Problem 9.19

R Program

```
prob9.12 <- read.table("http://www.stat.ufl.edu/~winner/sta4702/data/wichern/T9-12.DAT",
   header=F,col.names=c("x1","x2","x3","x4","x5","x6","x7"))
attach(prob9.12)

X <- as.matrix(prob9.12)
(R.X <- cor(X))
n <- nrow(X); p <- ncol(X)

f2nr <- factanal(X,factors=2,rotation="none")
f2nr
f3nr <- factanal(X,factors=3,rotation="none")
f3nr
f2r <- factanal(X,factors=2)
f2r
f3r <- factanal(X,factors=3)
f3r

comm2r <- f2r$loadings[,1]^2 + f2r$loadings[,2]^2
specvar2r <- 1 - comm2r
cbind(comm2r, specvar2r)

comm3r <- f3r$loadings[,1]^2 + f3r$loadings[,2]^2 + f3r$loadings[,3]^2
specvar3r <- 1 - comm3r
cbind(comm3r, specvar3r)

pred2r <- f2r$loadings %*% t(f2r$loadings) +
    diag(f2r$uniquenesses)
m <- 2

TS.2r <- (n-1-((2*p+4*m+5)/6)) * log(det(pred2r) / det(R.X))
df.2r <- ((p-m)^2-p-m)/2
CV.2r <- qchisq(.95,df.2r)
PV.2r <- 1-pchisq(TS.2r,df.2r)

test_out.2r <- cbind(m,TS.2r,df.2r,CV.2r,PV.2r)
colnames(test_out.2r) <- c("m","Test Stat","df","X2(.05)","Pr>TS")
round(test_out.2r,4)

pre3r <- f3r$loadings %*% t(f3r$loadings) +
    diag(f3r$uniquenesses)
m <- 3

TS.3r <- (n-1-((2*p+4*m+5)/6)) * log(det(pre3r) / det(R.X))
df.3r <- ((p-m)^2-p-m)/2
CV.3r <- qchisq(.95,df.3r)
PV.3r <- 1-pchisq(TS.3r,df.3r)

test_out.3r <- cbind(m,TS.3r,df.3r,CV.3r,PV.3r)
colnames(test_out.3r) <- c("m","Test Stat","df","X2(.05)","Pr>TS")
round(test_out.3r,4)
```

Continued

```
xnew <- c(110,98,105,15,18,12,35)
xbar <- (1/n) * (t(X) %*% rep(1,n))
xsd <- sqrt(diag(cov(X)))
znew <- (xnew - xbar) / xsd

(fscore.2r.wls <- solve(t(f2r$loadings) %*% solve(diag(f2r$uniquenesses)) %*%
  f2r$loadings) %*% t(f2r$loadings) %*% solve(diag(f2r$uniquenesses)) %*%
  znew)

(fscore.3r.wls <- solve(t(f3r$loadings) %*% solve(diag(f3r$uniquenesses)) %*%
  f3r$loadings) %*% t(f3r$loadings) %*% solve(diag(f3r$uniquenesses)) %*%
  znew)

(fscore.2r.reg <- t(f2r$loadings) %*%
 solve((f2r$loadings %*% t(f2r$loadings)) + diag(f2r$uniquenesses)) %*% znew)

(fscore.3r.reg <- t(f3r$loadings) %*%
 solve((f3r$loadings %*% t(f3r$loadings)) + diag(f3r$uniquenesses)) %*% znew)
```

R Output

```
> (R.X <- cor(X))
          x1        x2        x3        x4        x5        x6        x7
x1 1.0000000 0.9260758 0.8840023 0.5720363 0.7080738 0.6744073 0.9273116
x2 0.9260758 1.0000000 0.8425232 0.5415080 0.7459097 0.4653880 0.9442960
x3 0.8840023 0.8425232 1.0000000 0.7003630 0.6374712 0.6410886 0.8525682
x4 0.5720363 0.5415080 0.7003630 1.0000000 0.5907360 0.1469074 0.4126395
x5 0.7080738 0.7459097 0.6374712 0.5907360 1.0000000 0.3859502 0.5745533
x6 0.6744073 0.4653880 0.6410886 0.1469074 0.3859502 1.0000000 0.5663721
x7 0.9273116 0.9442960 0.8525682 0.4126395 0.5745533 0.5663721 1.0000000
> n <- nrow(X); p <- ncol(X)
>
> f2nr <- factanal(X,factors=2,rotation="none")
> f2nr

Call:
factanal(x = X, factors = 2, rotation = "none")

Uniquenesses:
   x1    x2    x3    x4    x5    x6    x7
0.069 0.070 0.123 0.005 0.474 0.614 0.029

Loadings:
   Factor1 Factor2
x1  0.695   0.669
x2  0.669   0.695
x3  0.795   0.494
x4  0.983  -0.167
x5  0.655   0.312
x6  0.250   0.569
x7  0.558   0.812

                Factor1 Factor2
SS loadings       3.333   2.283
Proportion Var    0.476   0.326
Cumulative Var    0.476   0.802

Test of the hypothesis that 2 factors are sufficient.
The chi square statistic is 117.2 on 8 degrees of freedom.
The p-value is 1.25e-21
```

```
> f3nr
Uniquenesses:
   x1    x2    x3    x4    x5    x6    x7
0.039 0.034 0.088 0.005 0.447 0.005 0.038

Loadings:
   Factor1 Factor2 Factor3
x1  0.901   0.381
x2  0.775   0.600
x3  0.931   0.202
x4  0.733  -0.118   0.666
x5  0.689   0.225   0.169
x6  0.757  -0.132  -0.636
x7  0.762   0.608  -0.110
                Factor1 Factor2 Factor3
SS loadings       4.445   0.998   0.901
Proportion Var    0.635   0.143   0.129
Cumulative Var    0.635   0.778   0.906

Test of the hypothesis that 3 factors are sufficient.
The chi square statistic is 62.18 on 3 degrees of freedom.
The p-value is 2.01e-13
> f2r
Uniquenesses:
   x1    x2    x3    x4    x5    x6    x7
0.069 0.070 0.123 0.005 0.474 0.614 0.029
Loadings:
   Factor1 Factor2
x1 0.852   0.452
x2 0.868   0.419
x3 0.717   0.602
x4 0.148   0.987
x5 0.501   0.525
x6 0.619
x7 0.946   0.277
                Factor1 Factor2
SS loadings       3.545   2.071
Proportion Var    0.506   0.296
Cumulative Var    0.506   0.802

Test of the hypothesis that 2 factors are sufficient.
The chi square statistic is 117.2 on 8 degrees of freedom.
The p-value is 1.25e-21
> f3r
Uniquenesses:
   x1    x2    x3    x4    x5    x6    x7
0.039 0.034 0.088 0.005 0.447 0.005 0.038
Loadings:
   Factor1 Factor2 Factor3
x1 0.793   0.374   0.438
x2 0.911   0.317   0.185
x3 0.651   0.544   0.438
x4 0.255   0.964
x5 0.542   0.465   0.207
x6 0.299           0.950
x7 0.917   0.180   0.298
                Factor1 Factor2 Factor3
SS loadings       3.175   1.718   1.453
Proportion Var    0.454   0.245   0.208
Cumulative Var    0.454   0.699   0.906

Test of the hypothesis that 3 factors are sufficient.
The chi square statistic is 62.18 on 3 degrees of freedom.
The p-value is 2.01e-13
```

```
> cbind(comm2r, specvar2r)
      comm2r   specvar2r
x1 0.9308083 0.069191662
x2 0.9296182 0.070381824
x3 0.8766896 0.123310363
x4 0.9950121 0.004987889
x5 0.5264121 0.473587898
x6 0.3863622 0.613637835
x7 0.9711829 0.028817139

> cbind(comm3r, specvar3r)
      comm3r   specvar3r
x1 0.9614284 0.038571561
x2 0.9655192 0.034480799
x3 0.9118756 0.088124391
x4 0.9950434 0.004956640
x5 0.5533820 0.446617961
x6 0.9950317 0.004968311
x7 0.9624901 0.037509883

> round(test_out.2r,4)
     m Test Stat df X2(.05) Pr>TS
[1,] 2  117.3065  8 15.5073     0

> round(test_out.3r,4)
     m Test Stat df X2(.05) Pr>TS
[1,] 3   62.8371  3  7.8147     0

> (fscore.2r.wls <- solve(t(f2r$loadings) %*% solve(diag(f2r$uniquenesses)) %*%
+   f2r$loadings) %*% t(f2r$loadings) %*% solve(diag(f2r$uniquenesses)) %*%
+   znew)
             [,1]
Factor1 0.1118669
Factor2 0.9507034
>
> (fscore.3r.wls <- solve(t(f3r$loadings) %*% solve(diag(f3r$uniquenesses)) %*%
+   f3r$loadings) %*% t(f3r$loadings) %*% solve(diag(f3r$uniquenesses)) %*%
+   znew)
              [,1]
Factor1 -0.3532855
Factor2  1.0766309
Factor3  0.8058558
>
> (fscore.2r.reg <- t(f2r$loadings) %*%
+  solve((f2r$loadings %*% t(f2r$loadings)) + diag(f2r$uniquenesses)) %*% znew)
             [,1]
Factor1 0.1145316
Factor2 0.9454347
>
> (fscore.3r.reg <- t(f3r$loadings) %*%
+  solve((f3r$loadings %*% t(f3r$loadings)) + diag(f3r$uniquenesses)) %*% znew)
              [,1]
Factor1 -0.3286063
Factor2  1.0631969
Factor3  0.7927693
>
```

Problem 9.24

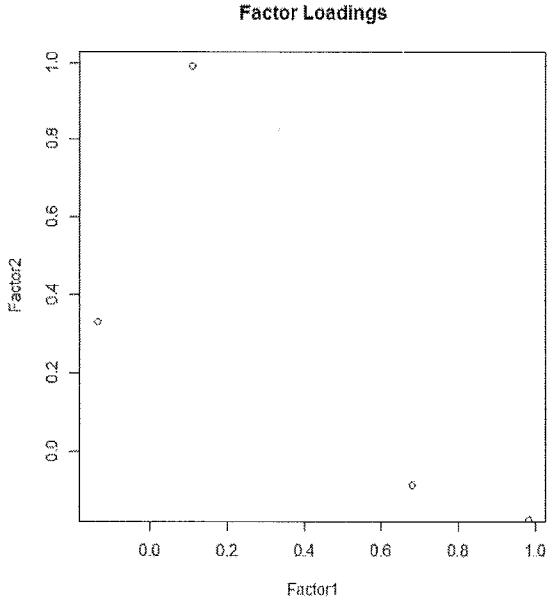
R Program

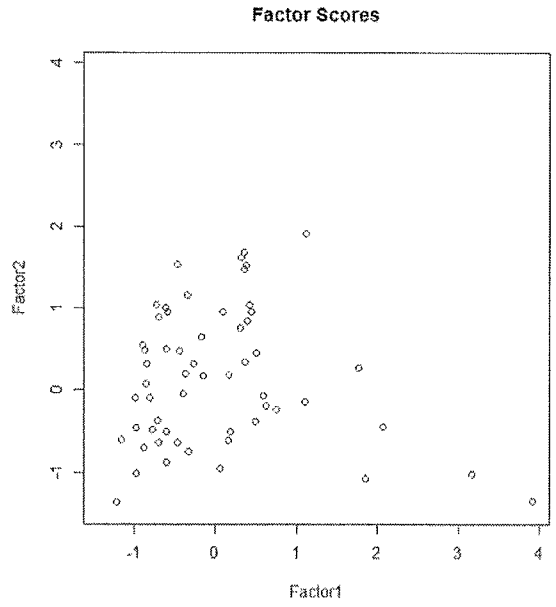
```
prob9.24 <-
read.table("http://www.stat.ufl.edu/~winner/sta4702/data/wichern/T8-
5.DAT",
  header=F,col.names=c("totpop","prodeg","empo16","govemp","medhv"))
attach(prob9.24)

X <- as.matrix(prob9.24)
(R <- cor(X))
n <- nrow(X); p <- ncol(X)

### Principal Components

(lambda <- eigen(R)$val)
(E <- eigen(R)$vec)

L <- matrix(rep(0,p^2),ncol=p)

for (i in 1:p) L[,i] <- sqrt(lambda[i]) * E[,i]

for (m in 1:2) {
psi <- diag(R) - diag(L[,1:m] %*% t(L[,1:m]))
print(cbind(m,L[,1:m],psi))
}

### Maximum Likelihood

ml.fa <- lapply(1:2, function(nf)
  factanal(X, factors=nf, scores="regression"))
ml.fa

par(pty="s")
plot(ml.fa[[2]]$loadings[,1],ml.fa[[2]]$loadings[,2],
  ylim=range(ml.fa[[2]]$loadings[,1]),xlab="Factor1",
ylab="Factor2",
  main="Factor Loadings")

ml.fa[[2]]$rotmat
(phi <- acos(ml.fa[[2]]$rotmat[1,1]))

par(pty="s")
plot(ml.fa[[2]]$scores[,1],ml.fa[[2]]$scores[,2],
  ylim=range(ml.fa[[2]]$scores[,1]),xlab="Factor1", ylab="Factor2",
   main="Factor Scores")
```

R Output

```
> (R <- cor(X))
            totpop     prodeg      empo16     govemp       medhv
totpop  1.00000000 -0.1922736  0.31321982 -0.1194831  0.02614869
prodeg -0.19227360  1.0000000 -0.06523680  0.3731722  0.68528795
empo16  0.31321982 -0.0652368  1.00000000 -0.4111161 -0.01034666
govemp -0.11948307  0.3731722 -0.41111605  1.0000000  0.17970100
medhv   0.02614869  0.6852879 -0.01034666  0.1797010  1.00000000
> n <- nrow(X); p <- ncol(X)
>
> ### Principal Components
>
> (lambda <- eigen(R)$val)
[1] 1.9919183 1.3675266 0.8641573 0.5350610 0.2413367
> (E <- eigen(R)$vec)
           [,1]       [,2]        [,3]       [,4]       [,5]
[1,]  0.2625829  0.4629936  0.78390268  0.2169291  0.2347882
[2,] -0.5933541  0.3256442 -0.16407255 -0.1446471  0.7028828
[3,]  0.3256978  0.6051419 -0.22487455 -0.6628689 -0.1943206
[4,] -0.4792022 -0.2524850  0.55070086 -0.5716730 -0.2766497
[5,] -0.4932213  0.4996473 -0.06882436  0.4072024 -0.5801162

+ print(cbind(m,L[,1:m],psi))
+ }
       m                   psi
totpop 1  0.3705973 0.8626576
prodeg 1 -0.8374323 0.2987071
empo16 1  0.4596747 0.7886992
govemp 1 -0.6763237 0.5425863
medhv  1 -0.6961095 0.5154315
       m                         psi
totpop 2  0.3705973  0.5414307 0.5695105
prodeg 2 -0.8374323  0.3808125 0.1536889
empo16 2  0.4596747  0.7076608 0.2879154
govemp 2 -0.6763237 -0.2952593 0.4554083
medhv  2 -0.6961095  0.5842940 0.1740320

> ml.fa
[[1]]

Uniquenesses:
totpop prodeg empo16 govemp  medhv
 0.963  0.005  0.996  0.860  0.528

Loadings:
       Factor1
totpop -0.192
prodeg  0.997
empo16
govemp  0.374
medhv   0.687

               Factor1
SS loadings      1.647
Proportion Var   0.329

Test of the hypothesis that 1 factor is sufficient.
The chi square statistic is 20.77 on 5 degrees of freedom.
The p-value is 0.000895
```

```
[[2]]

Call:
factanal(x = X, factors = nf, scores = "regression")

Uniquenesses:
totpop prodeg empo16 govemp  medhv
 0.872  0.005  0.005  0.710  0.527

Loadings:
       Factor1 Factor2
totpop -0.135   0.331
prodeg  0.982  -0.176
empo16  0.111   0.991
govemp  0.299  -0.448
medhv   0.682

               Factor1 Factor2
SS loadings      1.55   1.332
Proportion Var   0.31   0.266
Cumulative Var   0.31   0.576

Test of the hypothesis that 2 factors are sufficient.
The chi square statistic is 4.08 on 1 degree of freedom.
The p-value is 0.0434


> ml.fa[[2]]$rotmat
          [,1]       [,2]
[1,] 0.7927259 -0.6095782
[2,] 0.6095782  0.7927259

> (phi <- acos(ml.fa[[2]]$rotmat[1,1]))
[1] 0.6555284
```




Factor Scores
Factor2
Factor1
-1
0
1
2
3
4