 2.74 3.11  7.37 25.04 18.11 14.71 17.18
1983 12.59 3.05 3.85  7.12 23.40 19.17 11.86 18.97
1984 13.00 3.00 4.00  7.00 24.00 20.00 12.00 17.00
```

Min / Max / Mean

```
> summary(bfgp)
```

```
      NET             INT             SUB             LMT       
 Min.   : 9.46   Min.   :2.330   Min.   :0.880   Min.   : 6.760  
 1st Qu.:12.38   1st Qu.:2.715   1st Qu.:1.880   1st Qu.: 7.308  
 Median :13.23   Median :3.075   Median :2.270   Median : 8.745  
 Mean   :13.85   Mean   :3.135   Mean   :2.304   Mean   : 8.874  
 3rd Qu.:16.30   3rd Qu.:3.570   3rd Qu.:2.482   3rd Qu.:10.045  
 Max.   :19.01   Max.   :3.960   Max.   :4.000   Max.   :11.810  
      DCT             IMM             EXP             VRD       
 Min.   :17.87   Min.   :17.61   Min.   : 5.010   Min.   :15.49  
 1st Qu.:20.23   1st Qu.:19.40   1st Qu.: 5.992   1st Qu.:16.27  
 Median :23.28   Median :22.26   Median :11.240   Median :17.09  
 Mean   :22.37   Mean   :21.98   Mean   :10.323   Mean   :17.16  
 3rd Qu.:24.02   3rd Qu.:24.49   3rd Qu.:12.150   3rd Qu.:17.98  
 Max.   :25.05   Max.   :26.58   Max.   :16.670   Max.   :19.21  
```

Variance / Covariance

```
> cov(bfgp)
```

```
          NET        INT        SUB        LMT        DCT
NET  **8.423612**  1.0582800 -1.1314108 -1.0241508 -6.1256742
INT  1.058280  **0.2824400** -0.2018100 -0.4226967 -0.7564900
SUB -1.131411 -0.2018100  **0.7138796** -0.5979071  1.2265412
LMT -1.024151 -0.4226967 -0.5979071  **2.9894529** -0.7718187
DCT -6.125674 -0.7564900  1.2265412 -0.7718187  **5.6142629**
IMM  5.066636  0.4114500 -1.8469437  3.2582229 -6.1029954
EXP -7.870113 -0.6950833  1.9769854 -2.5981613  7.8835904
VRD  1.602360  0.3248200 -0.1385433 -0.8369100 -0.9665167

            IMM         EXP        VRD
NET   5.0666358  -7.8701125  1.6023600
INT   0.4114500  -0.6950833  0.3248200
SUB  -1.8469437   1.9769854 -0.1385433
LMT   3.2582229  -2.5981613 -0.8369100
DCT  -6.1029954   7.8835904 -0.9665167
IMM   **9.9446329** -11.4660313  0.7354367
EXP -11.4660313  **14.8165963** -2.0492900
VRD   0.7354367  -2.0492900  **1.3316267**
```

Correlation factor

```
> cor(bfgp)
```

```
           NET        INT        SUB        LMT        DCT
NET  1.0000000  0.6861014 -0.4613799 -0.2040887 -0.8907552
INT  0.6861014  1.0000000 -0.4494352 -0.4600127 -0.6007499
SUB -0.4613799 -0.4494352  1.0000000 -0.4092846  0.6126653
LMT -0.2040887 -0.4600127 -0.4092846  1.0000000 -0.1883966
DCT -0.8907552 -0.6007499  0.6126653 -0.1883966  1.0000000
IMM  0.5535746  0.2455046 -0.6931813  0.5975724 -0.8167747
EXP -0.7044620 -0.3397814  0.6078785 -0.3903876  0.8643775
VRD  0.4784314  0.5296493 -0.1420960 -0.4194611 -0.3534856

           IMM        EXP        VRD
NET  0.5535746 -0.7044620  0.4784314
INT  0.2455046 -0.3397814  0.5296493
```

```
SUB -0.6931813  0.6078785 -0.1420960
LMT  0.5975724 -0.3903876 -0.4194611
DCT -0.8167747  0.8643775 -0.3534856
IMM  1.0000000 -0.9445926  0.2020969
EXP -0.9445926  1.0000000 -0.4613579
VRD  0.2020969 -0.4613579  1.0000000
```

First analysis based on correlation factor

In the above, we can identify “high” values (closed to ‘-1’ or to ‘+1’) for the correlation factors:

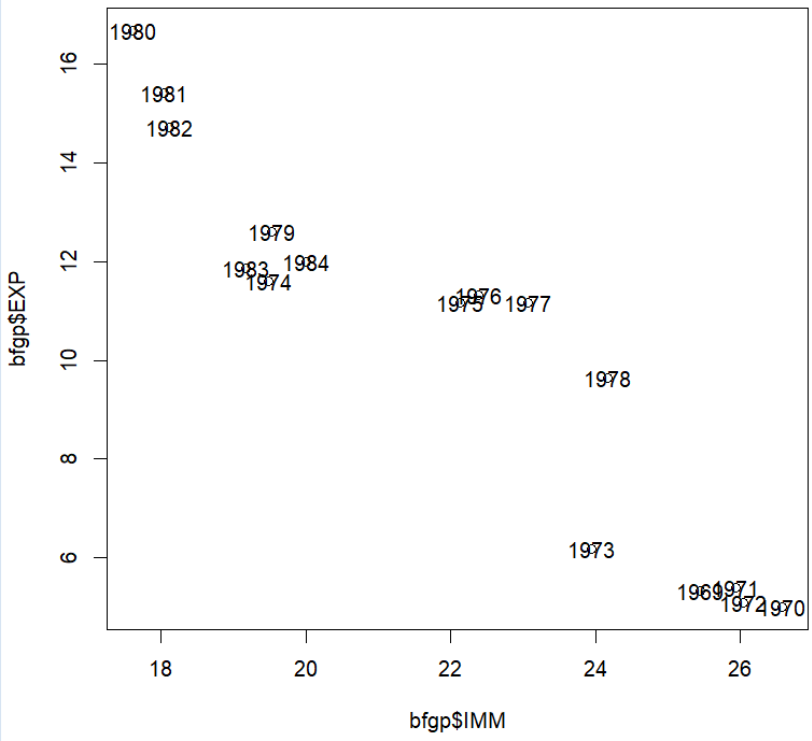
```
            NET        INT        SUB        LMT        DCT
NET  1.0000000  0.6861014 -0.4613799 -0.2040887 -0.8907552
INT  0.6861014  1.0000000 -0.4494352 -0.4600127 -0.6007499
SUB -0.4613799 -0.4494352  1.0000000 -0.4092846  0.6126653
LMT -0.2040887 -0.4600127 -0.4092846  1.0000000 -0.1883966
DCT -0.8907552 -0.6007499  0.6126653 -0.1883966  1.0000000
IMM  0.5535746  0.2455046 -0.6931813  0.5975724 -0.8167747
EXP -0.7044620 -0.3397814  0.6078785 -0.3903876  0.8643775
VRD  0.4784314  0.5296493 -0.1420960 -0.4194611 -0.3534856

            IMM        EXP        VRD
NET  0.5535746 -0.7044620  0.4784314
INT  0.2455046 -0.3397814  0.5296493
SUB -0.6931813  0.6078785 -0.1420960
LMT  0.5975724 -0.3903876 -0.4194611
DCT -0.8167747  0.8643775 -0.3534856
IMM  1.0000000 -0.9445926  0.2020969
EXP -0.9445926  1.0000000 -0.4613579
VRD  0.2020969 -0.4613579  1.0000000
```

The most significant is on the IMM / EXP correlation. We can decide to draw a 2D cloud based on these two variables. Two other drawings are also proposed:

IMM / EXP NET / DCT DCT / EXP DCT / IMM

IMM / EXP

```
> plot(bfgp$IMM,bfgp$EXP)
> text(bfgp$IMM,bfgp$EXP,rownames(bfgp))
```



Analysis :

… bring your own conclusions here…

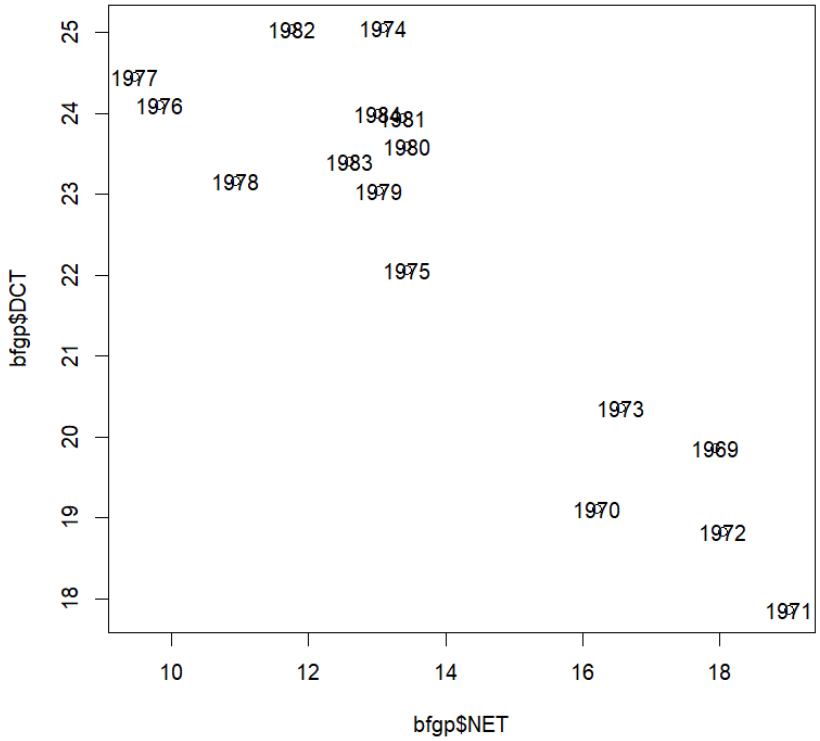
Remark:

2D drawing can be saved in a file as follow:

```
> png(filename="IMMEXP.png")
> plot(bfgp$IMM,bfgp$EXP)
> text(bfgp$IMM,bfgp$EXP,rownames(bfgp))
> dev.off()
windows
      2
>
```

NET / DCT

```
> plot(bfgp$NET,bfgp$DCT)
> text(bfgp$NET,bfgp$DCT,rownames(bfgp))
```

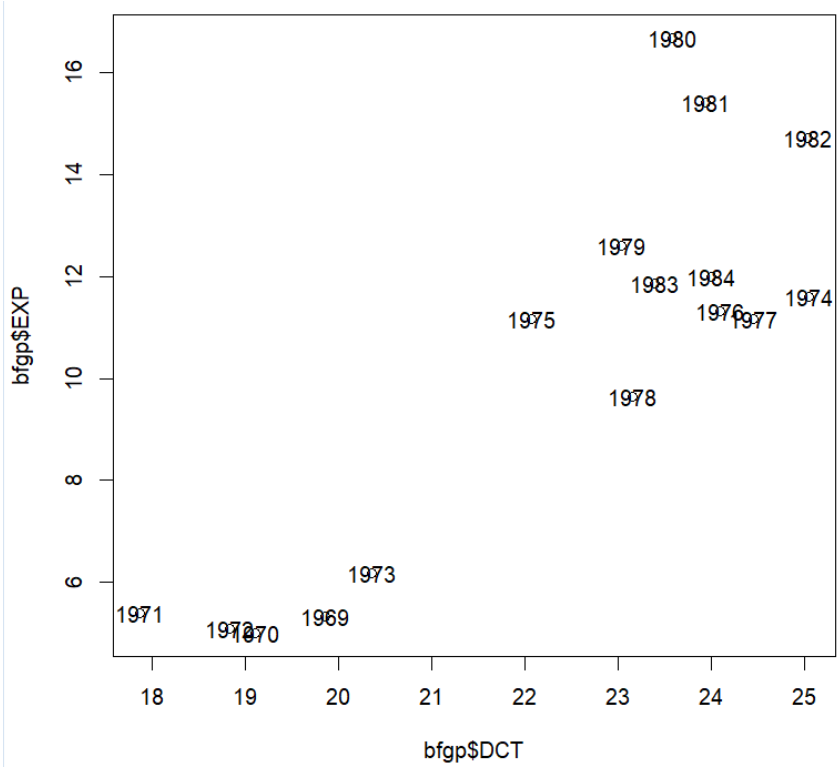


Analysis :

… bring your own conclusions here…

DCT / EXP

```
> plot(bfgp$DCT,bfgp$EXP)
> text(bfgp$DCT,bfgp$EXP,rownames(bfgp))
```

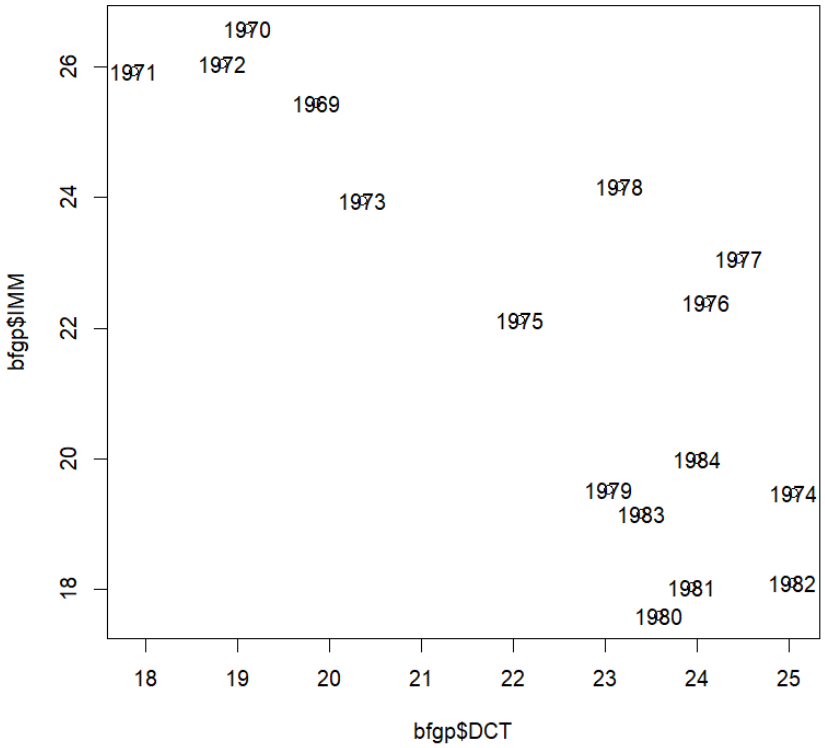


Analysis :

… bring your own conclusions here…

DCT / IMM

```
> plot(bfgp$DCT,bfgp$IMM)
> text(bfgp$DCT,bfgp$IMM,rownames(bfgp))
```



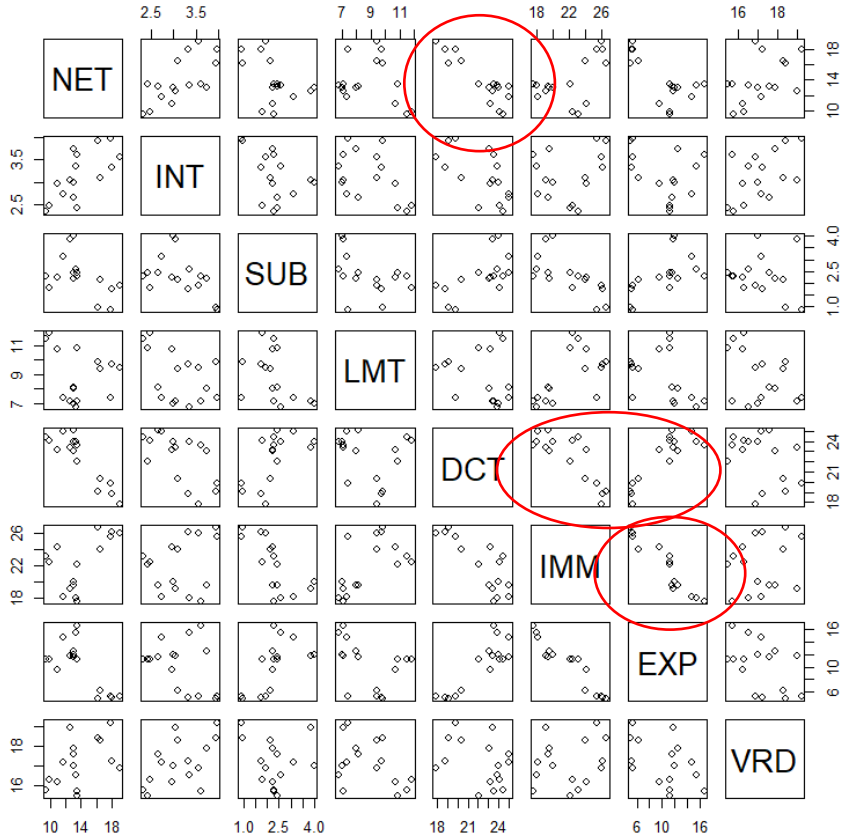
Analysis :

… bring your own conclusions here…

Remark:

These “high” correlation factors may be identified from the overview of 2D drawings using the “pairs” R command. If specific types of drawings are identified (lines for instance), we could identify high correlation values:

```
> pairs(bfgp)
```



Strong correlation (from matrix): IMM / EXP - NET / DCT - DCT / EXP - DCT / IMM

NORMALIZED DATA

The ACP will be executed with centered and reduced data ("normalized" data). We can generate que display theses transformed data.

Centered data

```
>  bfgpC<-scale(bfgp,center=TRUE,scale=FALSE)

> bfgpC
```

	NET	INT	SUB	LMT	DCT	IMM	EXP	VRD
1969	4.07625	0.825	-1.424375	-1.494375	-2.506875	3.470625	-4.983125	2.045
1970	2.35625	0.795	-1.364375	0.945625	-3.256875	4.600625	-5.313125	1.235
1971	5.15625	0.425	-0.394375	0.555625	-4.496875	3.960625	-4.923125	-0.285
1972	4.19625	0.195	-0.574375	0.845625	-3.536875	4.070625	-5.243125	0.045
1973	2.70625	-0.035	-0.164375	0.515625	-2.006875	1.970625	-4.133125	1.145
1974	-0.76375	-0.495	0.135625	-0.774375	2.683125	-2.499375	1.286875	0.425
1975	-0.42375	-0.715	0.145625	1.955625	-0.296875	0.150625	0.846875	-1.675
1976	-4.02375	-0.675	-0.514375	2.935625	1.733125	0.410625	0.986875	-0.865
1977	-4.39375	-0.805	-0.004375	2.585625	2.083125	1.090625	0.836875	-1.395
1978	-2.92375	-0.185	-0.054375	1.845625	0.793125	2.190625	-0.683125	-0.965
1979	-0.83375	0.605	-0.094375	-0.884375	0.673125	-2.449375	2.276875	0.705
1980	-0.42375	0.465	-0.014375	-1.784375	1.223125	-4.369375	6.346875	-1.445
1981	-0.48375	0.215	0.275625	-2.114375	1.573125	-3.939375	5.096875	-0.625
1982	-2.10375	-0.395	0.805625	-1.504375	2.673125	-3.869375	4.386875	0.015
1983	-1.26375	-0.085	1.545625	-1.754375	1.033125	-2.809375	1.536875	1.805
1984	-0.85375	-0.135	1.695625	-1.874375	1.633125	-1.979375	1.676875	-0.165

```
attr(,"scaled:center")
      NET       INT       SUB       LMT       DCT       IMM       EXP       VRD
13.853750  3.135000  2.304375  8.874375 22.366875 21.979375 10.323125 17.165000
```

```
> summary(bfgpC)
```

```
      NET                INT                SUB                LMT
 Min.   :-4.3937   Min.   :-0.805   Min.   :-1.42437   Min.   :-2.1144
 1st Qu.:-1.4737   1st Qu.:-0.420   1st Qu.:-0.42438   1st Qu.:-1.5669
 Median :-0.6238   Median :-0.060   Median :-0.03438   Median :-0.1294
 Mean   : 0.0000   Mean   : 0.000   Mean   : 0.00000   Mean   : 0.0000
 3rd Qu.: 2.4438   3rd Qu.: 0.435   3rd Qu.: 0.17813   3rd Qu.: 1.1706
 Max.   : 5.1562   Max.   : 0.825   Max.   : 1.69562   Max.   : 2.9356

      DCT                IMM                EXP                VRD
 Min.   :-4.4969   Min.   :-4.3694   Min.   :-5.3131   Min.   :-1.675
 1st Qu.:-2.1319   1st Qu.:-2.5769   1st Qu.:-4.3306   1st Qu.:-0.890
 Median : 0.9131   Median : 0.2806   Median : 0.9169   Median :-0.075
 Mean   : 0.0000   Mean   : 0.0000   Mean   : 0.0000   Mean   : 0.000
 3rd Qu.: 1.6581   3rd Qu.: 2.5106   3rd Qu.: 1.8269   3rd Qu.: 0.815
 Max.   : 2.6831   Max.   : 4.6006   Max.   : 6.3469   Max.   : 2.045
>
```

Reminder – Raw data:

```
> summary(bfgp)
```

```
      NET              INT              SUB              LMT
 Min.   : 9.46   Min.   :2.330   Min.   :0.880   Min.   : 6.760
 1st Qu.:12.38   1st Qu.:2.715   1st Qu.:1.880   1st Qu.: 7.308
 Median :13.23   Median :3.075   Median :2.270   Median : 8.745
 Mean   :13.85   Mean   :3.135   Mean   :2.304   Mean   : 8.874
 3rd Qu.:16.30   3rd Qu.:3.570   3rd Qu.:2.482   3rd Qu.:10.045
 Max.   :19.01   Max.   :3.960   Max.   :4.000   Max.   :11.810

      DCT              IMM              EXP              VRD
 Min.   :17.87   Min.   :17.61   Min.   : 5.010   Min.   :15.49
 1st Qu.:20.23   1st Qu.:19.40   1st Qu.: 5.992   1st Qu.:16.27
 Median :23.28   Median :22.26   Median :11.240   Median :17.09
 Mean   :22.37   Mean   :21.98   Mean   :10.323   Mean   :17.16
 3rd Qu.:24.02   3rd Qu.:24.49   3rd Qu.:12.150   3rd Qu.:17.98
 Max.   :25.05   Max.   :26.58   Max.   :16.670   Max.   :19.21
```

Variance / Covariance and Correlation factor

```
> cov(bfgpC)
```

```
          NET        INT        SUB        LMT        DCT         IMM
NET  8.423612  1.0582800 -1.1314108 -1.0241508 -6.1256742   5.0666358
INT  1.058280  0.2824400 -0.2018100 -0.4226967 -0.7564900   0.4114500
SUB -1.131411 -0.2018100  0.7138796 -0.5979071  1.2265412  -1.8469437
LMT -1.024151 -0.4226967 -0.5979071  2.9894529 -0.7718187   3.2582229
DCT -6.125674 -0.7564900  1.2265412 -0.7718187  5.6142629  -6.1029954
IMM  5.066636  0.4114500 -1.8469437  3.2582229 -6.1029954   9.9446329
EXP -7.870113 -0.6950833  1.9769854 -2.5981613  7.8835904 -11.4660313
VRD  1.602360  0.3248200 -0.1385433 -0.8369100 -0.9665167   0.7354367

            EXP        VRD
NET  -7.8701125  1.6023600
INT  -0.6950833  0.3248200
```

```
SUB   1.9769854 -0.1385433
LMT  -2.5981613 -0.8369100
DCT   7.8835904 -0.9665167
IMM -11.4660313  0.7354367
EXP  14.8165963 -2.0492900
VRD  -2.0492900  1.3316267
```

```
> cor(bfgpC)

           NET        INT        SUB        LMT        DCT        IMM
NET  1.0000000  0.6861014 -0.4613799 -0.2040887 -0.8907552  0.5535746
INT  0.6861014  1.0000000 -0.4494352 -0.4600127 -0.6007499  0.2455046
SUB -0.4613799 -0.4494352  1.0000000 -0.4092846  0.6126653 -0.6931813
LMT -0.2040887 -0.4600127 -0.4092846  1.0000000 -0.1883966  0.5975724
DCT -0.8907552 -0.6007499  0.6126653 -0.1883966  1.0000000 -0.8167747
IMM  0.5535746  0.2455046 -0.6931813  0.5975724 -0.8167747  1.0000000
EXP -0.7044620 -0.3397814  0.6078785 -0.3903876  0.8643775 -0.9445926
VRD  0.4784314  0.5296493 -0.1420960 -0.4194611 -0.3534856  0.2020969

           EXP        VRD
NET -0.7044620  0.4784314
INT -0.3397814  0.5296493
SUB  0.6078785 -0.1420960
LMT -0.3903876 -0.4194611
DCT  0.8643775 -0.3534856
IMM -0.9445926  0.2020969
EXP  1.0000000 -0.4613579
VRD -0.4613579  1.0000000
>
```

Centered and reduced data

```
> bfgpCR<-scale(bfgp,center=T,scale=T)
> bfgpCR

            NET         INT           SUB        LMT        DCT         IMM
1969  1.4044673  1.55235429 -1.685822562 -0.8642985 -1.0580016  1.10055896
1970  0.8118433  1.49590505 -1.614809413  0.5469191 -1.3745316  1.45888970
1971  1.7765801  0.79969767 -0.466763509  0.3213556 -1.8978612  1.25594131
1972  1.4458132  0.36692011 -0.679802955  0.4890823 -1.4927028  1.29082307
1973  0.9324354 -0.06585745 -0.194546439  0.2982209 -0.8469815  0.62489868
1974 -0.2631492 -0.93141258  0.160519305 -0.4478736  1.1323861 -0.79256893
1975 -0.1460026 -1.34537372  0.172354830  1.1310706 -0.1252931  0.04776422
1976 -1.3863785 -1.27010806 -0.608789806  1.6978711  0.7314481  0.13021200
1977 -1.5138616 -1.51472146 -0.005178042  1.4954424  0.8791621  0.34584466
1978 -1.0073748 -0.34810369 -0.064355666  1.0674502  0.3347305  0.69466219
1979 -0.2872676  1.13839315 -0.111697765 -0.5114941  0.2840857 -0.77671359
1980 -0.1460026  0.87496333 -0.017013567 -1.0320251  0.5162077 -1.38555874
1981 -0.1666755  0.40455294  0.326216652 -1.2228865  0.6639217 -1.24920279
1982 -0.7248447 -0.74324842  0.953499466 -0.8700821  1.1281657 -1.22700531
1983 -0.4354236 -0.15993953  1.829328300 -1.0146741  0.4360201 -0.89087205
1984 -0.2941586 -0.25402161  2.006861171 -1.0840782  0.6892441 -0.62767337

            EXP         VRD
1969 -1.2945760  1.77215649
1970 -1.3803074  1.07022654
1971 -1.2789885 -0.24697535
1972 -1.3621219  0.03899611
1973 -1.0737528  0.99223432
1974  0.3343198  0.36829658
1975  0.2200114 -1.45152182
```

```
1976  0.2563822 -0.74959187
1977  0.2174134 -1.20887937
1978 -0.1774704 -0.83624989
1979  0.5915139  0.61093904
1980  1.6488674 -1.25220838
1981  1.3241274 -0.54161262
1982  1.1396750  0.01299870
1983  0.3992678  1.56417725
1984  0.4356387 -0.14298573

attr(,"scaled:center")
      NET       INT       SUB       LMT       DCT       IMM       EXP       VRD
13.853750  3.135000  2.304375  8.874375 22.366875 21.979375 10.323125 17.165000

attr(,"scaled:scale")
      NET       INT       SUB       LMT       DCT       IMM       EXP       VRD
2.9023459 0.5314508 0.8449140 1.7290034 2.3694436 3.1535112 3.8492332 1.1539613

> summary(bfgpCR)

      NET                INT                SUB                LMT
 Min.   :-1.5139   Min.   :-1.5147   Min.   :-1.68582   Min.   :-1.22289
 1st Qu.:-0.5078   1st Qu.:-0.7903   1st Qu.:-0.50227   1st Qu.:-0.90623
 Median :-0.2149   Median :-0.1129   Median :-0.04068   Median :-0.07483
 Mean   : 0.0000   Mean   : 0.0000   Mean   : 0.00000   Mean   : 0.00000
 3rd Qu.: 0.8420   3rd Qu.: 0.8185   3rd Qu.: 0.21082   3rd Qu.: 0.67705
 Max.   : 1.7766   Max.   : 1.5524   Max.   : 2.00686   Max.   : 1.69787
      DCT                IMM                EXP                VRD
 Min.   :-1.8979   Min.   :-1.38556   Min.   :-1.3803   Min.   :-1.45152
 1st Qu.:-0.8997   1st Qu.:-0.81714   1st Qu.:-1.1251   1st Qu.:-0.77126
 Median : 0.3854   Median : 0.08899   Median : 0.2382   Median :-0.06499
 Mean   : 0.0000   Mean   : 0.00000   Mean   : 0.0000   Mean   : 0.00000
 3rd Qu.: 0.6998   3rd Qu.: 0.79614   3rd Qu.: 0.4746   3rd Qu.: 0.70626
 Max.   : 1.1324   Max.   : 1.45889   Max.   : 1.6489   Max.   : 1.77216

> cov(bfgpCR)

           NET        INT        SUB        LMT        DCT        IMM
NET  1.0000000  0.6861014 -0.4613799 -0.2040887 -0.8907552  0.5535746
INT  0.6861014  1.0000000 -0.4494352 -0.4600127 -0.6007499  0.2455046
SUB -0.4613799 -0.4494352  1.0000000 -0.4092846  0.6126653 -0.6931813
LMT -0.2040887 -0.4600127 -0.4092846  1.0000000 -0.1883966  0.5975724
DCT -0.8907552 -0.6007499  0.6126653 -0.1883966  1.0000000 -0.8167747
IMM  0.5535746  0.2455046 -0.6931813  0.5975724 -0.8167747  1.0000000
EXP -0.7044620 -0.3397814  0.6078785 -0.3903876  0.8643775 -0.9445926
VRD  0.4784314  0.5296493 -0.1420960 -0.4194611 -0.3534856  0.2020969

           EXP        VRD
NET -0.7044620  0.4784314
INT -0.3397814  0.5296493
SUB  0.6078785 -0.1420960
LMT -0.3903876 -0.4194611
DCT  0.8643775 -0.3534856
IMM -0.9445926  0.2020969
EXP  1.0000000 -0.4613579
VRD -0.4613579  1.0000000

> cor(bfgpCR)

           NET        INT        SUB        LMT        DCT        IMM
NET  1.0000000  0.6861014 -0.4613799 -0.2040887 -0.8907552  0.5535746
```

```
INT  0.6861014  1.0000000 -0.4494352 -0.4600127 -0.6007499  0.2455046
SUB -0.4613799 -0.4494352  1.0000000 -0.4092846  0.6126653 -0.6931813
LMT -0.2040887 -0.4600127 -0.4092846  1.0000000 -0.1883966  0.5975724
DCT -0.8907552 -0.6007499  0.6126653 -0.1883966  1.0000000 -0.8167747
IMM  0.5535746  0.2455046 -0.6931813  0.5975724 -0.8167747  1.0000000
EXP -0.7044620 -0.3397814  0.6078785 -0.3903876  0.8643775 -0.9445926
VRD  0.4784314  0.5296493 -0.1420960 -0.4194611 -0.3534856  0.2020969

           EXP        VRD
NET -0.7044620  0.4784314
INT -0.3397814  0.5296493
SUB  0.6078785 -0.1420960
LMT -0.3903876 -0.4194611
DCT  0.8643775 -0.3534856
IMM -0.9445926  0.2020969
EXP  1.0000000 -0.4613579
VRD -0.4613579  1.0000000
```

« High » correlation :

IMM / EXP NET / DCT DCT / EXP DCT / IMM

COMPUTING PRINCIPAL COMPONENTS

Eigen values / eigen vectors

Eigen values and eigen vectors can be computed if required (direct call of “princomp” library in R hides this computation).
This is usually done on centered and reduced data.

Correlation matrix

```
> bfgpCR_cor<-cor(bfgpCR)

> bfgpCR_cor

           NET        INT        SUB        LMT        DCT        IMM
NET  1.0000000  0.6861014 -0.4613799 -0.2040887 -0.8907552  0.5535746
INT  0.6861014  1.0000000 -0.4494352 -0.4600127 -0.6007499  0.2455046
SUB -0.4613799 -0.4494352  1.0000000 -0.4092846  0.6126653 -0.6931813
LMT -0.2040887 -0.4600127 -0.4092846  1.0000000 -0.1883966  0.5975724
DCT -0.8907552 -0.6007499  0.6126653 -0.1883966  1.0000000 -0.8167747
IMM  0.5535746  0.2455046 -0.6931813  0.5975724 -0.8167747  1.0000000
EXP -0.7044620 -0.3397814  0.6078785 -0.3903876  0.8643775 -0.9445926
VRD  0.4784314  0.5296493 -0.1420960 -0.4194611 -0.3534856  0.2020969

           EXP        VRD
NET -0.7044620  0.4784314
INT -0.3397814  0.5296493
SUB  0.6078785 -0.1420960
LMT -0.3903876 -0.4194611
DCT  0.8643775 -0.3534856
IMM -0.9445926  0.2020969
EXP  1.0000000 -0.4613579
VRD -0.4613579  1.0000000
```

Eigen computation

```
> bfgpCR_eigen<-eigen(bfgpCR_cor)
```

```
> bfgpCR_eigen

$values

[1] 4.470371e+00 2.114846e+00 6.806590e-01 5.007466e-01 1.595783e-01
[6] 6.414833e-02 9.649896e-03 6.169196e-07

$vectors

            [,1]         [,2]        [,3]        [,4]        [,5]        [,6]
[1,]  0.40208788  0.238456880  0.07644079  0.47231215  0.50034217  0.07509041
[2,]  0.29779237  0.427526054  0.45243556 -0.16589935 -0.62810989  0.02306718
[3,] -0.35100790  0.189653459 -0.49815138  0.57943955 -0.47376279  0.05137648
[4,]  0.09467414 -0.661253362 -0.05952318 -0.11123175 -0.18993506  0.55288819
[5,] -0.45114040  0.001153742 -0.06371899 -0.36720944  0.12964108 -0.47378080
[6,]  0.41047282 -0.307837381 -0.15054278  0.01741232 -0.27032179 -0.52671297
[7,] -0.43782955  0.144301518  0.36269954 -0.04513239  0.01901321  0.31260427
[8,]  0.23187280  0.414182609 -0.61602506 -0.51389983  0.06608320  0.29306690

            [,7]        [,8]
[1,] -0.3229319 -0.43849271
[2,] -0.3079450 -0.08022711
[3,] -0.1168870 -0.12714849
[4,] -0.3568773 -0.26218923
[5,] -0.5373128 -0.35719467
[6,]  0.3740488 -0.47261927
[7,]  0.4696220 -0.57902219
[8,]  0.1034338 -0.17364939
```

Eigen values / vectors are orders by decreasing eigen value.

The eigen values and vectors can be exported to excel for further usage:

```
>  write.table(bfgpCR_eigen,file="bfgpEigen.csv",sep=";",row.names=FALSE)
```

Excel file:

values	vectors.1	vectors.2	vectors.3	vectors.4	vectors.5	vectors.6	vectors.7	vectors.8
4.47037149636544	0.4020878756827	0.238456880270619	0.0764407947808057	0.472312152216604	0.500342165260708	0.075090407308644	-0.322931850950274	-0.43849271405196
2.11484576363664	0.297792373590449	0.427526053748763	0.452435557513672	-0.165899348062856	-0.628109894793856	0.0230671832784901	-0.307945001385358	-0.0802271116668709
0.680658989053217	-0.351007900958857	0.189653459297271	-0.498151377315259	0.579439554580461	-0.473762790441123	0.0513764761982466	-0.116886969769247	-0.127148492293842
0.500746595090833	0.0946741377998099	-0.661253362327704	-0.0595231766511395	-0.111231752250854	-0.189935057866426	0.552888194099404	-0.356877310042159	-0.262189228174197
0.159578309958799	-0.451140400659727	0.00115374206774771	-0.0637189921330961	-0.367209444446227	0.129641084891054	-0.473780802225921	-0.537312784506875	-0.357194672242966
0.0641483334480973	0.410472815480936	-0.30783738065422	-0.150542780592161	0.017412320604846	-0.270321785695665	-0.526712969620856	0.374048795319611	-0.472619271356207
0.00964989552733856	-0.437829550931114	0.14430151767877	0.362699543563684	-0.0451323922192291	0.0190132140666118	0.312604266981104	0.469622014445808	-0.57902219137288
6.16919624826957e-07	0.231872798171441	0.414182609184996	-0.616025056521777	-0.513899833851751	0.0660831967177032	0.293066903019291	0.103433802560456	-0.173649394399967

R command

The computation of eigen values and vectors is performed by the “princomp” command:

```
> princomp(bfgp,cor=T,scores=T)

Call:
princomp(x = bfgp, cor = T, scores = T)

Standard deviations:

      Comp.1       Comp.2       Comp.3       Comp.4       Comp.5       Comp.6
2.1143253052 1.4542509287 0.8250205992 0.7076345067 0.3994725397 0.2532752129

      Comp.7       Comp.8
0.0982338818 0.0007854423

 8  variables and  16 observations.
```

From “standard deviation” values indicated by the command, we can compute the “variance”.
The variance is equal to the eigen value.

Various information can be found in the result of the command:

- standard deviation

```
> bfgp_acp$sdev
      Comp.1       Comp.2       Comp.3       Comp.4       Comp.5       Comp.6
2.1143253052 1.4542509287 0.8250205992 0.7076345067 0.3994725397 0.2532752129
      Comp.7       Comp.8
0.0982338818 0.0007854423
```

- values (weight) in the principal components / eigen vectors

```
> bfgp_acp$loadings

Loadings:
    Comp.1 Comp.2 Comp.3 Comp.4 Comp.5 Comp.6 Comp.7 Comp.8
NET -0.402 -0.238        -0.472  0.500         0.323  0.438
INT -0.298 -0.428  0.452  0.166 -0.628         0.308
SUB  0.351 -0.190 -0.498 -0.579 -0.474         0.117  0.127
LMT         0.661         0.111 -0.190  0.553  0.357  0.262
DCT  0.451                0.367  0.130 -0.474  0.537  0.357
IMM -0.410  0.308 -0.151        -0.270 -0.527 -0.374  0.473
EXP  0.438 -0.144  0.363                0.313 -0.470  0.579
VRD -0.232 -0.414 -0.616  0.514         0.293 -0.103  0.174
```

Values close to ‘0’ are not printed.

- mean that has been substracted

```
> bfgp_acp$center
      NET       INT       SUB       LMT       DCT       IMM       EXP       VRD
13.853750  3.135000  2.304375  8.874375 22.366875 21.979375 10.323125 17.165000
```

- scaling applied to each variable

```
> bfgp_acp$scale
      NET       INT       SUB       LMT       DCT       IMM       EXP       VRD
2.8101843 0.5145751 0.8180844 1.6741004 2.2942039 3.0533741 3.7270040 1.1173182
```

- number of units (same as for the raw data, if no unit excluded from analysis)

```
> bfgp_acp$n.obs
[1] 16
```

- values of principal components for each unit

```
> bfgp_acp$scores
         Comp.1      Comp.2      Comp.3       Comp.4      Comp.5      Comp.6
1969 -3.5566217 -1.50534936  0.04279714  0.949649924  0.35975977 -0.39963920
1970 -3.5754644  0.04273311  0.22593331  0.845647304 -0.41345012  0.08247740
1971 -3.1202723  0.21808477  0.34197617 -1.346404626 -0.08209617  0.07332437
1972 -2.8755252  0.54758080 -0.03223800 -0.811512543  0.16145909 -0.03567640
1973 -1.8493595 -0.02351833 -0.95033062 -0.171319378  0.32001780  0.25887868
1974  1.4243206 -0.32194490 -0.57143008  0.476034445  0.88214246 -0.19336337
1975  0.7947650  1.97215358  0.20834135 -0.941025227  0.36624284  0.27864837
1976  1.1606959  2.50400214  0.01046878  0.907198076  0.08779416  0.22658406
1977  1.5972624  2.65758461 -0.17763740  0.349548743 -0.14757381 -0.21387089
1978  0.3791794  1.74803264  0.06083084  0.255236738 -0.68233233 -0.33238600
1979  0.3615042 -1.35612358  0.53312263  0.816804577 -0.42374367  0.36624725
1980  1.7596512 -1.20307377  2.06532003 -0.254306496 -0.02941139  0.07377831
1981  1.7500088 -1.40024594  1.07435475 -0.136569103  0.15839355 -0.06602290
1982  2.5183975 -0.84114725 -0.30642366  0.065610264  0.32860315 -0.03283370
1983  1.3791803 -1.88578529 -1.72369320  0.003868896 -0.39560331  0.35370661
1984  1.8522779 -1.15298322 -0.80139203 -1.008461594 -0.49020203 -0.43985259

            Comp.7        Comp.8
1969 -0.133635705 -1.977124e-04
```

```
1970 -0.018037581 -8.109141e-04
1971  0.017311217  3.190325e-04
1972  0.026586697  1.723611e-04
1973  0.079847206  4.261703e-04
1974  0.203388739  3.104798e-04
1975 -0.078517274 -3.423774e-04
1976 -0.002701754 -9.938383e-04
1977 -0.058127505 -4.726204e-04
1978  0.031783400  2.002024e-03
1979  0.169781183  2.866871e-05
1980  0.002896488  2.644305e-04
1981 -0.071631235 -3.916374e-04
1982 -0.137815470  8.932110e-04
1983 -0.123872713  3.121718e-04
1984  0.092744308 -1.519449e-03
```

- command used

```
> bfgp_acp$call
princomp(x = bfgp, cor = T, scores = T)
```

Select principal components

Principal components are selected according to:

- the eigen values,
- the cumulative proportion of “quality” / “quantity of information”.

Proportion / Cumulative proportion

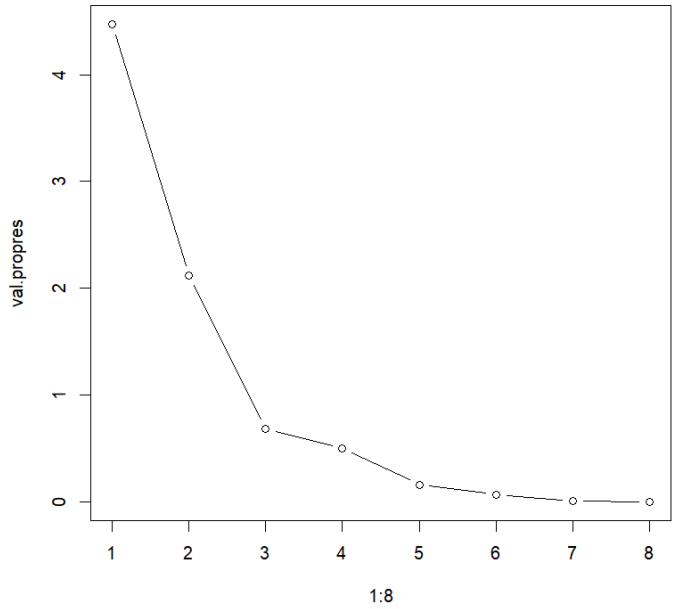
```
> bfgp_acp<-princomp(bfgp,cor=T,scores=T)
> summary(bfgp_acp)

Importance of components:
                          Comp.1    Comp.2     Comp.3     Comp.4     Comp.5
Standard deviation     2.1143253 1.4542509 0.82502060 0.70763451 0.39947254
Proportion of Variance 0.5587964 0.2643557 0.08508237 0.06259332 0.01994729
Cumulative Proportion  0.5587964 0.8231522 0.90823453 0.97082786 0.99077514

                           Comp.6      Comp.7       Comp.8
Standard deviation     0.253275213 0.098233882 7.854423e-04
Proportion of Variance 0.008018542 0.001206237 7.711495e-08
Cumulative Proportion  0.998793686 0.999999923 1.000000e+00
```

From the above, we can consider that the increase of the “cumulative proportion” is important from “comp 1” to “comp 1 & 2” (+27%).
The increase reduces if we go from “comp 1&2” to “comp 1&2&3” (+8%).

Elbow curve

The elbow drawing can show this:

```
> val.propres<-bfgp_acp$sdev^2
> val.propres

      Comp.1       Comp.2       Comp.3       Comp.4       Comp.5       Comp.6
4.470371e+00 2.114846e+00 6.806590e-01 5.007466e-01 1.595783e-01 6.414833e-02
      Comp.7       Comp.8
9.649896e-03 6.169196e-07

> plot(1:8,val.propres,type="b")
```



So :

Principal component 1 only: not useful for drawing
Principal components 1 & 2: yes
Principal components 1, 2 & 3: this adds some "quantity of information" but also brings complexity in the analysis
…?

ANALYSIS OF PRINCIPAL COMPONENTS

"+" / "-" table for each selected principal component

We start from the definition of the principal components (eigen vectors):

```
> bfgp_acp$loadings

Loadings:
    Comp.1 Comp.2 Comp.3 Comp.4 Comp.5 Comp.6 Comp.7 Comp.8
NET -0.402 -0.238        -0.472  0.500         0.323  0.438
INT -0.298 -0.428  0.452  0.166 -0.628         0.308
SUB  0.351 -0.190 -0.498 -0.579 -0.474         0.117  0.127
LMT         0.661         0.111 -0.190  0.553  0.357  0.262
DCT  0.451                0.367  0.130 -0.474  0.537  0.357
IMM -0.410  0.308 -0.151       -0.270 -0.527 -0.374  0.473
EXP  0.438 -0.144  0.363                0.313 -0.470  0.579
VRD -0.232 -0.414 -0.616  0.514         0.293 -0.103  0.174
```

Values close to '0' are not printed.

For each principal component, we identify raw variables with the most negative and most positive weights:

COMP	–	+
1	IMM NET …	… EXP DCT
2	INT VRD …	… IMM LMT
3	VRD SUB …	… EXP INT

We can consider only the 2 or 3 first components (highest eigen values) that we decided to use for the analysis. From this, we can attached some "meaning" to each principal components, explaining why a unit will have a positive or negative value on that axis.

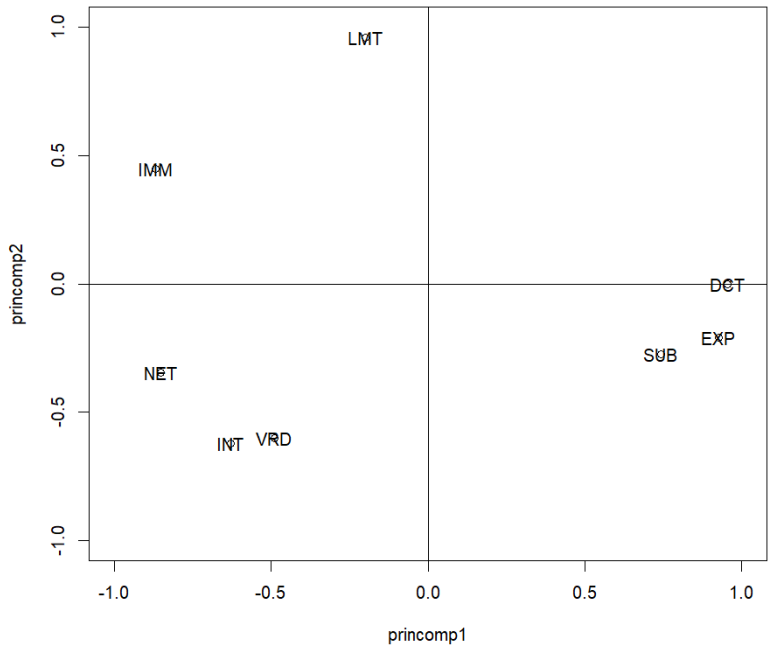
"Correlation circle" – Correlation factor between raw variables and principal components

Weights of raw variables are harmonized using standard deviations of each principal conponents:

```
> princomp1<-bfgp_acp$loadings[,1]*bfgp_acp$sdev[1]
> princomp2<-bfgp_acp$loadings[,2]*bfgp_acp$sdev[2]
> corrc1c2<-cbind(princomp1, princomp2)
> corrc1c2
     princomp1   princomp2
NET -0.8501446 -0.34677614
INT -0.6296300 -0.62173016
SUB  0.7421449 -0.27580372
LMT -0.2001719  0.96162832
DCT  0.9538576 -0.00167783
IMM -0.8678731  0.44767280
EXP  0.9257141 -0.20985062
VRD -0.4902545 -0.60232544
```

Impact of each raw variable is graphically shown in a 2D drawing for the principal components of interest.

```
> plot( princomp1, princomp2,xlim=c(-1,+1),ylim=c(-1,+1))
> abline(h=0,v=0)
> text(princomp1, princomp2,labels=colnames(bfgp))
```

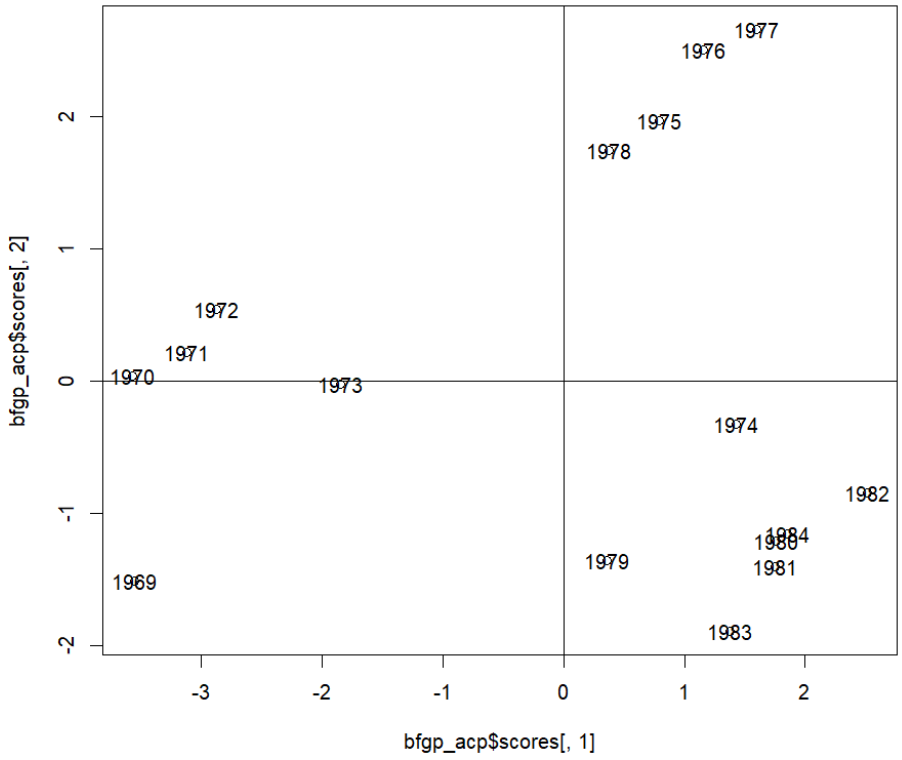


We see hare that principal component 1 opposes variables DCT/EXP/SUB to IMM/NET. The others have less impacts of the values of units on this principal component.
Principal component 2 opposes LMT to INT / VRD.

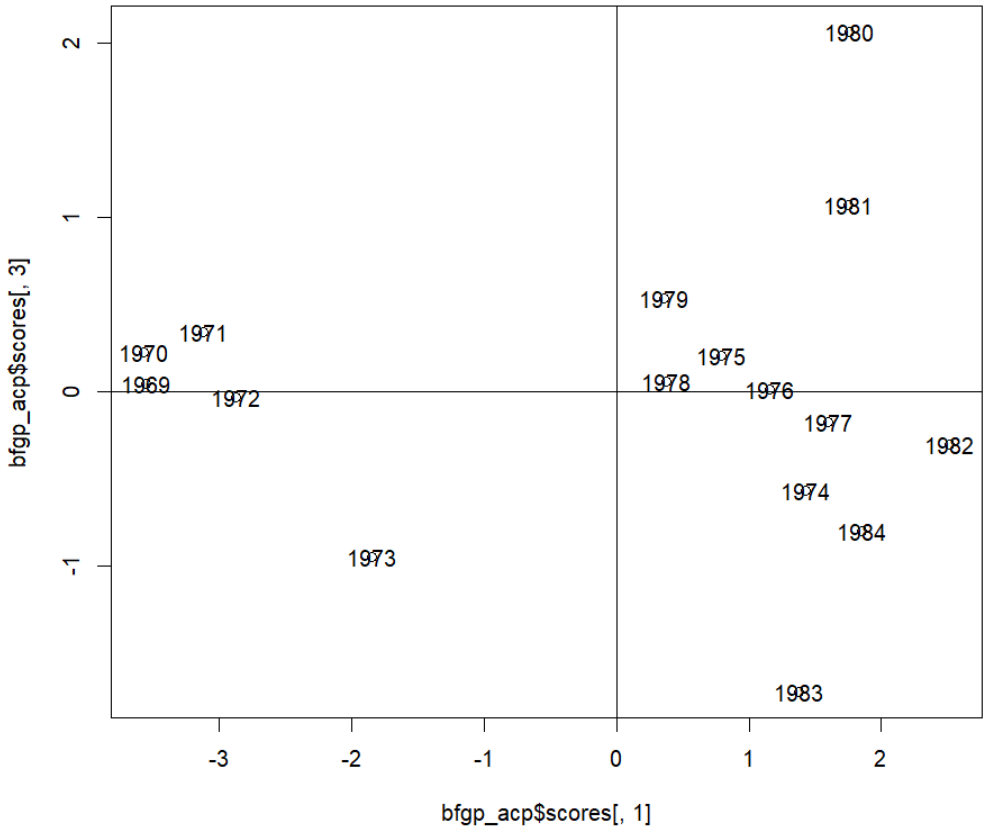
STUDY OF UNITS IN PRINCIPAL COMPONENTS

We can now produce 2D drawings associated to selected principal components.

```
> plot(bfgp_acp$scores[,1],bfgp_acp$scores[,2])
> text(bfgp_acp$scores[,1],bfgp_acp$scores[,2],labels=rownames(bfgp))
> abline(h=0,v=0)
```




```
> plot(bfgp_acp$scores[,1],bfgp_acp$scores[,3])
> text(bfgp_acp$scores[,1],bfgp_acp$scores[,3],labels=rownames(bfgp))
> abline(h=0,v=0)
```



```
>  plot(bfgp_acp$scores[,2],bfgp_acp$scores[,3])
> text(bfgp_acp$scores[,2],bfgp_acp$scores[,3],labels=rownames(bfgp))
>  abline(h=0,v=0)
```