Q1: hypothesis test.

H_0: $\Sigma = \Lambda\Lambda' + \Psi$ with a specified k, with Λ having 0 in specified position & Φ having unities in the diagonal

H_A: Σ is any positive definite matrix. (nested model)

under H_A, the log-likelihood is $L_A = -\frac{1}{2}n(\log|S| + q)$

under H_0, the log-likelihood is $L_0 = -\frac{1}{2}n[\log|\hat{\Sigma}| + tr(S\Sigma^{-1})]$

& the likelihood ratio is: (log ratio)

$$R_a = \frac{1}{2}n \cdot [\log|\hat{\Sigma}| - \log|S| + tr(S\hat{\Sigma}^{-1}) - p]$$

By Wiks's theorem

$-2 \cdot R_a \sim \chi^2(N)$, N is the number of restrictions on the free parameters imposed by the hypothesis (This is general conclusion of likelihood-ratio test)

$\Rightarrow$ $n \cdot [\log(|\hat{\Sigma}|) - (\log|S|) + tr(S\hat{\Sigma}^{-1}) - p] \sim \chi^2(N)$

to determine N:

We consider a othogonal rotation matrix M

under H_A: in Σ there are $\frac{p(p+1)}{2}$ elements (λ_{ij} as well as ψ_i in Ψ)

under H_0, there are p elements in Ψ and pk elements of Λ, but Λ can be replaced by $\Lambda \cdot M$, and these solutions are equivalent

M is a $k \cdot k$ matrix here, so it has $\frac{k \cdot (k-1)}{2}$ independent elements

which means in any solution, Λ can be made to satisfy $\frac{k \cdot (k-1)}{2}$ additional conditions

So, the number of all parameters minus the number of parameters specified in H_0 is N,

$$N = \frac{p(p+1)}{2} + \frac{k(k-1)}{2} - p - pk = \frac{p^2+p}{2} + \frac{k^2-k}{2} - \frac{2p+2pk}{2} = \frac{p^2-2pk+k^2-p-k}{2} = \frac{(p-k)^2}{2} - \frac{1}{2}(p+k)$$

$\Rightarrow$ when factors Num=3 in the problem

$$df = \frac{(6-3)^2}{2} - \frac{1}{2}(6+3) = 0$$

& the 4 factors specified will cause negative df, which is not allowed!

Q2:

$x_1 = \lambda_{11}f_1 + \lambda_{12}f_2 + \cdots + \lambda_{1k}f_k + u_1$

$x_2 = \lambda_{21}f_1 + \lambda_{22}f_2 + \cdots + \lambda_{2k}f_k + u_2$

$\vdots$

$x_q = \lambda_{q1}f_1 + \lambda_{q2}f_2 + \cdots + \lambda_{qk}f_k + u_q$

① $u_1, \cdots u_q$ indep ~~&~~ indep with f_i

② f_i has expectation 0 and variance 1

③ $f_1, f_2, \cdots f_k$ are uncorrelated

$\Rightarrow$

$var(x_i) = \sum_{j=1}^{k} \lambda_{ij}^2 \cdot var(f_k) + var(u_i) = \sum_{j=1}^{k} \lambda_{ij}^2 + \psi_i$ when $i \neq j$

$cov(x_i, x_j) = E(x_i - Ex_i)(x_j - Ex_j) = E(x_i \cdot x_j) = E(\sum_{m=1}^{k} \lambda_{im} f_m \cdot \sum_{n=1}^{k} \lambda_{jn} f_n)$

Because f_i are uncorrelated to each other
we have

$E(f_i f_j) = \begin{cases} 1 & \text{if } i=j \\ 0 & \text{if } i \neq j \end{cases} \Rightarrow cov(x_i, x_j) = \sum_{m=1}^{k} \lambda_{im}\lambda_{jm}$

$$\Rightarrow \Sigma_{(1)} = \begin{pmatrix} \sum_j \lambda_{1j}^2 & \sum \lambda_{1j}\lambda_{2j} & \cdots \\ \vdots & \sum_j \lambda_{2j}^2 & \\ & & \ddots \\ & & & \sum_j \lambda_{qj}^2 \end{pmatrix} = \begin{pmatrix} \lambda_{11} & \lambda_{12} & \cdots & \lambda_{1k} \\ \vdots & \vdots & & \\ \lambda_{q1} & \lambda_{q2} & \cdots & \lambda_{qk} \end{pmatrix} \begin{pmatrix} \lambda_{11} & \lambda_{21} & \cdots & \lambda_{q1} \\ \lambda_{12} & \lambda_{22} & \cdots & \lambda_{q2} \\ \lambda_{1k} & \lambda_{2k} & \cdots & \lambda_{qk} \end{pmatrix} = \Lambda \cdot \Lambda^T$$

$$\Sigma_{(2)} = \begin{pmatrix} \psi_1 & & 0 \\ & \psi_2 & \\ 0 & & \ddots \\ & & & \psi_q \end{pmatrix} \Rightarrow \Sigma = \Sigma_{(1)} + \Sigma_{(2)} = \Lambda\Lambda^T + \Psi$$

if factors are allowed to be correlated:
Assume the correlation matrix of factors is

$$\Sigma_F = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1k} \\ \sigma_{21} & \sigma_{22} & & \vdots \\ & & \ddots & \\ \sigma_{k1} & \sigma_{k2} & \cdots & \sigma_{kk} \end{pmatrix} \quad x = \Lambda f + \Psi \quad var(x) = var(\Lambda f + \Psi) = var(\Lambda f) + \Sigma_{(2)}$$

$var(\Lambda f) = \Lambda \cdot cov(f, f^T) \cdot \Lambda^T = \Lambda \cdot \Sigma_F \cdot \Lambda^T$

$\Rightarrow \Sigma^c = \Lambda \Sigma_F \Lambda^T + \Psi$

Q3: the communalities: $\sum_{j=1}^{k} \lambda_{ij}^2$ $\quad \Psi = \begin{pmatrix} u_1 & \\ & u_q \end{pmatrix}$

If $\Lambda^* = \Lambda M$

$x = \Lambda M \cdot f + u$

$var(x) = var(\Lambda M f + u) = var(\Lambda M f) + \Psi = \Lambda M \cdot var(f) \cdot M^T \Lambda^T + \Psi$

$cov(f, f^T) = var(f) = I$

$\Rightarrow var(x) = \Lambda M \cdot M^T \Lambda^T + \Psi$ if M is an orthogonal matrix $M \cdot M^T = I$

$\Rightarrow \Sigma = var(x) = \Lambda\Lambda^T + \Psi$ so the commualities are still $\sum_{j=1}^{k} \lambda_{ij}^2$

(Q4)

The result for male:

```
> mfact=factanal(mlife,factors=1,method="mle")
> mfact

Call:
factanal(x = mlife, factors = 1, method = "mle")

Uniquenesses:
   m0   m25   m50   m75 
0.279 0.005 0.148 0.655 

Loadings:
    Factor1
m0  0.849  
m25 0.998  
m50 0.923  
m75 0.587  

               Factor1
SS loadings      2.913
Proportion Var   0.728

Test of the hypothesis that 1 factor is sufficient.
The chi square statistic is 33 on 2 degrees of freedom.
The p-value is 6.83e-08 
```

The result for female:

```
> ffact

Call:
factanal(x = flife, factors = 1, method = "mle")

Uniquenesses:
   f0   f25   f50   f75 
0.220 0.005 0.115 0.526 

Loadings:
    Factor1
f0  0.883  
f25 0.998  
f50 0.941  
f75 0.689  

               Factor1
SS loadings      3.134
Proportion Var   0.784

Test of the hypothesis that 1 factor is sufficient.
The chi square statistic is 52.15 on 2 degrees of freedom.
The p-value is 4.74e-12 
```

The result shows that 1 factor for each sex is not sufficient to describe the data respectively. And more than 1 factor is not allowed because of insufficient degree of freedom related to Chi-square test statistic.

If we ignore the above problem which might imply the factor analysis questionable, we can conclude that the factors extracted from the data show that they are both dominated by life expectancy at age from 0 to 50.

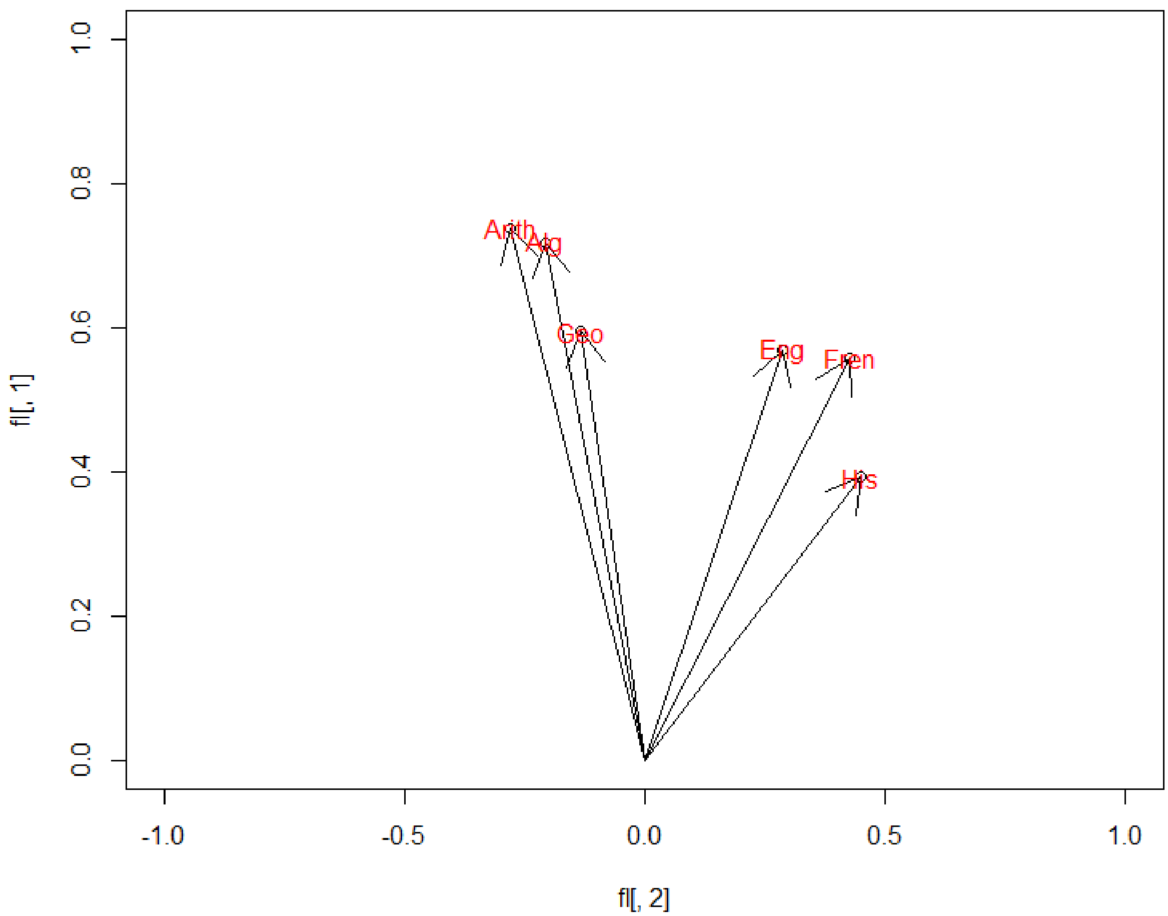
Q5

Without any rotation, the factor loadings are shown below:

```
> factanal(covmat=R,factors=2,method="mle",n.obs=220,rotation="none")

Call:
factanal(factors = 2, covmat = R, n.obs = 220, rotation = "none",     method
= "mle")

Uniquenesses:
[1] 0.508 0.595 0.644 0.377 0.440 0.628

Loadings:
     Factor1 Factor2
[1,]  0.558   0.425
[2,]  0.569   0.286
[3,]  0.392   0.450
[4,]  0.738  -0.279
[5,]  0.718  -0.209
[6,]  0.595  -0.133

               Factor1 Factor2
SS loadings      2.204   0.603
Proportion Var   0.367   0.101
Cumulative Var   0.367   0.468

Test of the hypothesis that 2 factors are sufficient.
The chi square statistic is 2.18 on 4 degrees of freedom.
The p-value is 0.703
```

Intuitively, we should rotate these two factor loadings so that they can labels as “verbal” and “math” respectively.

The loadings plot is shown below:



In this plot, we can find no indication of explicit interpretation of these two factors (as least not what we expected!).

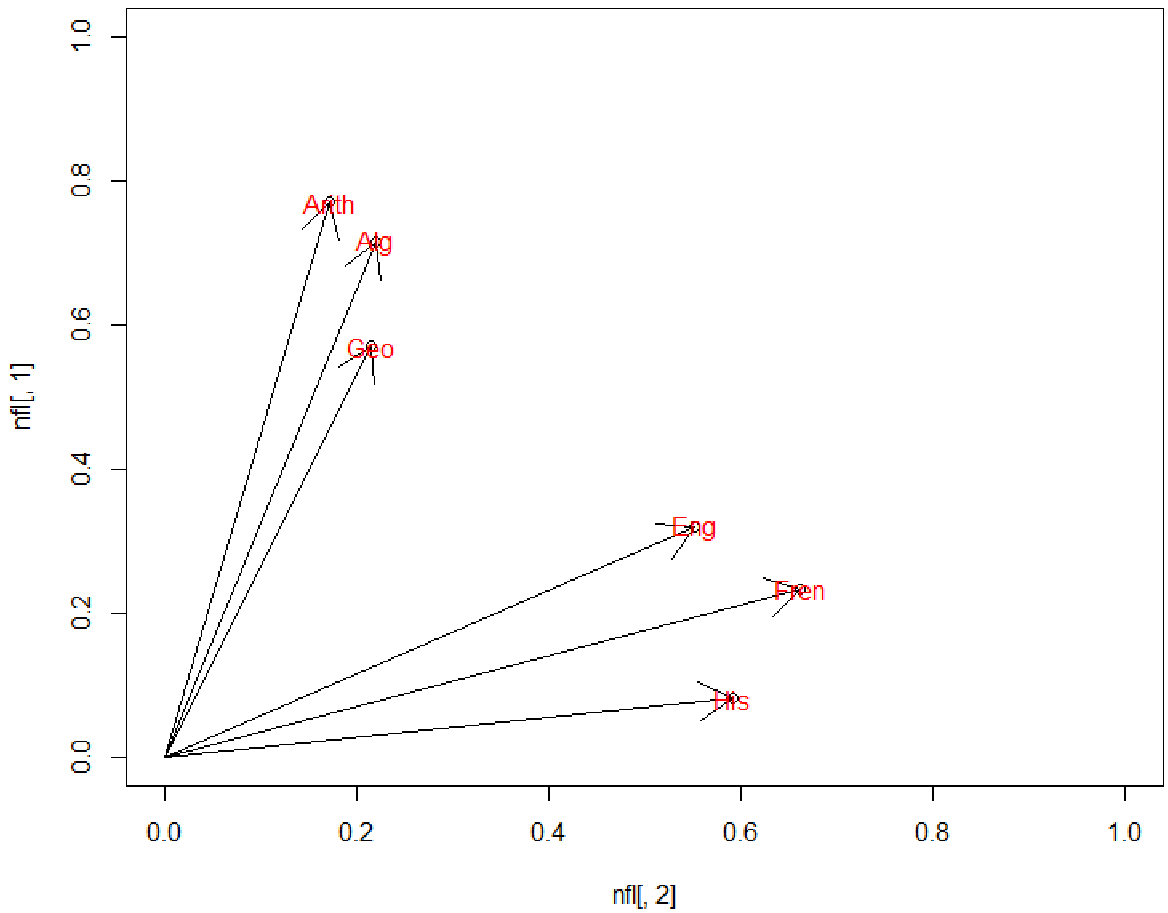
We need find a rotation matrix M to rotate the factor loadings matrix. Here, we choose to use "varimax" rotation. This matrix is:

```
> fvari$rotmat
           [,1]      [,2]
[1,]  0.8358396 0.5489738
[2,] -0.5489738 0.8358396
```

Rotate the factor loadings, we get the new factor loadings:

```
> fl%*%rot
          [,1]      [,2]
[1,] 0.2332771 0.6612268
[2,] 0.3187354 0.5512073
[3,] 0.0811948 0.5911398
[4,] 0.7702745 0.1715939
[5,] 0.7150775 0.2200083
[6,] 0.5704028 0.2151506
```

Then we plot new loading plot:



In this plot, “Arith”,”Alg”,”Geo” is mainly loaded on the first factor, and the other three on the second factor.

We can see from the plot that interpretation can be easier in the rotated factor loadings because the new factor loadings just meet our initial intuition, which is, two factors are dominated by “verbal” courses and “math” courses respectively.

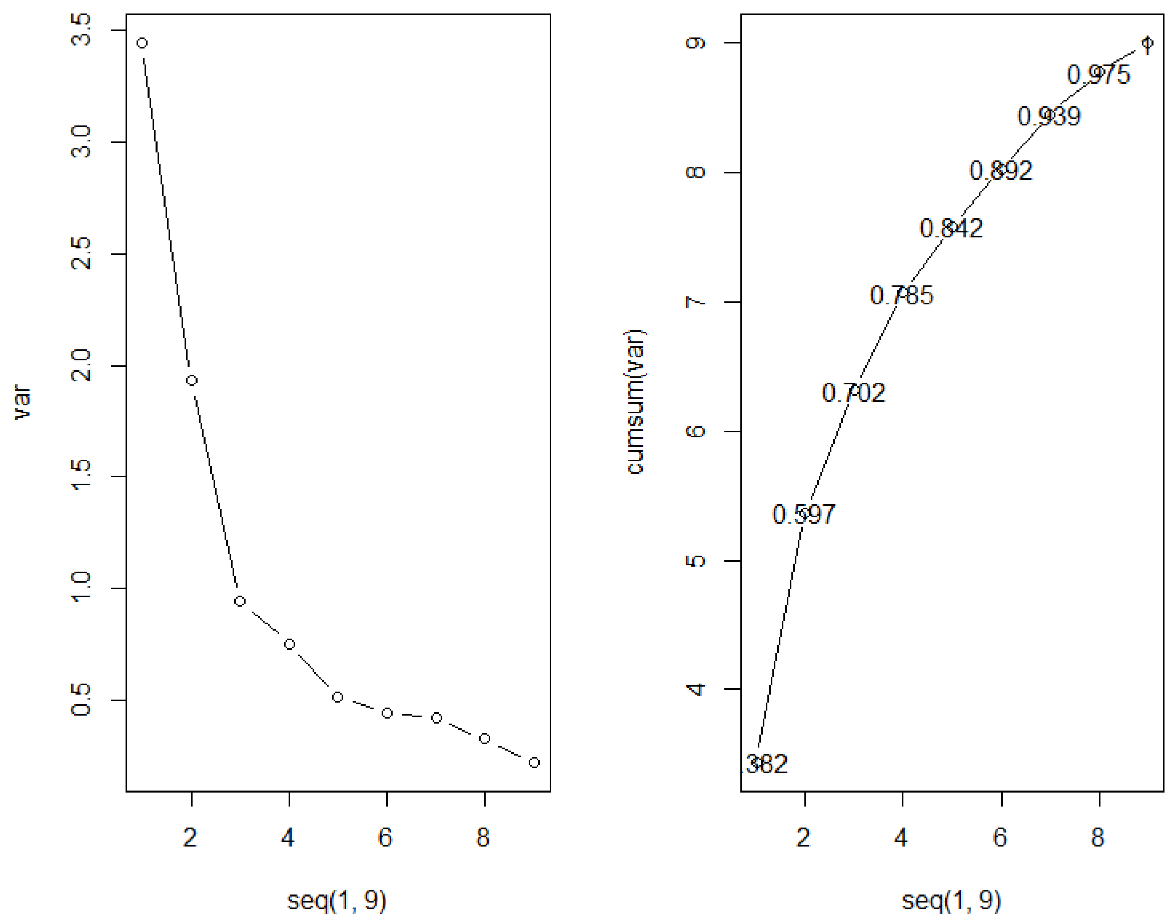
Q6

(a)

The result of PCA:

```
> summary(pr.out,loadings=T)
Importance of components:
                          Comp.1    Comp.2    Comp.3     Comp.4     Comp.5     Comp.6     Comp.7     Comp.8     Comp.9
Standard deviation     1.8551833 1.3891568 0.9707894 0.86823030 0.71739745 0.66647108 0.64981811 0.57103782 0.47443855
Proportion of Variance 0.3824117 0.2144174 0.1047147 0.08375821 0.05718435 0.04935374 0.04691818 0.03623158 0.02501022
Cumulative Proportion  0.3824117 0.5968291 0.7015437 0.78530194 0.84248629 0.89184003 0.93875821 0.97498978 1.00000000

Loadings:
      Comp.1 Comp.2 Comp.3 Comp.4 Comp.5 Comp.6 Comp.7 Comp.8 Comp.9
 [1,]  0.357 -0.268  0.420        -0.420  0.457 -0.108  0.407 -0.246
 [2,] -0.250 -0.394  0.399 -0.452  0.500        -0.230  0.184  0.285
 [3,]  0.374 -0.361         0.148               -0.400 -0.711  0.171
 [4,]  0.381 -0.280 -0.151  0.132  0.628  0.144  0.392        -0.409
 [5,] -0.225 -0.500  0.254  0.100 -0.311 -0.397  0.600
 [6,] -0.311 -0.408 -0.419 -0.103 -0.130 -0.199 -0.416        -0.561
 [7,] -0.339 -0.310 -0.410  0.217         0.647  0.122         0.369
 [8,]  0.404 -0.220 -0.336  0.216        -0.386 -0.121  0.513  0.448
 [9,]  0.316        -0.333 -0.801 -0.242         0.245 -0.117
>
```

The scree plot:



Five PC is sufficient to have a good summary of the data.

(b)

```
> sapply(c(1:5),function(f) factanal(covmat=R2,n.obs=123,factors=f)$PVAL)
   objective    objective    objective    objective    objective 
3.582705e-23 3.330276e-06 8.377428e-02 1.370264e-01 3.166230e-01 
```

We can find that 4 factors are sufficient and the result is shown below:

```
> r2f=factanal(covmat=R2,factors=4,method="mle",n.obs=123)
> r2f

Call:
factanal(factors = 4, covmat = R2, n.obs = 123, method = "mle")

Uniquenesses:
[1] 0.433 0.600 0.297 0.482 0.169 0.005 0.536 0.005 0.731

Loadings:
      Factor1 Factor2 Factor3 Factor4
 [1,]  0.707  -0.243                 
 [2,]          0.335   0.447  -0.296 
 [3,]  0.833                         
 [4,]  0.654  -0.126           0.265 
 [5,]          0.289   0.864         
 [6,] -0.108   0.966   0.224         
 [7,] -0.259   0.553   0.297         
 [8,]  0.643          -0.164   0.745 
 [9,]  0.421          -0.278   0.111 

               Factor1 Factor2 Factor3 Factor4
SS loadings      2.293   1.514   1.202   0.734
Proportion Var   0.255   0.168   0.134   0.082
Cumulative Var   0.255   0.423   0.557   0.638

Test of the hypothesis that 4 factors are sufficient.
The chi square statistic is 9.72 on 6 degrees of freedom.
The p-value is 0.137
```

How to interpret these factors?

The first factor is dominated by the 1st ,3rd, 4th and 8th, and we can find they are all “pain” and “doctor” relations.

The second factor is dominated by the 6th (maybe 7th included, too) statement, and we can find from these statements that these two is related to “pain” and “selfness”.

The third factor is dominated by the 5th (maybe 2nd included, too) statement, they are both “pain” and “inappropriateness”.

The fourth factor is dominated by 8th statement, and here a question raised.

As mentioned in the textbook, hypothesis test might be questionable, and it might just give us a clue on the upper boundary of the sufficient number of factors.

How about 3 factors?

We compare the fitted correlation matrix with the initial correlation matrix:

```
> round(tcrossprod(r2f$loadings)+diag(r2f$uniquenesses)-R2,3)
       [,1]   [,2]   [,3]   [,4]   [,5]  [,6]   [,7] [,8]   [,9]
 [1,]  0.000 -0.015 -0.007  0.040  0.000 0.000  0.010    0 -0.017
 [2,] -0.015  0.000  0.028 -0.039 -0.003 0.000  0.038    0 -0.009
 [3,] -0.007  0.028  0.000 -0.020  0.000 0.000  0.002    0  0.035
 [4,]  0.040 -0.039 -0.020  0.000  0.004 0.001 -0.080    0 -0.042
 [5,]  0.000 -0.003  0.000  0.004  0.000 0.940  0.001    0 -0.003
 [6,]  0.000  0.000  0.000  0.001  0.000 0.000  0.000    0  0.000
 [7,]  0.010  0.038  0.002 -0.080  0.001 0.000  0.000    0  0.069
 [8,]  0.000  0.000  0.000  0.000  0.000 0.000  0.000    0  0.000
 [9,] -0.017 -0.009  0.035 -0.042 -0.003 0.000  0.069    0  0.000
```

We can find that their difference is really small. So, we conclude that 3 factors are sufficient. And the final result is shown below:

```
> factanal(covmat=R2,factors=3,method="mle",n.obs=123,rotation="varimax")

Call:
factanal(factors = 3, covmat = R2, n.obs = 123, rotation = "varimax",     met
hod = "mle")

Uniquenesses:
[1] 0.404 0.518 0.336 0.455 0.499 0.171 0.496 0.239 0.754

Loadings:
      Factor1 Factor2 Factor3
 [1,]  0.649  -0.372   0.190
 [2,] -0.126   0.194   0.655
 [3,]  0.794  -0.144   0.116
 [4,]  0.725  -0.106
 [5,]          0.292   0.645
 [6,]          0.825   0.377
 [7,] -0.225   0.590   0.325
 [8,]  0.815          -0.304
 [9,]  0.437          -0.221

               Factor1 Factor2 Factor3
SS loadings      2.507   1.331   1.291
Proportion Var   0.279   0.148   0.143
Cumulative Var   0.279   0.426   0.570

Test of the hypothesis that 3 factors are sufficient.
The chi square statistic is 19.2 on 12 degrees of freedom.
The p-value is 0.0838
```

Code:

```
library(cluster.datasets)
data("life.expectancy.1971")
life=life.expectancy.1971
rm(life.expectancy.1971)
life=life[,-c(1,2)]
mlife=life[,1:4]
flife=life[,5:8]
sapply(1,function(f)
  factanal(mlife,factors=f,method="mle")$PVAL)
flife$f50=as.numeric(flife$f50)
ffact=factanal(flife,factors=1,method="mle")
mfact=factanal(mlife,factors=1,method="mle")
mfact
ffact
R=cbind(c(1,.44,.41,.29,.33,.25),
        c(.44,1,.35,.35,.32,.33),
        c(.41,.35,1,.16,.19,.18),
        c(.29,.35,.16,1,.59,.47),
        c(.33,.32,.19,.59,1,.46),
        c(.25,.33,.18,.47,.46,1)
)
#num of factors #
sapply(1:2,function(f) factanal(covmat=R,factors=f,method="mle",n.obs=220)$PVAL)
factanal(covmat=R,factors=2,method="mle",n.obs=220,rotation="varimax")
f=factanal(covmat=R,factors=2,method="mle",n.obs=220,rotation="none")
fvari=factanal(covmat=R,factors=2,method="mle",n.obs=220,rotation="varimax")
fl=cbind(f$loadings[,1],f$loadings[,2])
plot(fl[,1]~fl[,2],xlim=c(-1,1),ylim=c(-1,1))
for (i in 1:6){
```

```
arrows(-1,-1,fl[i,2],fl[i,1])
}
text(fl[,2],fl[,1],labels=c("Fren","Eng","His","Arith","Alg","Geo"),col="red")
rot=as.matrix(fvari$rotmat)
fl=as.matrix(fl)
nfl=fl%*%rot
plot(nfl[,1]~nfl[,2],xlim=c(0,1),ylim=c(0,1))
for (i in 1:6){
  arrows(0,0,nfl[i,2],nfl[i,1])
}
text(nfl[,2],nfl[,1],labels=c("Fren","Eng","His","Arith","Alg","Geo"),col="red")

R2=cbind(c(1,-0.04,0.61,0.45,0.03,-0.29,-0.3,0.45,0.3),
      c(-0.04,1,-0.07,-0.12,0.49,0.43,0.3,-0.31,-0.17),
      c(0.61,-0.07,1,0.59,0.03,-0.13,-0.24,0.59,0.32),
      c(0.45,-0.12,0.59,1,-0.08,-0.21,-0.19,0.63,0.37),
      c(0.03,0.49,0.03,-0.08,1,0.47,0.41,-0.14,-0.24),
      c(-0.29,0.43,-0.13,-0.21,-.47,1,0.63,-0.13,-0.15),
      c(-0.3,0.3,-0.24,-0.19,0.41,0.63,1,-0.26,-0.29),
      c(0.45,-0.31,0.59,0.63,-0.14,-0.13,-0.26,1,0.4),
      c(0.3,-0.17,0.32,0.37,-0.24,-0.15,-0.29,0.4,1)
)
#principle componet analysis#
pr.out=princomp(covmat=R2)
summary(pr.out,loadings=T)
var=(pr.out$sdev)^2
par(mfrow=c(1,2))
plot(var~seq(1,9),type="b")
plot(cumsum(var)~seq(1,9),type="b")
```

```
text(seq(1,9),cumsum(var),labels=round(cumsum(var)/sum(var),3))

sapply(c(1:5),function(f) factanal(covmat=R2,n.obs=123,factors=f)$PVAL)

r2f=factanal(covmat=R2,factors=4,method="mle",n.obs=123,rotation="varimax")

r2f

round(tcrossprod(r2f$loadings)+diag(r2f$uniquenesses)-R2,3)

factanal(covmat=R2,factors=3,method="mle",n.obs=123,rotation="varimax")
```